Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2864904 C 2023/04/25

(11)(21) 2 864 904

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2012/07/11
(87) Date publication PCT/PCT Publication Date: 2013/08/22
(45) Date de délivrance/Issue Date: 2023/04/25
(85) Entrée phase nationale/National Entry: 2014/08/15
(86) N° demande PCT/PCT Application No.: US 2012/046326
(87) N° publication PCT/PCT Publication No.: 2013/122617
(30) Priorité/Priority: 2012/02/15 (US61/599,400)

(51) Cl.Int./Int.Cl. *C07K 19/00* (2006.01), *A61K 38/37* (2006.01), *A61P 7/04* (2006.01), *C07K 14/755* (2006.01), *C12N 15/62* (2006.01)

(72) Inventeurs/Inventors:
SCHELLENBERGER, VOLKER, US;
CHANG, PEI-YUN, US;
VARFAJ, FATBARDHA, US;
DING, SHENG, US;
SILVERMAN, JOSHUA, US;
WANG, CHIA-WEI, US;
...

(73) Propriétaire/Owner:
BIOVERATIV THERAPEUTICS INC., US

(74) Agent: GOWLING WLG (CANADA) LLP

(54) Titre : COMPOSITIONS DU FACTEUR VIII ET LEURS PROCEDES DE FABRICATION ET D'UTILISATION
(54) Title: FACTOR VIII COMPOSITIONS AND METHODS OF MAKING AND USING SAME

(57) Abrégé/Abstract:
The present invention relates to compositions comprising factor VIII coagulation factors linked to extended recombinant polypeptide (XTEN), isolated nucleic acids encoding the compositions and vectors and host cells containing the same, and methods of making and using such compositions in treatment of factor VIII-related diseases, disorders, and conditions.

CA 2864904 C 2023/04/25
(11)(21) 2 864 904
(13) C

(72) **Inventeurs(suite)/Inventors(continued):** SPINK, BENJAMIN, US; STEMMER, WILLEM P., US; GEETHING, NATHAN, US; KULMAN, JOHN, US; LIU, TONGYAO, US; TOBY, GARABET G., US; JIANG, HAIYAN, US; PETERS, ROBERT, US; WANG, DEPING, US; MEI, BAISONG, US

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 August 2013 (22.08.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/122617 A1

(51) International Patent Classification:
C07K 14/755 (2006.01) *A61K 38/37* (2006.01)

(21) International Application Number:
PCT/US2012/046326

(22) International Filing Date:
11 July 2012 (11.07.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/599,400 15 February 2012 (15.02.2012) US

(71) Applicant *(for all designated States except US)*: **AMUNIX OPERATING INC.** [US/US]; 500 Ellis Street, Mountain View, CA 94043-2206 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SCHELLENBERGER, Volker** [DE/US]; 914 Moreno Avenue, Palo Alto, CA 94303 (US). **CHANG, Pei-yun** [—/US]; 2395 Sharon Road, Menlo Park, CA 94025 (US). **VARFAJ, Fatbardha** [AL/US]; 777 W Middlefield Road, Apt. 140, Mountain View, CA 94043 (US). **DING, Sheng** [CN/US]; 36 Duane Street, Apt. 1, Redwood City, CA 94062 (US). **SILVERMAN, Joshua** [US/US]; 829 Dartshire Way, Sunnyvale, CA 94087 (US). **WANG, Chia-wei** [—/US]; 1449 David Lane, Milpitas, CA 95035 (US). **SPINK, Benjamin** [GB/US]; 379 Northam Avenue, San Carlos, CA 94070 (US). **STEMMER, Willem, P.** [NL/US]; 108 Kathy Court, Los Gatos, CA 95051 (US). **GEETHING, Nathan** [US/US]; 2009 Garzoni Place, Santa Clara, CA 95054 (US). **KULMAN, John** [US/US]; 88 Creeley Road, Belmont, MA 02478 (US). **LIU, Tongyao** [US/US]; 53 Buckman Drive, Lexington, MI 02421 (US). **TOBY, Garabet, G.** [US/US]; 9 School Hill Lane, North Reading, MA 01864 (US). **JIANG, Haiyan** [US/US]; 3 Bayberry Lane, Bilmont, MA 02478 (US). **PETERS, Robert** [US/US]; 53 Parkinson Street, Needham, MA 02492 (US). **WANG, Deping** [US/US]; 52 Cannon Ball Road, Sharon, MA 02067 (US). **MEI, Baisong** [US/US]; 353 Woodward Street, Waban, MA 02468 (US).

(74) Agents: **ORBAI, Lucian** et al.; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, CA 94304-1050 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2013/122617 A1

(54) **Title**: FACTOR VIII COMPOSITIONS AND METHODS OF MAKING AND USING SAME

(57) **Abstract**: The present invention relates to compositions comprising factor VIII coagulation factors linked to extended recombinant polypeptide (XTEN), isolated nucleic acids encoding the compositions and vectors and host cells containing the same, and methods of making and using such compositions in treatment of factor Vlll-related diseases, disorders, and conditions.

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 180

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 180

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHIER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

FACTOR VIII COMPOSITIONS AND METHODS OF MAKING AND USING SAME

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web. Said ASCII copy, created on July 11, 2012, is named 32887346.txt and is 13,344,768 Bytes in size.

BACKGROUND OF THE INVENTION

[0003] Factor VIII is an important component of the intrinsic pathway of the blood coagulation cascade. In the circulation, factor VIII is mainly complexed to von Willebrand factor. Upon activation by thrombin, (Factor IIa), it dissociates from the complex to interact with factor IXa in the intrinsic coagulation cascade, which, in turn, activates factor X. Once removed from the von Willebrand factor complex, activated factor VIII is proteolytically inactivated by activated Protein C (APC), factor Xa, and factor IXa, and is quickly cleared from the blood stream. When complexed with normal von Willebrand factor protein, the half-life of factor VIII is approximately 12 hours, whereas in the absence of von Willebrand factor, the half-life of factor VIII is reduced to 2 hours (Tuddenham EG, et al., Br J Haematol. (1982) 52(2):259-267).

[0004] In hemophilia, the clotting of blood is disturbed by a lack of certain plasma blood clotting factors. Hemophilia A is a deficiency of factor VIII, and is a recessive sex-linked, X chromosome disorder that represents 80% of hemophilia cases. The standard of care for the management of hemophilia A is replacement therapy with recombinant factor VIII concentrates. Subjects with severe hemophilia A have circulating procoagulant factor VIII levels below 1-2% of normal , and are generally on prophylactic therapy with the aim of keeping factor VIII above 1% between doses, which can usually be achieved by giving factor VIII two to three times a week. Persons with moderately severe hemophilia (factor VIII levels of 2-5% of normal) constitute 25-30% hemophilia incidents and manifest bleeding after minor trauma. Persons with mild hemophilia A (factor VIII levels of 5-40% of normal) comprise 15-20% of all hemophilia incidents, and develop bleeding only after significant trauma or surgery.

[0005] The *in vivo* activity of exogenously supplied factor VIII is limited both by a short protein half-life and inhibitors that bind to the factor VIII and diminish or destroy hemostatic function.

[0006] Up to 30% of hemophilia A patients receiving exogenously-supplied factor VIII mount an IgG immune response towards factor VIII (Towfighi, F., et al. Comparative measurement of anti-factor VIII

antibody by Bethesda assay and ELISA reveals restricted isotype profile and epitope specificity. Acta Haematol (2005) 114:84-90), which can result in the complete inhibition of its procoagulant activity and/or promote more rapid clearance of the factor VIII (Briët E et al. High titer inhibitors in severe haemophilia A. A meta-analysis based on eight long-term follow-up studies concerning inhibitors associated with crude or intermediate purity factor VIII products. Throm. Haemost. (1994) 72: 162-164). The IgG antibodies, called FVIII inhibitors, are primarily directed towards the A2, A3 and C2 domains (Scandella D et al. Localization of epitopes for human factor VIII inhibitor antibodies by immunoblotting and antibody neutralization. Blood (1989) 74:1618-1626), but can arise against the A1, B and C1 domains, as well. As such, treatment options for patients with FVIII inhibitors are limited.

[0007] Large proteins such as factor VIII are normally given intravenously so that the medicament is directly available in the blood stream. It has been previously demonstrated that an unmodified factor VIII injected intramuscularly yielded a maximum circulating level of only 1.4% of the normal plasma level (Pool et al, Ineffectiveness of Intramuscularly Injected Factor VIII Concentrate in Two Hemophilic Patients. New England J. Medicine (1966) 275(10):547-548). Formulations that could be administered other than by the intravenous route would greatly simplify their use, increase safety, and result in substantial cost savings.

[0008] Chemical modifications to a therapeutic protein can modify its in vivo clearance rate and subsequent serum half life. One example of a common modification is the addition of a polyethylene glycol (PEG) moiety, typically coupled to the protein via an aldehyde or N-hydroxysuccinimide (NHS) group on the PEG reacting with an amine group (e.g. lysine side chain or the N-terminus). However, the conjugation step can result in the formation of heterogeneous product mixtures that require extraction, purification and/or other further processes, all of which inevitably affect product yield and quality control. Also, the pharmacologic function of coagulation factors may be hampered if amino acid side chains in the vicinity of its binding site become modified by the PEGylation process. Other approaches include the genetic fusion of an Fc domain to the therapeutic protein, which increases the size of the therapeutic protein, hence reducing the rate of clearance through the kidney. In some cases, the Fc domain confers the ability to bind to, and be recycled from lysosomes by the FcRn receptor, resulting in increased pharmacokinetic half-life. Unfortunately, the Fc domain does not fold efficiently during recombinant expression, and tends to form insoluble precipitates known as inclusion bodies. These inclusion bodies must be solubilized and functional protein must be renatured from the misfolded aggregate, which is a time-consuming, inefficient, and expensive process.

SUMMARY OF THE INVENTION

[0009] The present invention relates to novel coagulation factor VIII fusion protein compositions and the uses thereof. Specifically, the compositions provided herein are particularly used for the treatment or improvement of a condition associated with hemophilia A, deficiencies of factor VIII, bleeding disorders and coagulopathies. In one aspect, the present invention provides compositions of isolated fusion proteins comprising a factor VIII (FVIII) and one or more extended recombinant polypeptides (XTEN)

wherein the fusion protein exhibits procoagulant activity. A subject XTEN useful for constructing such fusion proteins is typically a polypeptide with a non-repetitive sequence and unstructured conformation. In one embodiment, one or more XTEN is linked to a coagulation factor FVIII ("CF") selected from native human factor VIII, factor VIII B-domain deleted sequences ("FVIII BDD"), and sequence variants thereof (all the foregoing collectively "FVIII" or "CF"), resulting in a recombinant factor VIII-XTEN fusion protein ("CFXTEN"). The factor VIII polypeptide component of the CFXTEN comprises an A1 domain, an A2 domain, a C1 domain, a C2 domain, and optionally a B domain or a portion thereof. In some embodiments, the FVIII is further characterized by delineation of the aforementioned domains to comprise an acidic a1, a2 and a3 spacer. In another embodiment, the present disclosure is directed to pharmaceutical compositions comprising the fusion proteins and the uses thereof in methods and regimens for treating factor VIII-related conditions. The CFXTEN compositions have enhanced pharmacokinetic and pharmacologic properties compared to FVIII not linked to XTEN, which may permit more convenient dosing and improved efficacy.

[0010] In a first aspect, the invention relates to recombinant factor VIII fusion proteins comprising a factor VIII polypeptide and one or more extended recombinant polypeptide (XTEN) linked to the factor VIII. In some embodiments, the invention provides recombinant factor VIII fusion proteins comprising a factor VIII polypeptide and at least one extended recombinant polypeptide (XTEN), wherein said factor VIII polypeptide comprises an A1 domain including an a1 acidic spacer region, an A2 domain including an a2 acidic spacer region, an A3 domain including an a3 acidic spacer region, C1 domain, C2 domain and optionally all or a portion of B domain, and wherein said at least one XTEN is linked to said factor VIII polypeptide at (i) the C-terminus of said factor VIII polypeptide; (ii) within B domain of said factor VIII polypeptide if all or a portion of B domain is present; (iii) within the A1 domain of said factor VIII polypeptide; (iv) within the A2 domain of said factor VIII polypeptide; (v) within the A3 domain of said factor VIII polypeptide; (vi) within the C1 domain of said factor VIII polypeptide; (vii) within the C2 domain of said factor VIII polypeptide; (viii) at the N-terminus of said factor VIII polypeptide, or (ix) between two domains of said factor VIII polypeptide, wherein the fusion protein retains at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 400%, or 500% of the procoagulant activity, when measured by an in vitro coagulation assay, compared to a corresponding factor VIII not linked to XTEN. In one embodiment, in the foregoing recombinant factor VIII fusion protein the at least one XTEN is linked to said factor VIII polypeptide at a site at or within 1 to 6 amino acids of a site selected from Table 5, Table 6, Table 7, Table 8, and Table 9. In other embodiments, the invention provides recombinant factor VIII fusion proteins comprising a factor VIII polypeptide and at least a first extended recombinant polypeptide (XTEN), wherein said factor VIII polypeptide comprises an A1 domain including an a1 acidic spacer region, an A2 domain including an a2 acidic spacer region, an A3 domain including an a3 acidic spacer region, a C1 domain, a C2 domain and optionally all or a portion of a B domain, and wherein said first XTEN is linked to said factor VIII polypeptide at (i) the C-terminus of said factor VIII polypeptide; (ii) within the B domain of said factor VIII polypeptide if all or a portion of the B domain is present; (iii) within the A1 domain of said factor VIII polypeptide; (iv)

within the A2 domain of said factor VIII polypeptide; (v) within the A3 domain of said factor VIII polypeptide; (vi) within the C1 domain of said factor VIII polypeptide; or (vii) within the C2 domain of said factor VIII polypeptide; and when compared to a corresponding factor VIII protein not linked to XTEN, the fusion protein (a) retains at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 100%, 200%, 300%, 400%, or 500% of the procoagulant activity in an in vitro coagulation assay described herein or other such assays known in the art, and/or (b) exhibits reduced binding to an anti-factor VIII antibody in an *in vitro* binding assay described herein or other such assays known in the art. I n one embodiment, in the foregoing recombinant factor VIII fusion protein the at least one XTEN is linked to said factor VIII polypeptide at a site at or within 1 to 6 amino acids of a site selected from Table 5, Table 6, Table 7, Table 8, and Table 9. In other embodiments, the invention provides recombinant factor VIII fusion proteins comprising a factor VIII polypeptide and at least a first extended recombinant polypeptide (XTEN), wherein said factor VIII polypeptide comprises an A1 domain including an a1 acidic spacer region, an A2 domain including an a2 acidic spacer region, an A3 domain including an a3 acidic spacer region, a C1 domain, a C2 domain and optionally all or a portion of a B domain, and wherein said first XTEN is linked to said factor VIII polypeptide at an insertion site selected from Table 6 and Table 7 and wherein the fusion protein retains at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%,100%, 200%, 300%, 400%, or 500% of the procoagulant activity, when measured by an *in vitro* coagulation assay described herein or other such assays known in the art, compared to a corresponding factor VIII protein not linked to XTEN. Non-limiting examples of the factor VIII protein not linked to XTEN includes native FVIII, BDD FVIII, pBC100 and sequences from Table 1. In another embodiment of the recombinant factor VIII fusion protein, the factor VIII polypeptide has at least about 80% sequence identity, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% sequence identity to a sequence selected from the group consisting of the sequences of Table 1, the sequence depicted in FIG. 3, and the sequence depicted in FIG. 4, when optimally aligned. In yet another embodiment, the fusion protein comprises at least another XTEN linked to said factor VIII polypeptide at the C-terminus of said factor VIII polypeptideor within or optionally replacing the B domain of said factor VIII polypeptide. In a specific embodiment, the fusion protein comprises at least one XTEN sequence located within or optionally replacing the B domain of said factor VIII polypeptide. In another specific embodiment, the fusion protein comprises at least one XTEN sequence linked to said factor VIII polypeptide at the C-terminus of said factor VIII polypeptide. In one embodiment, the recombinant factor VIII fusion protein comprises a B-domain deleted variant of human factor VIII, wherein the B-domain deletion starts from a first position at about amino acid residue number 741 to about 750 and ending at a second position at amino acid residue number 1635 to about 1648 with reference to full-length human factor VIII sequence as set forth in FIG. 3. In another embodiment, the recombinant factor VIII fusion protein comprises a first XTEN sequence linked to said factor VIII polypeptide at the C-terminus of said factor VIII polypeptide, and at least a second XTEN within or replacing the B domain of said factor VIII polypeptide, wherein the second XTEN is linked to the C-terminal end of about amino acid residue

number 741 to about 750 and to the N-terminal end of amino acid residue numbers 1635 to about 1648 with reference to full-length human factor VIII sequence as set forth in FIG. 3, wherein the cumulative length of the XTEN is at least about 100 amino acid residues. In one embodiment, in the foregoing fusion protein, the second XTEN links the factor VIII amino acids between N745 to P1640 or between S743 to Q1638 or between P747 to V1642 or between N745 and Q1656 or between N745 and S1657 or between N745 and T1667 or between N745 and Q1686 or between R747 and V1642 or between T751 and T1667. In one embodiment, the recombinant factor VIII fusion protein comprises a sequence having at least about 80% sequence identity, or at least about 90%, or at least about 91%, or at least about 92%, or at least about 93%, or at least about 94%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99%, to about 100% sequence identity compared to a sequence of comparable length selected from Table 21, when optimally aligned. In another embodiment, the recombinant factor VIII fusion protein comprises at least a second XTEN, optionally a third XTEN, optionally a fourth XTEN, optionally a fifth XTEN and optionally a sixth XTEN, wherein each of the second, third, fourth, fifth, or sixth XTEN is linked to said factor VIII polypeptide at a second, third, fourth, fifth, or sixth site selected from the group consisting of an insertion site from Table 5, Table 6, Table 7 Table 8, and Table 9; a location within 6 amino acids of amino acid residue 32, 220, 224, 336, 339, 390, 399, 416, 603, 1656, 1711, 1725, 1905 and 1910 of mature factor VIII; a location between any two adjacent domains of said factor VIII polypeptide, wherein said two adjacent domains are selected from the group consisting of A1 and A2 domains, A2 and B domains, B and A3 domains, A3 and C1 domains, and C1 and C2 domains; a location within the B domain of said factor VIII polypeptide, wherein the second XTEN is linked to the C-terminal end of about amino acid residue number 741 to about 750 and to the N-terminal end of amino acid residue numbers 1635 to about 1648 of a native factor VIII sequence; and the C-terminus of said factor VIII polypeptide. In one embodiment, the first XTEN is separated from the second XTEN by at least 10 amino acids, at least 50 amino acids, at least 100 amino acids, at least 200 amino acids, at least 300 amino acids, or at least 400 amino acids. In one embodiment of the recombinant factor VIII fusion protein that comprises at least a second XTEN, optionally a third XTEN, optionally a fourth XTEN, optionally a fifth XTEN and optionally a sixth XTEN, each XTEN has at least about 80% sequence identity, or at least about 90%, or at least about 91%, or at least about 92%, or at least about 93%, or at least about 94%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99%, or about 100% sequence identity compared to an XTEN of comparable length selected from the group consisting of the sequences in Table 4, Table 13, Table 14, Table 15, Table 16, and Table 17, when optimally aligned. In yet another embodiment of the recombinant factor VIII fusion protein that comprises at least a second XTEN, optionally a third XTEN, optionally a fourth XTEN, optionally a fifth XTEN and optionally a sixth XTEN, In preferred embodiments, the recombinant factor VIII fusion protein exhibits a terminal half-life at least about 3 hours, or 4 hours, or 6 hours, or 12 hours, or 13 hours, or 14 hours, or 16 hours, or 24 hours, or 48 hours, or 72 hours, or 96 hours, or 120 hours, or 144 hours, or 7 days, or 14 days, or 21 days when administered to a subject, wherein said subject is selected from human and factor VIII/von Willebrand factor double

knock-out mouse. Further, in the embodiments of this paragraph, the fusion protein exhibits reduced binding to anti-factor VIII antibody or greater retained procoagulant activity, or both as compared to a corresponding factor VIII not linked to XTEN. In one embodiment, the procoagulant activity of the recombinant factor VIII fusion protein is at least 30%, or 40%, 50%, 80%, 100%, 200%, 300%, 400%, or 500% greater procoagulant activity in the presence of the anti-FVIII antibody compared to a corresponding factor VIII not linked to XTEN when each are assayed by an in vitro coagulation assay. In one embodiment, the reduced binding of the fusion protein to anti-factor VIII antibody is determined using a Bethesda assay using anti-factor VIII antibody selected from the group consisting of the antibodies of Table 10 and polyclonal antibody from a hemophilia A patient with factor VIII inhibitors, wherein the reduced binding and retained procoagulant activity of the fusion protein is evidenced by a lower Bethesda titer of at least about 2, 4, 6, 8, 10, 12, 15, 20, 30, 40, 50, 60, 70, 80, 100, or 200 Bethesda units for the fusion protein compared to that for the factor VIII not linked to XTEN.

[0011] In one embodiment, the recombinant factor VIII fusion protein can, for example, comprise one or more XTEN wherein the XTEN has at least about 80% sequence identity, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% sequence identity compared to one or more XTEN of comparable length selected from Table 4, Table 13, Table 14, Table 15, Table 16, and Table 17, when optimally aligned.

[0012] In another aspect, the invention relates to recombinant factor VIII fusion proteins comprising FVIII and one or more XTEN in specific N- to C-terminus configurations. In one embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula I:

$$(XTEN)_x\text{-}CF\text{-}(XTEN)_y \qquad I$$

wherein independently for each occurrence, CF is a factor VIII as defined herein, including sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity with sequenced from Table 1; x is either 0 or 1 and y is either 0 or 1 wherein $x+y \geq 1$; and XTEN is an extended recombinant polypeptide as described herein, including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4. Accordingly, the CFXTEN fusion composition can have XTEN-CF, XTEN-CF-XTEN, or CF-XTEN configurations.

[0013] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula II:

$$(XTEN)_x\text{-}(S)_x\text{-}(CF)\text{-}(XTEN)_y \qquad II$$

wherein independently for each occurrence, CF is a factor VIII as defined herein, including sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 1; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; x is either 0 or 1 and y is

either 0 or 1 wherein x+y ≥1; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4.

[0014] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein, wherein the fusion protein is of formula III:

$(XTEN)_x$-$(S)_x$-(CF)-$(S)_y$-$(XTEN)_y$ III

wherein independently for each occurrence, CF is a factor VIII as defined herein, including sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequence set for in Table 1; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; x is either 0 or 1 and y is either 0 or 1 wherein x+y ≥1; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4.

[0015] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula IV:

(A1)-$(XTEN)_u$-(A2)-$(XTEN)_v$-(B)-$(XTEN)_w$-(A3)-$(XTEN)_x$-(C1)-$(XTEN)_y$-(C2)-$(XTEN)_z$ IV

wherein independently for each occurrence, A1 is an A1 domain of FVIII; A2 is an A2 domain of FVIII; A3 is an A3 domain of FVIII; B is a B domain of FVIII which can be a fragment or a splice variant of the B domain; C1 is a C1 domain of FVIII; C2 is a C2 domain of FVIII; v is either 0 or 1; w is either 0 or 1; x is either 0 or 1; y is either 0 or 1; y is either 0 or 1 with the proviso that u + v + x + y+z ≥1; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4.

[0016] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula V:

$(XTEN)_t$-$(S)_a$ -(A1)-$(S)_b$-$(XTEN)_u$-$(S)_b$-(A2)-$(S)_c$-$(XTEN)_v$-$(S)_c$-(B)-$(S)_d$-$(XTEN)_w$-$(S)_d$-(A3)-$(S)_e$-$(XTEN)_x$-$(S)_e$-(C1)-$(S)_f$-$(XTEN)_y$-$(S)_f$-(C2)-$(S)_g$-$(XTEN)_z$ V

wherein independently for each occurrence, A1 is an A1 domain of FVIII; A2 is an A2 domain of FVIII; A3 is an A3 domain of FVIII; B is a B domain of FVIII which can be a fragment or a splice variant of the B domain; C1 is a C1 domain of FVIII; C2 is a C2 domain of FVIII; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; a is either 0 or 1; b is either 0 or 1; c is either 0 or 1; d is either 0 or 1; e is either 0 or 1; f is either 0 or 1; g is either 0 or 1; t is either 0 or 1; u is either 0 or 1; v is either 0

or 1; w is 0 or 1, x is either 0 or 1; y is either 0 or 1; z is either 0 or 1 with the proviso that t + u + v + w+ x + y + z ≥1; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4. In another embodiment of formula V, the spacer sequence is glycine or a sequence selected from Tables 11 and 12.

[0017] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula VI:

$(XTEN)_u$-$(S)_a$-(A1)-$(S)_b$-$(XTEN)_v$-$(S)_b$-(A2)-$(S)_c$-$(XTEN)_w$-$(S)_c$-(A3)-$(S)_d$-$(XTEN)_x$-$(S)_d$-(C1)-$(S)_e$-$(XTEN)_y$-$(S)_e$-(C2)-$(S)_f$-$(XTEN)_z$ VI

wherein independently for each occurrence, A1 is an A1 domain of FVIII; A2 is an A2 domain of FVIII; A3 is an A3 domain of FVIII; C1 is a C1 domain of FVIII; C2 is a C2 domain of FVIII; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; a is either 0 or 1; b is either 0 or 1; c is either 0 or 1; d is either 0 or 1; e is either 0 or 1; f is either 0 or 1; u is either 0 or 1; v is either 0 or 1; w is 0 or 1, x is either 0 or 1; y is either 0 or 1; z is either 0 or 1 with the proviso that u + v + w+ x + y + z ≥1; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4. In another embodiment of formula V, the spacer sequence is glycine or a sequence selected from Tables 11 and 12.

[0018] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula VII:

(SP)-$(XTEN)_x$-$(CS)_x$-$(S)_x$-(FVIII_1-745)-$(S)_y$-$(XTEN)_y$-$(S)_y$-(FVIII_1640-2332)-$(S)_z$-$(CS)_z$-$(XTEN)_z$ VII

wherein independently for each occurrence, SP is a signal peptide, preferably with sequence MQIELSTCFFLCLLRFCFS (SEQ ID NO: 1611), CS is a cleavage sequence listed in Table 12, S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include amino acids compatible with restrictions sites, "FVIII_1-745" is residues 1-745 of Factor FVIII and "FVIII_1640-2332" is residues 1640-2332 of FVIII, x is either 0 or 1, y is either 0 or 1, and z is either 0 or 1, wherein x+y+z >2; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity sequences set forth in Table 4. In one embodiment of formula VII, the spacer sequence is GPEGPS (SEQ ID NO: 1612). In another embodiment of formula V, the spacer sequence is glycine or a sequence selected from Tables 11 and 12.

[0019] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula VIII:

$(A1)-(S)_a-(XTEN)_v-(S)_a-(A2)-(B1)-(S)_b-(XTEN)_w-(S)_b-(B2)-(A3)-(S)_c-(XTEN)_x-(S)_c-(C1)-(S)_d-(XTEN)_y-(S)_d-(C2)-(S)_e-(XTEN)_z$ VIII

wherein independently for each occurrence, A1 is an A1 domain of FVIII; A2 is an A2 domain of FVIII; B1 is a fragment of the B domain that can have from residue 741 to 743-750 of FVIII or alternatively from about residue 741 to about residues 745 of FVIII; B2 is a fragment of the B domain that can have from residues 1635-1686 to 1689 of FVIII or alternatively from about residue 1640 to about residues 1689 of FVIII; A3 is an A3 domain of FVIII; C1 is a C1 domain of FVIII; C2 is a C2 domain of FVIII; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; a is either 0 or 1; b is either 0 or 1; c is either 0 or 1; d is either 0 or 1; e is either 0 or 1; f is either 0 or 1; u is either 0 or 1; v is either 0 or 1; w is 0 or 1, x is either 0 or 1; y is either 0 or 1; z is either 0 or 1 with the proviso that $u + v + w + x + y + z \geq 1$; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4. In one embodiment of formula VIII, the spacer sequence is GPEGPS (SEQ ID NO: 1612). In another embodiment of formula V, the spacer sequence is glycine or a sequence selected from Tables 11 and 12.

[0020] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula IX:

$(A1_N)-(S)_a-(XTEN)_t-(S)_b-(A1_C)-(A2_N)-(S)_c-(XTEN)_u-(S)_d-(A2_C)-(B_N)-(S)_e-(XTEN)_v-(S)_f-(B_C)-(A3_N)-(S)_g-(XTEN)_w-(S)_h-(A3_C)-(C1_N)-(S)_i-(XTEN)_x-(S)_j-(C1_C)-(C2_N)-(S)_k-(XTEN)_y-(S)_l-(C2_C)-(S)_m-(XTEN)_z$

IX

wherein independently for each occurrence, $A1_N$ is a fragment of the A1 domain from at least residue number 1 (numbered relative to native, mature FVIII) to no more than residue number 371, $A1_c$ is a fragment of the A1 domain from at least residue number 2 to no more than residue number 372, with the priviso that no sequence of the $A1_N$ fragment is duplicated in the $A1_c$ is a fragment; $A2_N$ is a fragment of the A2 domain from at least residue number 373 to no more than residue number 739, $A2_c$ is a fragment of the A2 domain from at least residue number 374 to no more than residue number 740, with the priviso that no sequence of the $A2_N$ fragment is duplicated in the $A2_c$ is a fragment; B_N is a fragment of the B domain from at least residue number 741 to no more than residue number 1647, B_c is a fragment of the B domain from at least residue number 742 to no more than residue number 1648, with the priviso that no sequence of the B_N fragment is duplicated in the B_c is a fragment; $A3_N$ is a fragment of the A3 domain from at least residue number 1649 to no more than residue number 2019, $A3_c$ is a fragment of the A3 domain from at least residue number 1650 to no more than residue number 2019, with the priviso that no sequence of the $A3_N$ fragment is duplicated in the $A3_c$ is a fragment; $C1_N$ is a fragment of the C1 domain from at least residue number 2020 to no more than residue number 2171, $C1_c$ is a fragment of the C1 domain from at least residue number 2021 to no more than residue number 2172, with the priviso that no sequence of the $C1_N$ fragment is duplicated in the $C1_c$ is a fragment; $C2_N$ is a fragment of the C2 domain

from at least residue number 2173 to no more than residue number 2331, $C2_c$ is a fragment of the C2 domain from at least residue number 2174 to no more than residue number 2332, with the priviso that no sequence of the $C2_N$ fragment is duplicated in the $C2_c$ is a fragment; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; a is either 0 or 1; b is either 0 or 1; c is either 0 or 1; d is either 0 or 1; e is either 0 or 1; f is either 0 or 1; g is either 0 or 1; h is either 0 or 1; i is either 0 or 1; j is either 0 or 1; k is either 0 or 1; l is either 0 or 1; m is either 0 or 1; t is either 0 or 1; u is either 0 or 1; v is either 0 or 1; w is 0 or 1, x is either 0 or 1; y is either 0 or 1; z is either 0 or 1 with the proviso that t + u + v + w+ x + y + z $\geq$1; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80% sequence identity, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% sequence identity compared to one or more XTEN of comparable length selected from Table 4. In one embodiment of formula IX, the spacer sequence is GPEGPS (SEQ ID NO: 1612). In another embodiment of formula V, the spacer sequence is glycine or a sequence selected from Tables 11 and 12. In another embodiment of formula IX, Z is 1. In another embodiment of the fusion protein of formula IX V is 1 and the XTEN is linked to the C-terminal end of about amino acid residue number 741 to about 750 and to the N-terminal end of amino acid residue numbers 1635 to about 1648 with reference to full length human factor VIII sequence as set forth in FIG. 3. In another embodiment of the fusion protein of formula IX, the sum of t, u, v, w, x, y, and z equals 2, 3, 4, 5, or 6. In another embodiment of formula IX, the sum of t, u, v, w, x, y, and z equals 2, and v is 1 and z is 1. In another embodiment of the fusion protein of formula IX, the sum of t, u, v, w, x, y, and z equals 3, v and z each equal 1, and either t, u, w, x or y is 1. In another embodiment of formula IX, the sum of t, u, v, w, x, y, and z equals 4, v and w and z each equal 1, and two of t, u, x or y is 1. In another embodiment of the fusion protein of formula IX, the cumulative length of the XTENs is between about 84 to about 3000 amino acid residues. In another embodiment of formula IX, at least one XTEN is inserted immediately downstream of an amino acid which corresponds to an amino acid in mature native human factor VIII selected from the group consisting of amino acid residue number 32, 220, 224, 336, 339, 399, 416, 603, 1656, 1711, 1725, 1905 and 1910. In another embodiment of the fusion protein formula IX, each XTEN is linked to said fusion protein at sites selected from Table 5, Table 6, Table 7, Table 8, and Table 9. In another embodiment of the fusion protein formula IX, each XTEN has at least about 80%, or about 90%, or at least about 91%, or at least about 92%, or at least about 93%, or at least about 94%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99%, or about 100% sequence identity compared to an XTEN of comparable length selected from the group consisting of the sequences in Table 4, Table 13, Table 14, Table 15, Table 16, and Table 17, when optimally aligned.

[0021] In another embodiment of the CFXTEN composition, the invention provides a first recombinant factor VIII polypeptide of formula X:

(A1) – a1 – (A2) – a2 – [B] X

and a second polypeptide comprising Formula XI:

a3 – (A3) – (C1) - (C2) XI

wherein the first polypeptide and the second polypeptide are fused or exist as a heterodimer; wherein, A1 is an A1 domain of factor VIII; A2 is an A2 domain of factor VIII; [B] is a B domain of factor VIII, a fragment thereof, or is deleted; A3 is an A3 domain of factor VIII; C1 is a C1 domain of factor VIII; C2 is a C2 domain of factor VIII; a1, a2, and a3 are acidic spacer regions; wherein the A1 domain comprises an XTEN permissive loop-1 (A1-1) region and an XTEN permissive loop-2 (A1-2) region; wherein the A2 domain comprises an XTEN permissive loop-1 (A2-1) region and an XTEN permissive loop-2 (A2-2) region; wherein the A3 domain comprises an XTEN permissive loop-1 (A3-1) region and an XTEN permissive loop-2 (A3-2) region; wherein an XTEN sequence is inserted into at least one of the regions A1-1, A1-2, A2-1, A2-2, A3-1, or A3-2; and wherein the recombinant factor VIII protein exhibits procoagulant activity. In one embodiment of the heterodimer, the first polypeptide and the second polypeptide form a single polypeptide chain comprising the formula (A1) – a1 – (A2) – a2 – [B] – [a3] – (A3) – (C1) – (C2). In one embodiment of the foregoing, "fused" means a peptidic bond.

[0022] In another embodiment of the CFXTEN composition, the invention provides a first recombinant factor VIII polypeptide of formula X:

(A1) – a1 – (A2) – a2 – [B] X

and a second polypeptide comprising Formula XI:

a3 (A3) (C1) (C2) XI

wherein the first polypeptide and the second polypeptide are fused or exist as a heterodimer; wherein, A1 is an A1 domain of factor VIII; A2 is an A2 domain of factor VIII; [B] is a B domain of factor VIII, a fragment thereof, or is deleted; A3 is an A3 domain of factor VIII; C1 is a C1 domain of factor VIII; C2 is a C2 domain of factor VIII; a1, a2, and a3 are acidic spacer regions; wherein an XTEN sequence is inserted into a3; and wherein the recombinant factor VIII protein exhibits procoagulant activity. In one embodiment of the heterodimer, the first polypeptide and the second polypeptide form a single polypeptide chain comprising the formula (A1) – a1 – (A2) – a2 – [B] – [a3] – (A3) – (C1) – (C2). In one embodiment of the foregoing, "fused" means a peptidic bond.

[0023] In embodiments of the foregoing formulae X and XI polypeptides, the XTEN permissive loops are contained within surface-exposed, flexible loop structures, and wherein A1-1 is located between beta strand 1 and beta strand 2, A1-2 is located between beta strand 11 and beta strand 12, A2-1 is located between beta strand 22 and beta strand 23, A2-2 is located between beta strand 32 and beta strand 33, A3-1 is located between beta strand 38 and beta strand 39 and A3-2 is located between beta strand 45 and beta strand 46, according to the secondary structure of mature factor VIII stored as Accession Number 2R7E of the DSSP database. In other embodiments of the foregoing formulae X and XI polypeptides, the surface-exposed, flexible loop structure comprising A1-1 corresponds to a region in native mature human factor VIII from about amino acid 15 to about amino acid 45. In other embodiments of the foregoing formulae X and XI polypeptides the A1-1 corresponds to a region in native mature human factor VIII from about amino acid 18 to about amino acid 41. In other embodiments of the foregoing formulae X and XI polypeptides, the surface-exposed, flexible loop structure comprising A1-2 corresponds to a region in

native mature human factor VIII from about amino acid 201 to about amino acid 232. In other embodiments of the foregoing formulae X and XI polypeptides the A1-2 corresponds to a region in native mature human factor VIII from about amino acid 218 to about amino acid 229. In other embodiments of the foregoing formulae X and XI polypeptides, the surface-exposed, flexible loop structure comprising A2-1 corresponds to a region in native mature human factor VIII from about amino acid 395 to about amino acid 421. In other embodiments of the foregoing formulae X and XI polypeptides, the A2-1 corresponds to a region in native mature human factor VIII from about amino acid 397 to about amino acid 418. In other embodiments of the foregoing formulae X and XI polypeptides, the surface-exposed, flexible loop structure comprising A2-2 corresponds to a region in native mature human factor VIII from about amino acid 577 to about amino acid 635. In other embodiments of the foregoing formulae X and XI polypeptides, the A2-2 corresponds to a region in native mature human factor VIII from about amino acid 595 to about amino acid 607. In other embodiments of the foregoing formulae X and XI polypeptides, the surface-exposed, flexible loop structure comprising A3-1 corresponds to a region in native mature human factor VIII from about amino acid 1705 to about amino acid 1732. In other embodiments of the foregoing formulae X and XI polypeptides, the A3-1 corresponds to a region in native mature human factor VIII from about amino acid 1711 to about amino acid 1725. In other embodiments of the foregoing formulae X and XI polypeptides, the the surface exposed, flexible loop structure comprising A3-2 corresponds to a region in native mature human factor VIII from about amino acid 1884 to about amino acid 1917. In other embodiments of the foregoing formulae X and XI polypeptides, the A3-2 corresponds to a region in native mature human factor VIII from about amino acid 1899 to about amino acid 1911. In other embodiments of the foregoing formulae X and XI polypeptides, an XTEN sequence is inserted into at least two of the regions A1-1, A1-2, A2-1, A2-2, A3-1, or A3-2. In other embodiments of the foregoing formulae X and XI polypeptides, an XTEN sequence is inserted immediately downstream of an amino acid which corresponds to an amino acid in mature native human factor VIII selected from the group consisting of amino acid residue number 32, 220, 224, 336, 339, 399, 416, 603, 1656, 1711, 1725, 1905 and 1910. In other embodiments of the foregoing formulae X and XI polypeptides, an additional XTEN sequence is inserted into the a3 acidic spacer region. In other embodiments of the foregoing formulae X and XI polypeptides, an additional XTEN sequence is inserted into the a3 acide spacer immediately downstream of an amino acid which corresponds to amino acid 1656. In other embodiments of the foregoing formulae X and XI polypeptides, the A1 domain comprises an XTEN permissive loop-1 (A1-1) region and an XTEN permissive loop-2 (A1-2) region wherein the A2 domain comprises an XTEN permissive loop-1 (A2-1) region and an XTEN permissive loop-2 (A2-2) region, and wherein the A3 domain comprises an XTEN permissive loop-1 (A3-1) region and an XTEN permissive loop-2 (A3-2) region, and wherein an additional XTEN sequence is inserted into at least one of the regions A1-1, A1-2, A2-1, A2-2, A3-1, or A3-2. In other embodiments of the foregoing formulae X and XI polypeptides, an additional XTEN sequence is inserted immediately downstream of an amino acid which corresponds to an amino acid in mature native human factor VIII selected from the group consisting of amino acid

residue number 32, 220, 224, 336, 339, 390, 399, 416, 603, 1656, 1711, 1725, 1905 and 1910. In the foregoing embodiments of formulae X and XI polypeptides, the fusion protein exhibits at least about 30%, 40%, 50%, 60%, 70%, or 80%, or 90% of the procoagulant activity of the corresponding factor VIII not linked to XTEN, wherein the procoagulant activity is assayed by an in vitro coagulation assay.

[0024] In all embodiments, the polypeptide can, for example, exhibit an in vitro procoagulant activity exceeding 0.5 IU/ml, or 1.0, or 1.5, or 2.0 IU/ml when expressed in cell-culture medium and assayed by an in vitro coagulation assay. The procoagulant activity can be measured by a chromogenic assay, a one stage clotting assay (e.g., a aPTT) or both.

[0025] In some embodiments, wherein the recombinant factor VIII fusion protein comprises a factor VIII and at least a first and a second XTEN, the at least first XTEN is separated from the at least second XTEN by at least 10 amino acids, at least 50 amino acids, at least 100 amino acids, at least 200 amino acids, at least 300 amino acids, or at least 400 amino acids.

[0026] In preferred embodiments, the recombinant factor VIII fusion protein comprising a factor VIII and at least a first XTEN and, optionally, at least a second, or optionally at least a third, or optionally at least a fourth XTEN, the fusion protein exhibits reduced binding to an anti-factor VIII antibody as compared to the corresponding factor VIII not linked to XTEN. The reduced binding can be assessed either in vivo or by an in vitro assay. In one embodiment, the in vitro assay is an ELISA assay, wherein the binding of an anti FVIII antibody to the fusion protein is reduced at least about 5%, 10%, 15%, 20%, 25%, 30%, 35% or at least about 40% or more compared to a FVIII not linked to XTEN. In another embodiment, the in vitro assay is a Bethesda assay wherein the reduced binding of the fusion protein is evidenced by a lower Bethesda titer of at least about 2, 4, 6, 8, 10, 12, 15, 20, 30, 40, 50, 60, 70, 80, 100, or 200 Bethesda units for the fusion protein compared to that for a factor VIII not linked to XTEN. In the in vitro assays, the anti-factor VIII antibody is selected from an antibody of Table 10 and polyclonal antibody from a hemophilia A patient with factor VIII inhibitors. In particular embodiments of a recombinant factor VIII fusion protein comprising a factor VIII and at least a first and a second XTEN exhibiting reduced binding to a factor VIII inhibitor antibody, the first XTEN is linked to said factor VIII polypeptide within a C2 domain of said factor VIII polypeptide, and the second XTEN is linked to said factor VIII polypeptide within an A1 or A2 domain of said factor VIII polypeptide, wherein said fusion protein exhibits reduced binding to a factor VIII inhibitor antibody as compared to the corresponding factor VIII not linked to XTEN, wherein the factor VIII inhibitor antibody is capable of binding to an epitope located within the A1, A2 or C2 domain, and further wherein the fusion protein exhibits procoagulant activity. In one embodiment of the foregoing fusion protein, the second XTEN is linked to said factor VIII polypeptide within the A2 domain of the factor VIII polypeptide and the factor VIII inhibitor antibody binds to the A2 domain of the factor VIII polypeptide. In another embodiment of the foregoing fusion protein, the second XTEN is linked to said factor VIII polypeptide within the C2 domain of the factor VIII polypeptide and the factor VIII inhibitor antibody binds to the C2 domain of the factor VIII polypeptide. The binding of an anti-factor VIII antibody to the fusion protein is reduced by at least about 5%, 10%, 15%, 20%, 25%, 30%, 35% or 40% compared to the corresponding factor

VIII not linked to XTEN when assayed by an ELISA assay, wherein the anti-factor VIII antibody is selected from the group consisting of the antibodies in Table 10 and a polyclonal antibody from a hemophilia A subject with factor VIII inhibitors. The foregoing fusion proteins can further comprise at least three XTENs, wherein the at least third XTEN is linked to the factor VIII at a site selected from within or replacing the B domain, at the C-terminus, and at or within 1, 2, 3, 4, 5, or 6 amino acids of an insertion site selected from Table 7 or Table 9. In the embodiments with reduced binding to anti-factor VIII antibodies, the fusion protein has greater procoagulant activity in the presence of the anti-FVIII antibody of at least 10%, 20%, 30%, 40%, 50%, 80%, 100%, 200%, 300%, 400%, or 500% or more compared to a corresponding factor VIII not linked to XTEN when assayed by an in vitro coagulation assay (e.g., a chromogenic or one-stage clotting assay).

[0027] In all embodiments, the XTEN of the fusion protein can, for example, be characterized in that the XTEN comprise at least 36, or at least 42, or at least 72, or at least 96, or at least 144, or at least 288, or at least 400, or at least 500, or at least 576, or at least 600, or at least 700, or at least 800, or at least 864, or at least 900, or at least 1000, or at least 2000, to about 3000 amino acid residues or even more residues; the sum of glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P) residues constitutes at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% of the total amino acid residues of the XTEN; the XTEN is substantially non repetitive such that (i) the XTEN contains no three contiguous amino acids that are identical unless the amino acids are serine; (ii) at least about 80% of the XTEN sequence consists of non-overlapping sequence motifs, each of the sequence motifs comprising about 9 to about 14, or about 12 amino acid residues consisting of four to six amino acids selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), wherein any two contiguous amino acid residues do not occur more than twice in each of the non-overlapping sequence motifs; or (iii) the XTEN sequence has a subsequence score of less than 10; the XTEN has greater than 90%, or greater than 95%, or greater than 99% random coil formation as determined by GOR algorithm; the XTEN has less than 2% alpha helices and 2% beta-sheets as determined by Chou-Fasman algorithm; the XTEN lacks a predicted T-cell epitope when analyzed by TEPITOPE algorithm, wherein the TEPITOPE threshold score for said prediction by said algorithm has a threshold of –9, and wherein said fusion protein exhibits a terminal half-life that is longer than at least about 12 h, or at least about 24 h, or at least about 48 h, or at least about 72 h, or at least about 96 h, or at least about 120 h, or at least about 144 h, or at least about 21 days or greater. In one embodiment, the recombinant factor VIII fusion protein comprises at least a second, or at least a third, or at least a fourth XTEN, which can be identical or different to the other XTEN. According to a different approach, the at least one, at least a second, or at least a third, or at least a fourth XTEN of the CFXTEN fusion protein each have at least about 80% sequence identity, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% sequence identity compared to one or more XTEN of comparable length selected from Table 4, Table 13, Table 14, Table 15, Table 16, and Table 17, when optimally aligned. In yet another different approach, the at least one, at

least a second, or at least a third, or at least a fourth XTEN of the CFXTEN fusion protein each have at least 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% sequence identity compared to a sequence selected from AE42_1, AE42_2, AE42_3, AG42_1, AG42_2, AG42_3, AG42_4, AE144_1A, AE144_2A, AE144_2B, AE144_3A, AE144_3B, AE144_4A, AE144_4B, AE144_5A, AE144_6B, AG144_1, AG144_2, AG144_A, AG144_B, AG144_C, AG144_F, AG144_3, AG144_4, AE288_1, AE288_2, AG288_1, and AG288_2.

[0028] In one embodiment, the factor VIII component of the CFXTEN recombinant factor VIII fusion protein comprisies one, two or three amino acid substitutions selected from residues R1648, Y1680, and R1689, numbered relative to mature human factor VIII, wherein the substitutions are selected from alanine, glycine, and phenylalanine. Non-limiting examples of said substitutions include R1648A, Y1680F, and R1689A.

[0029] In another embodiment, the CFXTEN fusion protein exhibits an apparent molecular weight factor of at least about 1.3, or at least about two, or at least about three, or at least about four, or at least about five, or at least about six, or at least about seven, or at least about eight, or at least about nine, or at least about 10, when measured by size exclusion chromatography or comparable method.

[0030] In some embodiments of the CFXTEN fusion proteins, one or more of the XTEN is to the FVIII via one or two cleavage sequences that each is cleavable by a mammalian protease selected from the group consisting of factor XIa, factor XIIa, kallikrein, factor VIIa, factor IXa, factor Xa, factor IIa (thrombin), Elastase-2, MMP-12, MMP13, MMP-17 and MMP-20, wherein cleavage at the cleavage sequence by the mammalian protease releases the factor VIII sequence from the XTEN sequence, and wherein the released factor VIII sequence exhibits an increase in procoagulant activity compared to the uncleaved fusion protein. In one embodiment, the cleavage sequence(s) are cleavable by factor XIa.

[0031] According to a different approach, the CFXTEN fusion proteins comprise at least three XTENs located at different locations of the factor VIII polypeptide, wherein said different locations are selected from: an insertion location at or within 1 to 6 amino acids from a site selected from Table 5, Table 6, Table 7 Table 8, and Table 9; a location at or within 1 to 6 amino acids of amino acid residue 32, 220, 224, 336, 339, 390, 399, 416, 603, 1656, 1711, 1725, 1905 and 1910 of mature factor VIII; a location between any two adjacent domains in the factor VIII sequence, wherein said two adjacent domains are selected from the group consisting of A1 and A2, A2 and B, B and A3, A3 and C1, and C1 and C2; a location within an internal B domain deletion starting from a first position at about amino acid residue number 741 to about 750 and ending at a second position at amino acid residue number 1635 to about 1648 with reference to full-length human factor VIII sequence as set forth in FIG. 3 and the C-terminus of the factor VIII sequence, wherein the cumulative length of the multiple XTENs is at least about 100 to about 3000 amino acid residues and wherein the fusion protein retains at least about 30%, or about 40%, or about 50%, or about 60%, or about 70%, or about 80%, or about 90% of the procoagulant activity compared to the corresponding factor VIII not linked to XTEN, wherein the procoagulant activity is assayed by an in vitro coagulation assay. In one embodiment of the foregoing, the fusion protein exhibits

a prolonged terminal half-life when administered to a subject as compared to a corresponding factor VIII polypeptide lacking said XTEN, wherein said fusion protein exhibits a terminal half-life at least about 3 hours, or 4 hours, or 6 hours, or 12 hours, or 13 hours, or 14 hours, or 16 hours, or 24 hours, or 48 hours, or 72 hours, or 96 hours, or 120 hours, or 144 hours, or 7 days, or 14 days, or 21 days when administered to a subject. In one embodiment, the subject is selected from the group consisting of human and a factor VIII/von Willebrand factor double knock-out mouse. In one embodiment of the foregoing, the fusion protein does not comprise a sequence selected from GTPGSGTASSSP (SEQ ID NO: 31), GSSTPSGATGSP (SEQ ID NO: 32), GSSPSASTGTGP (SEQ ID NO: 33), GASPGTSSTGSP (SEQ ID NO: 34), and GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP (SEQ ID NO: 59). In another embodiment of the foregoing, the fusion protein does not contain an XTEN sequence consisting of GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP (SEQ ID NO: 59), PGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSS (SEQ ID NO: 71), or PGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGS (SEQ ID NO: 80).

[0032] In a further aspect, the invention concerns CFXTEN fusion proteins with enhanced pharmacokinetic properties, including enhanced parameters compared to FVIII not linked to XTEN, wherein the enhanced properties include but are not limited to longer terminal half-life, larger area under the curve, increased time in which the blood concentration remains within the therapeutic window, increased time between consecutive doses results in blood concentrations within the therapeutic window, and decreased dose in IU over time that can be administered compared to a FVIII not linked to XTEN, yet still result in a blood concentration above a threshold concentration needed for a procoagulant effect. In some embodiments, a CFXTEN fusion proteins exhibit a prolonged terminal half-life when administered to a subject as compared to a corresponding factor VIII polypeptide lacking said XTEN. The subject can be a human or a mouse, such as a factor VIII/von Willebrand factor double knock-out mouse. In one embodiment of the foregoing, the CFXTEN exhibits a terminal half-life that is at least about two-fold, or about three fold, or about four-fold, or about five-fold, or about 10-fold, or about 20-fold longer when administered to a subject compared to the corresponding factor VIII not linked to XTEN. In one embodiment, the CFXTEN fusion protein exhibits a terminal half-life at least about 3

hours, or 4 hours, or 6 hours, or 12 hours, or 13 hours, or 14 hours, or 16 hours, or 24 hours, or 48 hours, or 72 hours, or 96 hours, or 120 hours, or 144 hours, or 7 days, or 14 days, or 21 days when administered to the subject. In other embodiments, the enhanced pharmacokinetic property of the fusion proteins of the embodiments is the property of maintaining a circulating blood concentration of procoagulant fusion protein in a subject in need thereof above a threshold concentration of 0.01 IU/ml, or 0.05 IU/ml, or 0.1 IU/ml, or 0.2 IU/ml, or 0.3 IU/ml, or 0.4 IU/ml or 0.5 IU/ml for a period that is at least about two fold, or at least about three-fold, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about eight-fold, or at least about ten-fold, or at least about 20-fold, or at least about 40-fold, or at least about 60-fold longer compared to the corresponding FVIII not linked to XTEN and administered to a subject at a comparable dose. The increase in half-life and time spent above the threshold concentration permits less frequent dosing and decreased amounts of the fusion protein (in moles equivalent) that are administered to a subject, compared to the corresponding FVIII not linked to XTEN. In one embodiment, administration of a subject fusion protein to a subject using a therapeutically-effective dose regimen results in a gain in time of at least two-fold, or at least three-fold, or at least four-fold, or at least five-fold, or at least six-fold, or at least eight-fold, or at least 10-fold, or at least about 20-fold, or at least about 40-fold, or at least about 60-fold or higher between at least two consecutive Cmax peaks and/or Cmin troughs for blood levels of the fusion protein compared to the corresponding FVIII not linked to the XTEN and administered using a comparable dose regimen to a subject.

[0033] In preferred embodiments, the CFXTEN fusion proteins retain at least about 30%, or about 40%, or about 50%, or about 60%, or about 70%, or about 80%, or about 90% of the procoagulant activity compared to the corresponding factor VIII not linked to XTEN, wherein the procoagulant activity is assayed by an in vitro coagulation assay such as, but not limited to a chromogenic assay or a one- or two-stage clotting assay.

[0034] According to a different approach, the invention provides recombinant factor VIII fusion proteins comprising a factor VIII polypeptide and at least one extended recombinant polypeptide (XTEN), wherein said factor VIII polypeptide comprises A1 domain, A2 domain, A3 domain, C1 domain, C2 domain and optionally all or a portion of B domain, and wherein said at least one XTEN is linked to said factor VIII polypeptide at an insertion site selected form residue numbers 18-32, or 40, or 211-224, or 336-403, or 599, or 745-1640, or 1656-1728, or 1796-1804, or 1900-1912, or 2171-2332; and wherein the fusion protein retains at least about 30%, or about 40%, or about 50%, or about 60%, or about 70%, or about 80%, or about 90% of the procoagulant activity compared to the corresponding factor VIII not linked to XTEN. In one embodiment of the foregoing, the fusion protein comprises at least a second XTEN, or at least a third, or at least a fourth XTEN wherein the XTEN are linked to the factor VIII at a site at or within 1 to 6 amino acids of a site selected from Table 5, Table 6, Table 7, Table 8, and Table 9. In another embodiment, the invention provides an recombinant factor VIII fusion protein further comprising at least a second XTEN, or at least a third, or at least a fourth XTEN linked to said FVIII polypeptide at an insertion site selected from Table 5, Table 6, Table 7, Table 8, Table 9, at or within 6 amino acids to the N- or C-terminus side of an insertion location at one or more insertion

locations from Figure 8 and within one or more insertion ranges from Figure 9 wherein at least two XTEN are separated by an amino acid sequence of at least 100 to about 400 amino acids.

[0035] The invention provides CFXTEN wherein the XTEN have a Ratio XTEN Radii of at least 2.3 or at least 2.5, and are separated by an amino acid sequence of at least about 20 amino acid residues, or at least about 50, or at least about 100, or at least about 200, or at least about 300, or at least about 400 amino acid residues. In other embodiments, the CFXTEN comprise at least four XTEN wherein the XTEN have a Ratio XTEN Radii of at least 2.3, or at least 2.5, or at least 2.8, and wherein at least three of the four of the XTEN linked to the fusion protein are separated by an amino acid sequence of at least about 20 amino acid residues, or at least about 50, or at least about 100, or at least about 200, or at least about 300, or at least about 400 amino acid residues, and the fourth XTEN is linked within the B domain (or a fragment thereof) or within the C domain (or the terminus thereof).

[0036] In some embodiments, the subject compositions are configured to have reduced binding affinity for a clearance receptor in a subject as compared to the corresponding FVIII not linked to the XTEN. In one embodiment, the CFXTEN fusion protein exhibits binding affinity for a clearance receptor of the FVIII in the range of about 0.01%-30%, or about 0.1% to about 20%, or about 1% to about 15%, or about 2% to about 10% of the binding affinity of the corresponding FVIII not linked to the XTEN. In another embodiment, a fusion protein with reduced affinity for a clearance receptor has reduced active clearance and a corresponding increase in half life of at least about 2 fold, or 3 fold, or at least 4 fold, or at least about 5-fold, or at least about 6-fold, or at least about 7-fold, or at least about 8-fold,or at least about 9-fold, or at least about 10-fold, or at least about 12-fold, or at least about 15-fold, or at least about 17-fold, or at least about 20-fold longer compared to the corresponding FVIII that is not linked to the XTEN.

[0037] In an embodiment, the invention provides a recombinant factor VIII fusion protein comprising FVIII and one or more XTEN wherein the fusion protein exhibits increased solubility of at least three-fold, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about seven-fold, or at least about eight-fold, or at least about nine-fold, or at least about ten-fold, or at least about 15-fold, or at least a 20-fold, or at least 40-fold, or at least 60-fold at physiologic conditions compared to the FVIII not linked to XTEN.

[0038] In a further aspect, the invention provides a pharmaceutical composition comprising the fusion protein of any of the embodiments described herein and a pharmaceutically acceptable carrier.

[0039] In another embodiment, the invention provides a method of treating a coagulopathy in a subject, comprising administering to said subject a composition comprising a clotting effective amount of the pharmaceutical composition. In one embodiment of the method, after said administration, a blood concentration of procoagulant factor VIII is maintained at about 0.05, or 1, or 1.5 IU/ml or more for at least 48 hours after said administration. In another embodiment, the invention provides a method of clotting blood in a subject, comprising contacting a clotting effective amount of the pharmaceutical composition with the blood.

[0040] In another embodiment, the invention provides a method of treating a coagulopathy in a subject with circulating inhibitors of factor VIII, comprising administering to said subject a composition

comprising a therapeutically effective amount of the pharmaceutical composition of CFXTEN, wherein the composition exhibits greater procoagulant activity in said subject compared to a composition comprising the corresponding factor VIII not linked to XTEN and administered using a comparable amount. In one embodiment of the method, the coagulopathy is hemophilia A. In another embodiment, the coagulopathy is the result of trauma or surgery or infection.

[0041] The invention provides a method of treating a bleeding episode in a subject, comprising administering to said subject a composition comprising a clotting effective amount of the CFXTEN pharmaceutical composition, wherein the clotting effective amount of the fusion protein arrests a bleeding episode for a period that is at least three-fold, or at least four-fold, or at least five-fold longer compared to a corresponding factor VIII not linked to XTEN and administered using a comparable amount to said subject. Non-limiting examples of a corresponsing factor VIII not linked to XTEN include native FVIII, the sequences of Table 1, BDD-FVIII, and the pCB0114 FVIII.

[0042] In another embodiment, the invention provides a CFXTEN recombinant factor VIII fusion protein for use in a pharmaceutical regimen for treating a hemophilia A patient, said regimen comprising a pharmaceutical composition comprising a CFXTEN fusion protein. In one embodiment of the pharmaceutical regimen, the regimen further comprises the step of determining the amount of pharmaceutical composition comprising the CFXTEN needed to achieve hemostasis in the hemophilia A patient. In another embodiment, the pharmaceutical regimen for treating a hemophilia A subject comprises administering the pharmaceutical composition in two or more successive doses to the subject at an effective amount, wherein the administration results in at least a 10%, or 20%, or 30%, or 40%, or 50%, or 60%, or 70%, or 80%, or 90% greater improvement of at least one, two, or three parameters associated with the hemophilia A disease compared to the factor VIII not linked to XTEN and administered using a comparable dose. Non-limited examples of parameters improved include blood concentration of procoagulant FVIII, a reduced activated partial prothrombin (aPTT) assay time, a reduced one-stage or two-stage clotting assay time, delayed onset of a bleeding episode, a reduced chromogenic assay time, a reduced bleeding assay time, resolution of a bleeding event, or a reduced Bethesda titer to native FVIII.

[0043] In another aspect, the invention provides isolated nucleic acid sequences encoding the fusion proteins of any one of the embodiments of the CFXTEN fusion protein. In one embodiment, the isolated nucleic acid is the complement of a sequence encoding a CFXTEN fusion protein of the embodiments. In one embodiment, the isolated nucleic acid further comprises a sequence encoding a signal peptide, wherein said sequence is

ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCGATTCTGCTTTAGT (SEQ ID NO: 1613), or the complement thereof. In another embodiment, the invention provides an expression vector comprising the nucleic acid encoding the fusion protein, or the complement thereof. In another embodiment, the invention provides an isolated host cell comprising the foregoing expression vector. In another embodiment, the invention provides a method of producing the fusion protein of any of the

embodiments, comprising providing a host cell comprising the expression vector; culturing the host cell to effect production of the fusion protein; and recovering the fusion protein.

[0044] In one embodiment, the invention provides an isolated fusion protein comprising a polypeptide having at least about 80% sequence identity, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% sequence identity compared to a sequence of comparable length selected from Table 21, when optimally aligned.

[0045] In another embodiment, the invention provides an isolated nucleic acid comprising a polynucleotide sequence selected from (a) a sequence having at least about 80% sequence identity, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% sequence identity compared to a sequence of comparable length selected from Table 21, when optimally aligned, or (b) the complement of the polynucleotide of (a). In another embodiment, the isolated nucleic acid comprises the sequence ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCGATTCTGCTTTAGT (SEQ ID NO: 1613) linked to the 5' end of the nucleic acid of (a) or the complement of the sequence linked to the 3' end of (b).

[0046] It is specifically contemplated that the recombinant factor VIII fusion proteins can exhibit one or more or any combination of the properties disclosed herein.

Various embodiments of the invention relate to a recombinant factor VIII (FVIII) fusion protein comprising a FVIII polypeptide and a first extended recombinant polypeptide (XTEN), wherein the FVIII polypeptide comprises an A3 domain, an A2 domain, an A1 domain, a C2 domain, a C1 domain, and a B domain or a portion of the B domain, wherein the first XTEN is inserted into the FVIII polypeptide at a site immediately downstream of amino acid 745 of the sequence set forth in SEQ ID NO:2; and wherein the recombinant FVIII fusion protein exhibits a prolonged terminal half-life when administered to a subject as compared to a corresponding FVIII polypeptide lacking the first XTEN. Various embodiments relate to an isolated nucleic acid encoding the recombinant FVIII fusion protein. Various embodiments relate to a host cell comprising an expression vector comprising the nucleic acid. Various embodiments relate to a method of making the recombinant FVIII fusion protein, wherein the method comprises culturing the host cell in media under conditions suitable for making the fusion protein and recovering the fusion protein. Various embodiments relate to a pharmaceutical composition comprising (i) the recombinant factor VIII fusion protein, and (ii) a pharmaceutically acceptable carrier. The recombinant FVIII fusion protein or the pharmaceutical composition may be used for treating a bleeding disease or condition in a subject in need thereof.

Date Reçue/Date Received 2022-02-04

Various embodiments of the invention relate to a recombinant factor VIII fusion protein comprising a factor VIII polypeptide fused to an extended recombinant polypeptide (XTEN) and a Fc fragment of immunoglobulin; wherein the factor VIII polypeptide comprises amino acid residues 1-745 and 1649-2332 of mature human factor VIII (SEQ ID NO: 2); wherein an FVIII processing site corresponding to amino acid residue R1648 of SEQ ID NO: 2 is eliminated in the FVIII polypeptide; wherein the XTEN is inserted within the factor VIII polypeptide immediately downstream of an amino acid corresponding to residue 745 of SEQ ID NO: 2; and wherein the XTEN comprises the amino acid sequence set forth in SEQ ID NO: 78. Various embodiments relate to an isolated nucleic acid encoding the recombinant FVIII fusion protein. Various embodiments relate to a host cell comprising an expression vector comprising the nucleic acid. Various embodiments relate to a method of making the recombinant FVIII fusion protein, wherein the method comprises culturing the host cell in media under conditions suitable for making the fusion protein and recovering the fusion protein. Various embodiments relate to a pharmaceutical composition comprising (i) the recombinant factor VIII fusion protein, and (ii) a pharmaceutically acceptable carrier. The recombinant FVIII fusion protein or the pharmaceutical composition may be used for treating a bleeding disease or condition in a subject in need thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

[0048] The features and advantages of the invention may be further explained by reference to the following detailed description and accompanying drawings that sets forth illustrative embodiments.

[0049] FIG. 1 shows a schematic representation of the FVIII architecture and spatial arrangement of the domains during processing and clotting, and is intended to represent both native FVIII and B domain deleted variants. The A1 domain ranges from residue 1 to 372 (numbering relative to the mature form of FVIII sequence NCBI Protein RefSeq NP_000123 and encompassing a1 residues), A2 domain ranges from residue 373 to 740, B domain ranges from residue 741 to 1648, A3 domain ranges from residue 1649 to 2019 (encompassing a3 acidic region), C1 domain ranges from 2020 to 2172, and the C2 domain ranges from residue 2173 to 2332. BDD variants include deletions between the range 741 to 1648, leaving some or no remnant residues, with a non-limiting BDD remnant sequence being SFSQNPPVLKRHQR (SEQ ID NO: 1614). FIG. 1A shows the domain architecture of a single chain FVIII prior to processing. Arrows indicate the sites at residues R372, R740, R1648, and R1689 that are cleaved in the processing and conversion of FVIII to FVIIIa. FIG. 1B shows the FVIII molecule that has been processed into the heterodimer by the cleavage at the R1648 residue, with the a3 acidic region of

Date Reçue/Date Received 2021-02-26

the A3 domain indicated on the N-terminus of the A3. FIG. 1C shows the FVIII molecule processed into the FVIIIa heterotrimer by the cleavage at the R372, R740, and R1689 residues.

[0050] FIG. 2 is a schematic of the coagulation cascade, showing the intrinsic and extrinsic arms leading to the common pathway.

[0051] FIG. 3 depicts the amino acid sequence of mature human factor VIII (SEQ ID NO: 1592).

[0052] FIG. 4 depicts a factor VIII sequence with a deletion of a portion of the B domain (SEQ ID NO: 1593).

[0053] FIG. 5 illustrates several examples of CFXTEN configurations of FVIII linked to XTEN (the latter shown as thick, wavy lines). In all cases, the FVIII can be either native or a BDD form of FVIII, or a single chain form in which the entire B domain, including the native cleavage sites are removed. FIG. 5A shows, left to right, three variations of single chain factor VIII with XTEN linked to the N-terminus, the C-terminus, and two XTEN linked to the N- and C-terminus. FIG. 5B shows six variations of mature heterodimer FVIII with, left to right, an XTEN linked to the N-terminus of the A1 domain; an XTEN linked to the C-terminus of the C2 domain; an XTEN linked to the N-terminus of the A1 domain and the C-terminus of the C2 domain; an XTEN linked to the N-terminus of the A1 domain and to the N-terminus of the A3 domain; an XTEN linked to the C-terminus of the C2 domain and to the N-terminus of the A3 domain via residual B domain amino acids; and an XTEN linked to the N-terminus of the A1 domain, the C terminus of the A2 domain via residual B domain amino acids, and to the C terminus of the C2 domain. FIG. 5C shows, left to right, three variations of single chain factor VIII: an XTEN linked to the N-terminus of the A1 domain, an XTEN linked within a surface loop of the A1 domain and an XTEN linked within a surface loop of the A3 domain; an XTEN linked within a surface loop of the A2 domain, an XTEN linked within a surface loop of the C2 domain and an XTEN linked to the C terminus of the C2 domain; an XTEN linked to the N-terminus of the A1 domain and within a surface loop of the C1 domain and to the C-terminus of the C domain. FIG. 5D shows six variations of mature heterodimer FVIII with, left to right, an XTEN linked to the N-terminus of the A1 domain, an XTEN linked within a surface loop of the A1 domain, and an XTEN linked within a surface loop of the A3 domain; an XTEN linked within a surface loop of the A2 domain, and an XTEN linked within a surface loop of the C1 domain, and an XTEN linked to the C-terminus of the C2 domain; an XTEN linked to the N-terminus of the A1 domain, an XTEN linked within a surface loop of the A1 domain, an XTEN linked within a surface loop of the A3 domain, and an XTEN linked to the C-terminus of the C2 domain; an XTEN linked to the N-terminus of the A1 domain, an XTEN linked to the N-terminus of the A3 domain via residual amino acids of the B domain, and an XTEN linked within a surface loop of the C2 domain; an XTEN linked within a surface loop of the A2 domain, an XTEN linked to the N-terminus of the A3 domain via residual amino acids of the B domain, an XTEN linked within a surface loop of the C1 domain, and an XTEN linked to the C-terminus of the C2 domain; and an XTEN linked within the B domain or between the residual B domain residues of the BDD variant (and the invention also contemplates a variation in which the XTEN replaces the entirety of the B domain, including all native cleavage sites, linking the A2 and A3 domains, resulting in a single chain form of factor VIII). This

figure also embodies all variations in which one or more XTEN sequences are inserted within the B domain and the resulting fusions are cleaved at one or more sites (e.g., at R1648 site) during intracellular processing.

[0054] FIG. 6 is a graphic portrayal of a CFXTEN construct with an XTEN inserted within the B domain and linked to the C-terminus of the C2 domain illustrating the unstructured characteristic of the XTEN leading to random coil formation that can cover portions of the factor VIII proximal to the XTEN. In the lower panel, the drawing depicts that when XTEN is in random coil, it can adopt a conformation resulting in steric hindrance that blocks binding of factor VIII inhibitor antibodies that would otherwise have affinity for epitopes proximal to the XTEN site of insertion.

[0055] FIG. 7 is a graphic portrayal of the various analyses performed on a FVIII B-domain deleted sequence to identify insertion sites for XTEN within the FVIII sequence. Each of lines A-H are on an arbitrary scale of Y axis values across the FVIII BDD sequence such that low values represent areas with a high predicted tolerance for XTEN insertion, with the residue numbers on the X axis. Line A shows the domain boundaries; all discontinuities in this line represent boundaries that are likely to accept XTEN. Line B shows exon boundaries; i.e., each step in the line represents a new exon. Line C shown regions that were not visible in the X-ray structure due to a lack of order in the crystal. Lines labeled D represents multiple predictions of order that were calculated using the respective programs FoldIndex found on the World Wide web site bip.weizmann.ac.il/fldbin/findex (last accessed February 23, 2011) (*see* Jaime Prilusky, Clifford E. Felder, Tzviya Zeev-Ben-Mordehai, Edwin Rydberg, Orna Man, Jacques S. Beckmann, Israel Silman, and Joel L. Sussman, 2005, Bioinformatics based on the Kyte & Doolitlle algorithm, as well as RONN found on the World-Wide web site strubi.ox.ac.uk/RONN (last accessed February 23, 2011) (*see* Yang,Z.R., Thomson, R., McMeil, P. and Esnouf, R.M. (2005) RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins Bioinformatics 21: 3369-3376. Lines E and F were calculated based on multiple sequence alignments of FVIII genes from 11 mammals available in GenBank. Line E represents the conservation of individual residues. Line F represent the conservation of 3 amino acid segments of FVIII. Lines G and H represent gaps and insertions observed in the multiple sequence alignment of 11 mammalian FVIII genes. Line J lists the XTEN insertion points by amino acid number that were obtained based by combining the multiple measurements above.

[0056] FIG. 8 depicts the sites in a FVIII B-domain deleted sequence (SEQ ID NO: 1594) identified as active insertion points for XTEN using the information depicted in FIG. 8 and as confirmed in the assays of Example 34.

[0057] FIG. 9 depicts the range of sites in a FVIII B-domain deleted sequence (SEQ ID NO: 1595) identified for insertion of XTEN using the information depicted in FIG. 8 and or Example 34 plus a span of amino acids around each insertion point that are considered suitable for insertion of XTEN.

[0058] FIG. 10 is a schematic of the assembly of a CFXTEN library created by identifying insertion points as described for FIGS. 7 followed by insertion of single XTEN (black bars) at the various insertion points using molecular biology techniques. The constructs are expressed and recovered, then evaluated

for FVIII activity and pharmacokinetic properties to identify those CFXTEN configurations that result in enhanced properties.

[0059] FIG. 11 is a schematic of the assembly of a CFXTEN component library in which segments of FVIII BDD domains, either singly or linked to various lengths of XTEN (black bars) are assembled in a combinatorial fashion into libraries of genes encoding the CFXTEN, which can then be evaluated for FVIII activity and pharmacokinetic properties to identify those CFXTEN configurations that result in enhanced properties.

[0060] FIG. 12 illustrates several examples of CFXTEN configurations with XTEN (shown as thick, wavy lines), with certain XTEN releasable by inserting cleavage sequences (indicated by black triangles) that are cleavable by procoagulant proteases. FIG. 12A illustrates a scFVIII with two terminal releasable XTENS. FIG. 12B illustrates the same configuration as FIG. 12A but with an additional non-releasable XTEN linking the A3 and C1 domains. FIG. 12C illustrates a mature heterodimer FVIII with two terminal releasable XTEN. FIG. 12D illustrates the same configuration as 10C but with an additional non-releasable XTEN linking the A3 and C1 domains.

[0061] FIG. 13 is a schematic flowchart of representative steps in the assembly, production and the evaluation of an XTEN.

[0062] FIG. 14 is a schematic flowchart of representative steps in the assembly of a CFXTEN polynucleotide construct encoding a fusion protein. Individual oligonucleotides **501** are annealed into sequence motifs **502** such as a 12 amino acid motif ("12-mer"), which is ligated to additional sequence motifs from a library to create a pool that encompasses the desired length of the XTEN **504**, as well as ligated to a smaller concentration of an oligo containing BbsI, and KpnI restriction sites **503.** The resulting pool of ligation products is gel-purified and the band with the desired length of XTEN is cut, resulting in an isolated XTEN gene with a stopper sequence **505**. The XTEN gene is cloned into a stuffer vector. In this case, the vector encodes an optional CBD sequence **506** and a GFP gene **508.** Digestion is then performed with BbsI/HindIII to remove **507** and **508** and place the stop codon. The resulting product is then cloned into a BsaI/HindIII digested vector containing a gene encoding the FVIII, resulting in the gene **500** encoding an FVIII-XTEN fusion protein.

[0063] FIG. 15 is a schematic flowchart of representative steps in the assembly of a gene encoding fusion protein comprising a CF and XTEN, its expression and recovery as a fusion protein, and its evaluation as a candidate CFXTEN product.

[0064] FIG. 16 illustrates the use of donor XTEN sequences to produce truncated XTENs. FIG. 16A provides the sequence of AG864 (SEQ ID NO: 1596), with the underlined sequence used to generate a sequence length of 576 (SEQ ID NO: 1597). FIG. 16B provides the sequence of AG864 (SEQ ID NO: 1598), with the underlined sequence used to generate a sequence length of 288 (SEQ ID NO: 1599). FIG. 16C provides the sequence of AG864 (SEQ ID NO: 1600), with the underlined sequence used to generate a sequence length of 144 (SEQ ID NO: 1601). FIG. 16D provides the sequence of AE864 (SEQ ID NO: 1602), with the underlined sequence used to generate a sequence length of 576 (SEQ ID NO: 1603). FIG. 16E provides the sequence of AE864 (SEQ ID NO: 1604), with the underlined sequence

used to generate a sequence length of 288 (SEQ ID NO: 1605). FIG. 16F provides the sequence of AE864 (SEQ ID NO: 1606) used to generate four sequences of 144 length (SEQ ID NOS 1607-1610, respectively, in order of appearance) (the double underline indicates the first amino acid in the 144 sequence with the single underline representing the balance of that sequence).

[0065] FIG. 17 is a schematic representation of the design of Factor VIII-XTEN expression vectors with different strategies introducing XTEN elements into the FVIII coding sequence. FIG. 17A shows an expression vector encoding XTEN fused to the 3' end of the sequence encoding FVIII. FIG. 17B depicts an expression vector encoding an XTEN element inserted into the middle of the coding sequence encoding a single FVIII. FIG. 17C depicts an expression vector encoding two XTEN elements: one inserted internal to the FVIII coding sequence, and the other fused to the 3' end of the FVIII coding sequence.

[0066] FIG. 18 illustrates the process of combinatorial gene assembly of genes encoding XTEN. In this case, the genes are assembled from 6 base fragments and each fragment is available in 4 different codon versions (A, B, C and D). This allows for a theoretical diversity of 4096 in the assembly of a 12 amino acid motif.

[0067] FIG. 19 shows the pharmacokinetic profile (plasma concentrations) in cynomolgus monkeys after single doses of different compositions of GFP linked to unstructured polypeptides of varying length, administered either subcutaneously or intravenously, as described in Example 41. The compositions were GFP-L288, GFP-L576, GFP-XTEN_AF576, GFP-Y576 and XTEN_AD836-GFP. Blood samples were analyzed at various times after injection and the concentration of GFP in plasma was measured by ELISA using a polyclonal antibody against GFP for capture and a biotinylated preparation of the same polyclonal antibody for detection. Results are presented as the plasma concentration versus time (h) after dosing and show, in particular, a considerable increase in half-life for the XTEN_AD836-GFP, the composition with the longest sequence length of XTEN. The construct with the shortest sequence length, the GFP-L288 had the shortest half-life.

[0068] FIG. 20 shows an SDS-PAGE gel of samples from a stability study of the fusion protein of XTEN_AE864 fused to the N-terminus of GFP (see Example 42). The GFP-XTEN was incubated in cynomolgus plasma and rat kidney lysate for up to 7 days at 37°C. In addition, GFP-XTEN administered to cynomolgus monkeys was also assessed. Samples were withdrawn at 0, 1 and 7 days and analyzed by SDS PAGE followed by detection using Western analysis with antibodies against GFP.

[0069] FIG. 21 shows results of a size exclusion chromatography analysis of glucagon-XTEN construct samples measured against protein standards of known molecular weight, with the graph output as absorbance versus retention volume, as described in Example 40. The glucagon-XTEN constructs are 1) glucagon-Y288; 2) glucagonY-144; 3) glucagon-Y72; and 4) glucagon-Y36. The results indicate an increase in apparent molecular weight with increasing length of XTEN moiety (see Example 40 for data).

[0070] FIG. 22 shows results of a Western blot of proteins expressed by cell culture of cells transformed with constructs as designated (Example 25). The samples in lanes 1-12 were: MW Standards, FVIII (42.5 ng), pBC0100B, pBC0114A, pBC0100, pBC0114, pBC0135, pBC0136,

pBC0137, pBC0145, pBC0149, and pBC0146, respectively. Lanes 8, 9 and 12 show bands consistent with a FVIII with a C-terminal XTEN288, with an estimated MW of 95 kDa. Lanes 7 and 11 show bands consistent with a FVIII with a C-terminal XTEN42, with an estimated MW of 175 kDa. Lanes 2-6 show bands consistent with FVIII and heavy chain. Lanes 10 and 23 show bands consistent with heavy chain. Lane 7 shows a band consistent with heavy chain and an attached XTEN42.

[0071] FIG. 23 shows the results of FVIII assay on samples obtained from FVIII and von Willebrand factor double knock-out mice with hydrodynamic plasmid DNA injection, as detailed in Example 36.

[0072] FIG. 24 is a graphic and tabular portrayal of the pharmacokinetic properties of rBDD-FVIII and the purified CFXTEN fusion proteins pBC0145 and pBC0146 (with C-terminal XTEN) administered to either HemA or FVIII/VWF double knock-out mice as described in Example 30, showing the enhanced half-life of the CFXTEN in both strains of mice.

[0073] FIG. 25 is a graphic and tabular portrayal of the pharmacokinetic properties of rBDD-FVIII and the CFXTEN fusion proteins pSD0050 and pSD0062 (with internal inserted XTEN) administered to either HemA (FIG. 25A) or FVIII/VWF double knock-out mice (FIG. 25B) using a cell culture PK assay in HemA mice. Dose, 5-minute recovery, and half-life (T1/2) are shown, as described in Example 32, underscoring the enhanced recovery and half-life of the CFXTEN compared to the positive control FVIII in both strains of mice.

[0074] FIG. 26 is a graphic depiction of a titration of GMA8021 FVIII inhibitor using the pBC0114 BDD-FVIII AND CFXTEN construct LSD0049.002 with three 144 amino acid XTEN insertions at residues 18, 745 and 2332. The data indicate a right-shift of approximately 0.7 order of magnitude in the amount of antibody in μg/ml required to inhibit the CFXTEN to the 50% level, compared to FVIII positive control.

[0075] FIG. 27 is a schematic of the logic flow chart of the algorithm SegScore. In the figure the following legend applies: i, j - counters used in the control loops that run through the entire sequence; HitCount- this variable is a counter that keeps track of how many times a subsequence encounters an identical subsequence in a block; SubSeqX - this variable holds the subsequence that is being checked for redundancy; SubSeqY - this variable holds the subsequence that the SubSeqX is checked against; BlockLen - this variable holds the user determined length of the block; SegLen - this variable holds the length of a segment. The program is hardcoded to generate scores for subsequences of lengths 3, 4, 5, 6, 7, 8, 9, and 10; Block - this variable holds a string of length BlockLen. The string is composed of letters from an input XTEN sequence and is determined by the position of the i counter; SubSeqList - this is a list that holds all of the generated subsequence scores.

[0076] FIG. 28 depicts the application of the algorithm SegScore to a hypothetical XTEN of 11 amino acids (SEQ ID NO: 1591) in order to determine the repetitiveness. An XTEN sequence consisting of N amino acids is divided into N-S+1 subsequences of length S (S=3 in this case). A pair-wise comparison of all subsequences is performed and the average number of identical subsequences is calculated to result in the subsequence score of 1.89.

[0077] FIG. 29 is a graph of the individual construct values of the ratio of FVIII activity in the assayed CFXTEN to that of the pBC114 FVIII positive control after exposure to the GMA8021 antibody to FVIII, grouped according to the number of XTEN in the construct fusion protein (see Example 28). The results show an essentially linear relationship in the ability of the CFXTEN to retain FVIII activity with increasing number of incorporated XTEN.

[0078] FIG. 30 depicts the primary sequence and domain structure of mature B-domain deleted (BDD) human FVIII construct (Example 46). The location of the introduced *NheI* and *ClaI* restriction sites is shown. Note that the amino acid numbering corresponds to the amino acid positions in the primary sequence of mature FVIII (FIG. 30). Individual domains are bounded by gray lines/boxes with domain identification in gray text. Acidic regions (a1, a2, a3) are indicated with dashed boxes. Solid wedges/triangles indicate sites of thrombin cleavage in the activation of FVIII to FVIIIa. Unfilled wedges/triangle indicates the site of intracellular proteolytic processing to the two-chained form of FVIII. Hexagons indicate sites of N-linked glycosylation. Circles indicate sites of Tyr sulfation. Unique non-native restriction sites (*NheI*, GCTAG; *ClaI*, ATCGAT) introduced into cDNA to facilitate XTEN insertion/recombination are highlighted in gray with double underline.

[0079] FIG. 31 provides graphical representation of the FVIII construct described in FIG. 30, indicating the domain organization and the location of native and non-native restriction sites.

[0080] FIG. 32 shows the graphical ASAView outputs for structural datasets 2R7E, 3CDZ, and PM0076106. Accessible Solvent Areas (ASA) for the amino acids in domains A1, A2, A3, C1 and C2 are shown. Analyses were performed on X-ray crystallographic coordinates 3CDZ (Ngo et al., Structure 16: 597-606 (2008)) and 2R7E (Shen et al., Blood 111:1240-1247 (2008)) deposited in the Protein Data Bank maintained by the Research Collaboratory for Structural Bioinformatics (RCSB; http://www.rcsb.org/pdb), as well as on atomic coordinates PM0076106 for the predicted refined FVIII structure derived from a molecular dynamics simulation study (Venkateswarlu, BMC Struct. Biol. 10:7 (2010)) deposited in the Protein Model Database (http://mi.caspur.it/PMDB/main.php) maintained by Consorzio Interuniversitario per le Applicazioni di Supercalcolo per Università e Riserca (CASPUR) and the Department of Biochemical Sciences of the University of Rome.

[0081] FIG. 33 shows a structural representation of the location of XTEN insertion sites. The central drawing corresponding to the crystal structure of FVIII (PDB: 2R7E) is surrounded by detailed view of domains A1, A2, A3, C1 and C2. Beta strands and alpha helices are shown as ribbon representation. Loops are shown as alpha carbon pipes. The amino acids at XTEN insertion sites are shown as CPK sphere representation. The number in each graph indicate the location of the XTEN insertion sites according to the numbering in FIG. 30.

[0082] FIG. 34 shows a structural representation of the location of XTEN insertion sites shown in FIG. 33 wherein the resulting recombinant FVIII protein displays FVIII activity.

[0083] FIG. 35 shows a structural representation of the location of XTEN insertion sites shown in FIG. 34 wherein the resulting recombinant FVIII protein displays FVIII activity.

[0084] FIG. 36 shows a structural representation of the location of XTEN insertion sites shown in FIG. 35 wherein the resulting recombinant FVIII protein displays FVIII activity.

[0085] FIG. 37 shows a ClustalW multiple sequence alignment of domains A1, A2, A3, C1 and C2 of FVIII showing the location of XTEN insertions resulting in recombinant FVIII proteins displaying FVIII activity (black box, white text) or displaying no FVIII activity (grey box, bold text).

[0086] FIG. 38 shows a DSSP graphical representation of the secondary structure of the two polypeptide chains in a native active human FVIII crystal structure deposited under the identifier 2R7E at the Protein Data Bank (see Example 47). Amino acid sequence numbering is the same as in the protein sequence in FIG. 30. The beta sheet regions are shown as filled arrows and are designated β1 to β66. The location of the XTEN permissive loops is denoted by crosshatched boxes. Domain A1 XTEN permissive loops are designated Loop A1-1 and Loop A1-2. Domain A2 XTEN permissive loops are designated Loop A2-1 and Loop A2-2. Domain A3 XTEN permissive loops are designated Loop A3-1 and Loop A3-2.

[0087] FIG. 39 shows a DSSP graphical representation of the secondary structure of the two polypeptide chains in a native active human FVIII crystal structure deposited under the identifier 2R7E at the Protein Data Bank (see Example 47). Amino acid sequence numbering is the same as in the protein sequence in FIG. 30. The beta sheet regions are shown as filled arrows and are designated β1 to β66. The location of the XTEN permissive loops is denoted by crosshatched boxes. Domain A1 XTEN permissive loops are designated Loop A1-1 and Loop A1-2. Domain A2 XTEN permissive loops are designated Loop A2-1 and Loop A2-2. Domain A3 XTEN permissive loops are designated Loop A3-1 and Loop A3-2.

[0088] FIG. 40 shows a ClustalW multiple sequence alignment of domains A1, A2, A3, C1 and C2 of FVIII showing the location of XTEN insertions resulting in recombinant FVIII proteins displaying FVIII activity (black box, white text) or displaying no FVIII activity (grey box, bold text). The locations of the XTEN permissive loops are indicated by dashed rectangles (see Example 47).

[0089] FIG. 41. FIG. 41A presents a front view structural representation of human FVIII (PDB:2R7E) showing the location of domains A1, A2, A3, C1 and C2 (circled in dashed lined) and the locations of XTEN permissive loops A1-1, A1-2, A2-1, A2-2, A3-1 and A3-2 highlighted as CPK sphere representations. FIG. 41B presents a side view structural representation of human FVIII (PDB:2R7E) showing the location of domains A1, A2, A3, C1 and C2 (circled in dashed lined) and the locations of XTEN permissive loops A1-1, A1-2, A2-1, A2-2, A3-1 and A3-2 highlighted as CPK sphere representations.

[0090] FIG. 42 shows the top view structural representations of isolated human FVIII (PDB:2R7E) A domains showing the location of XTEN permissive loops highlighted as CPK sphere representations. FIG. 42B, 42D and 42F show side view structural representations of isolated human FVIII (PDB:2R7E) A domains showing the location of XTEN permissive loops highlighted as CPK sphere representations.

[0091] FIG. 43 shows sequences of various factor VIII B-domain deletions and individual mutations. Lines 4-10 show various B-domain deletions with indicated XTEN linking the flanking B-domain

residual or A3 domain residues. The R1648A mutation is indicated by arrow in line 5 and 8, while the Y1680F mutation is indicated by arrow in lines 8-10.

[0092] FIG. 44 is a bar graph of chromogenic and aPTT assay activity of various CFXTEN with single XTEN insertions (Example 49).

[0093] FIG. 45 is a bar graph of chromogenic and aPTT assay activity of various CFXTEN with 2 XTEN insertions (Example 49).

[0094] FIG 46 is a bar graph of chromogenic and aPTT assay activity of various CFXTEN with 3 XTEN insertions (Example 49).

[0095] FIG. 47 is a graph of plasma levels in DKO mice of various administered CFXTEN with single XTEN insertions compared to a BDD-FVIII control, demonstrating the 10- to 20-fold longer half-life achieved by the XTEN insertions at various locations (Example 50).

[0096] FIG. 48 is a graph of plasma levels in DKO mice of various administered CFXTEN with one, two, and three XTEN insertions compared to a BDD-FVIII control, demonstrating the increases in half-life achieved by the inclusion of additional XTEN insertions compared to single or two insertions (Example 51).

[0097] FIG. 49 are graphs of the plotted inhbition curves for remaining factor VIII procoagulant activity in samples assayed in the Bethesda assay with three hemophilia patient sera (FIGS. 49A-C) or sheep anti FVII (FIG. 49D) described in Example 52, demonstrating a clear left shift of the inhibition curve for the two CFXTEN molecules compared to the FVIII not linked to XTEN.

DETAILED DESCRIPTION OF THE INVENTION

[0098] Before the embodiments of the invention are described, it is to be understood that such embodiments are provided by way of example only, and that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention.

[0099] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. In case of conflict, the patent specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention.

DEFINITIONS

[00100] In the context of the present application, the following terms have the meanings ascribed to them unless specified otherwise:

[00101] As used in the specification and claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise. For example, the term "a cell" includes a plurality of cells, including mixtures thereof.

[00102] The terms "polypeptide", "peptide", and "protein" are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified, for example, by disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, such as conjugation with a labeling component.

[00103] As used herein, the term "amino acid" refers to either natural and/or unnatural or synthetic amino acids, including but not limited to both the D or L optical isomers, and amino acid analogs and peptidomimetics. Standard single or three letter codes are used to designate amino acids.

[00104] The term "domain," when used in reference to a factor VIII polypeptide refers to either a full length domain or a functional fragment thereof, for example, full length or functional fragments of the A1 domain, A2 domain, A3 domain, B domain, C1 domain, and/or C2 domain of factor VIII.

[00105] The term "natural L-amino acid" means the L optical isomer forms of glycine (G), proline (P), alanine (A), valine (V), leucine (L), isoleucine (I), methionine (M), cysteine (C), phenylalanine (F), tyrosine (Y), tryptophan (W), histidine (H), lysine (K), arginine (R), glutamine (Q), asparagine (N), glutamic acid (E), aspartic acid (D), serine (S), and threonine (T).

[00106] The term "non-naturally occurring," as applied to sequences and as used herein, means polypeptide or polynucleotide sequences that do not have a counterpart to, are not complementary to, or do not have a high degree of homology with a wild-type or naturally-occurring sequence found in a mammal. For example, a non-naturally occurring polypeptide or fragment may share no more than 99%, 98%, 95%, 90%, 80%, 70%, 60%, 50% or even less amino acid sequence identity as compared to a natural sequence when suitably aligned.

[00107] The terms "hydrophilic" and "hydrophobic" refer to the degree of affinity that a substance has with water. A hydrophilic substance has a strong affinity for water, tending to dissolve in, mix with, or be wetted by water, while a hydrophobic substance substantially lacks affinity for water, tending to repel and not absorb water and tending not to dissolve in or mix with or be wetted by water. Amino acids can be characterized based on their hydrophobicity. A number of scales have been developed. An example is a scale developed by Levitt, M, et al., J Mol Biol (1976) 104:59, which is listed in Hopp, TP, et al., Proc Natl Acad Sci U S A (1981) 78:3824. Examples of "hydrophilic amino acids" are arginine, lysine, threonine, alanine, asparagine, and glutamine. Of particular interest are the hydrophilic amino acids aspartate, glutamate, and serine, and glycine. Examples of "hydrophobic amino acids" are tryptophan, tyrosine, phenylalanine, methionine, leucine, isoleucine, and valine.

[00108] A "fragment" when applied to a protein, is a truncated form of a native biologically active protein that retains at least a portion of the therapeutic and/or biological activity. A "variant". when applied to a protein is a protein with sequence homology to the native biologically active protein that

retains at least a portion of the therapeutic and/or biological activity of the biologically active protein. For example, a variant protein may share at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% amino acid sequence identity compared with the reference biologically active protein. As used herein, the term "biologically active protein moiety" includes proteins modified deliberately, as for example, by site directed mutagenesis, synthesis of the encoding gene, insertions, or accidentally through mutations.

[00109] The term "sequence variant" means polypeptides that have been modified compared to their native or original sequence by one or more amino acid insertions, deletions, or substitutions. Insertions may be located at either or both termini of the protein, and/or may be positioned within internal regions of the amino acid sequence. A non-limiting example is insertion of an XTEN sequence within the sequence of the biologically-active payload protein. In deletion variants, one or more amino acid residues in a polypeptide as described herein are removed. Deletion variants, therefore, include all fragments of a payload polypeptide sequence. In substitution variants, one or more amino acid residues of a polypeptide are removed and replaced with alternative residues. In one aspect, the substitutions are conservative in nature and conservative substitutions of this type are well known in the art.

[00110] As used herein, "internal XTEN" refers to XTEN sequences that have been inserted into the sequence of the coagulation factor. Internal XTENs can be constructed by insertion of an XTEN sequence into the sequence of a coagulation factor such as FVIII, either by insertion between two adjacent amino acids within a domain ("intradomain") or between two domains ("interdomain") of the coagulation factor or wherein XTEN replaces a partial, internal sequence of the coagulation factor.

[00111] As used herein, "terminal XTEN" refers to XTEN sequences that have been fused to or in the N- or C-terminus of the coagulation factor or to a proteolytic cleavage sequence or linker at the N- or C-terminus of the coagulation factor. Terminal XTENs can be fused to the native termini of the coagulation factor. Alternatively, terminal XTENs can replace a portion of a terminal sequence of the coagulation factor.

[00112] The term "XTEN release site" refers to a cleavage sequence in CFXTEN fusion proteins that can be recognized and cleaved by a mammalian protease, effecting release of an XTEN or a portion of an XTEN from the CFXTEN fusion protein. As used herein, "mammalian protease" means a protease that normally exists in the body fluids, cells or tissues of a mammal. XTEN release sites can be engineered to be cleaved by various mammalian proteases (a.k.a. "XTEN release proteases") such as FXIa, FXIIa, kallikrein, FVIIIa, FVIIIa, FXa, FIIa (thrombin), Elastase-2, MMP-12, MMP13, MMP-17, MMP-20, or any protease that is present during a clotting event. Other equivalent proteases (endogenous or exogenous) that are capable of recognizing a defined cleavage site can be utilized. The cleavage sites can be adjusted and tailored to the protease utilized.

[00113] The term "within", when referring to a first polypeptide being linked to a second polypeptide, encompasses linking that connects the N-terminus of the first or second polypeptide to the C-terminus of the second or first polypeptide, respectively, as well as insertion of the first polypeptide into the sequence of the second polypeptide. For example, when an XTEN is linked "within" a domain of a factor VIII

polypeptide, the XTEN may be linked to the N-terminus, the C-terminus, or may be inserted in said domain.

[00114] As used herein, the term "site," when used to refer to an insertion site of an XTEN within or to a biological polypeptide such as a factor VIII, represents the amino acid position at which the XTEN is linked. When numbered sites are described, such as a first, second, third, fourth, fifth, or sixth site for the insertion of an XTEN within or to the factor VIII, each site will be understood to represent a distinct site in the factor VIII; e.g., the second site is a different factor VIII location from the first site, the third site is different from the second and the first, etc.

[00115] "Activity" or "procoagulant activity" as applied to form(s) of a CFXTEN polypeptide provided herein, refers to the ability to bind to a target coagulation protein substrate or cofactor and promote a clotting event, whether measured by an *in vitro*, *ex vivo* or *in vivo* assay. Such assays include, but are not limited to, one-stage clotting assays, two-stage clotting assays, chromogenic assays, and ELISA assays. "Biological activity" refers to an *in vitro* or *in vivo* biological function or effect, including but not limited to either receptor or ligand binding, or an effect on coagulation generally known in the art for the FVIII coagulation factor, or a cellular, physiologic, or clinical response, including arrest of a bleeding episode.

[00116] As used herein, the term "ELISA" refers to an enzyme-linked immunosorbent assay as described herein or as otherwise known in the art.

[00117] A "host cell" includes an individual cell or cell culture which can be or has been a recipient for the subject vectors. Host cells include progeny of a single host cell. The progeny may not necessarily be completely identical (in morphology or in genomic of total DNA complement) to the original parent cell due to natural, accidental, or deliberate mutation. A host cell includes cells transfected in vivo with a vector of this invention.

[00118] "Isolated" when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. As is apparent to those of skill in the art, a non-naturally occurring polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, does not require "isolation" to distinguish it from its naturally occurring counterpart. In addition, a "concentrated", "separated" or "diluted" polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, is distinguishable from its naturally occurring counterpart in that the concentration or number of molecules per volume is generally greater than that of its naturally occurring counterpart. In general, a polypeptide made by recombinant means and expressed in a host cell is considered to be "isolated."

[00119] An "isolated" polynucleotide or polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated polypeptide-encoding nucleic acid

molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal or extra-chromosomal location different from that of natural cells.

[00120] A "chimeric" protein contains at least one fusion polypeptide comprising at least one region in a different position in the sequence than that which occurs in nature. The regions may normally exist in separate proteins and are brought together in the fusion polypeptide; or they may normally exist in the same protein but are placed in a new arrangement in the fusion polypeptide. A chimeric protein may be created, for example, by chemical synthesis, or by creating and translating a polynucleotide in which the peptide regions are encoded in the desired relationship.

[00121] "Conjugated", "linked," "fused," and "fusion" are used interchangeably herein. These terms refer to the joining together of two or more chemical elements, sequences or components, by whatever means including chemical conjugation or recombinant means. For example, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and in reading phase or in-frame. An "in-frame fusion" refers to the joining of two or more open reading frames (ORFs) to form a continuous longer ORF, in a manner that maintains the correct reading frame of the original ORFs. Thus, the resulting recombinant fusion protein is a single protein containing two or more segments that correspond to polypeptides encoded by the original ORFs (which segments are not normally so joined in nature).

[00122] In the context of polypeptides, a "linear sequence" or a "sequence" is an order of amino acids in a polypeptide in an amino to carboxyl terminus direction in which residues that neighbor each other in the sequence are contiguous in the primary structure of the polypeptide. A "partial sequence" is a linear sequence of part of a polypeptide that is known to comprise additional residues in one or both directions.

[00123] "Heterologous" means derived from a genotypically distinct entity from the rest of the entity to which it is being compared. For example, a glycine rich sequence removed from its native coding sequence and operatively linked to a coding sequence other than the native sequence is a heterologous glycine rich sequence. The term "heterologous" as applied to a polynucleotide, a polypeptide, means that the polynucleotide or polypeptide is derived from a genotypically distinct entity from that of the rest of the entity to which it is being compared.

[00124] The terms "polynucleotides", "nucleic acids", "nucleotides" and "oligonucleotides" are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three-dimensional structure, and may perform any function, known or unknown. The following are non-limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors,

isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component.

[00125] The term "complement of a polynucleotide" denotes a polynucleotide molecule having a complementary base sequence and reverse orientation as compared to a reference sequence, such that it could hybridize with a reference sequence with complete fidelity.

[00126] "Recombinant" as applied to a polynucleotide means that the polynucleotide is the product of various combinations of *in vitro* cloning, restriction and/or ligation steps, and other procedures that result in a construct that can potentially be expressed as a recombinant protein in a host cell.

[00127] The terms "gene" and "gene fragment" are used interchangeably herein. They refer to a polynucleotide containing at least one open reading frame that is capable of encoding a particular protein after being transcribed and translated. A gene or gene fragment may be genomic or cDNA, as long as the polynucleotide contains at least one open reading frame, which may cover the entire coding region or a segment thereof. A "fusion gene" is a gene composed of at least two heterologous polynucleotides that are linked together.

[00128] "Homology" or "homologous" or "sequence identity" refers to sequence similarity or interchangeability between two or more polynucleotide sequences or between two or more polypeptide sequences. When using a program such as BestFit to determine sequence identity, similarity or homology between two different amino acid sequences, the default settings may be used, or an appropriate scoring matrix, such as blosum45 or blosum80, may be selected to optimize identity, similarity or homology scores. Preferably, polynucleotides that are homologous are those which hybridize under stringent conditions as defined herein and have at least 70%, preferably at least 80%, more preferably at least 90%, more preferably 95%, more preferably 97%, more preferably 98%, and even more preferably 99% sequence identity compared to those sequences. Polypeptides that are homologous preferably have sequence identities that are at least 70%, preferably at least 80%, even more preferably at least 90%, even more preferably at least 95-99%, and most preferably 100% identical.

[00129] "Ligation" refers to the process of forming phosphodiester bonds between two nucleic acid fragments or genes, linking them together. To ligate the DNA fragments or genes together, the ends of the DNA must be compatible with each other. In some cases, the ends will be directly compatible after endonuclease digestion. However, it may be necessary to first convert the staggered ends commonly produced after endonuclease digestion to blunt ends to make them compatible for ligation.

[00130] The terms "stringent conditions" or "stringent hybridization conditions" includes reference to conditions under which a polynucleotide will hybridize to its target sequence, to a detectably greater degree than other sequences (e.g., at least 2-fold over background). Generally, stringency of hybridization is expressed, in part, with reference to the temperature and salt concentration under which

the wash step is carried out. Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short polynucleotides (e.g., 10 to 50 nucleotides) and at least about 60°C for long polynucleotides (e.g., greater than 50 nucleotides)—for example, "stringent conditions" can include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and three washes for 15 min each in 0.1×SSC/1% SDS at 60°C to 65°C. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2×SSC, with SDS being present at about 0.1%. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating Tm and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. *et al.*, "Molecular Cloning: A Laboratory Manual," 3rd edition, Cold Spring Harbor Laboratory Press, 2001. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art.

[00131] The terms "percent identity," percentage of sequence identity," and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences. Percent identity may be measured over the length of an entire defined polynucleotide sequence, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polynucleotide sequence, for instance, a fragment of at least 45, at least 60, at least 90, at least 120, at least 150, at least 210 or at least 450 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured. The percentage of sequence identity is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of matched positions (at which identical residues occur in both polypeptide sequences), dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. When sequences of different length are to be compared, the shortest sequence defines the length of the window of comparison. Conservative substitutions are not considered when calculating sequence identity.

[00132] "Percent (%) sequence identity," with respect to the polypeptide sequences identified herein, is defined as the percentage of amino acid residues in a query sequence that are identical with the amino

acid residues of a second, reference polypeptide sequence or a portion thereof, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Percent identity may be measured over the length of an entire defined polypeptide sequence, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

[00133] The term "non-repetitiveness" as used herein in the context of a polypeptide refers to a lack or limited degree of internal homology in a peptide or polypeptide sequence. The term "substantially non-repetitive" can mean, for example, that there are few or no instances of four contiguous amino acids in the sequence that are identical amino acid types or that the polypeptide has a subsequence score (defined infra) of 10 or less or that there is no a pattern in the order, from N- to C-terminus, of the sequence motifs that constitute the polypeptide sequence. The term "repetitiveness" as used herein in the context of a polypeptide refers to the degree of internal homology in a peptide or polypeptide sequence. In contrast, a "repetitive" sequence may contain multiple identical copies of short amino acid sequences. For instance, a polypeptide sequence of interest may be divided into n-mer sequences and the number of identical sequences can be counted. Highly repetitive sequences contain a large fraction of identical sequences while non-repetitive sequences contain few identical sequences. In the context of a polypeptide, a sequence can contain multiple copies of shorter sequences of defined or variable length, or motifs, in which the motifs themselves have non-repetitive sequences, rendering the full-length polypeptide substantially non-repetitive. The length of polypeptide within which the non-repetitiveness is measured can vary from 3 amino acids to about 200 amino acids, about from 6 to about 50 amino acids, or from about 9 to about 14 amino acids. "Repetitiveness" used in the context of polynucleotide sequences refers to the degree of internal homology in the sequence such as, for example, the frequency of identical nucleotide sequences of a given length. Repetitiveness can, for example, be measured by analyzing the frequency of identical sequences.

[00134] A "vector" is a nucleic acid molecule, preferably self-replicating in an appropriate host, which transfers an inserted nucleic acid molecule into and/or between host cells. The term includes vectors that function primarily for insertion of DNA or RNA into a cell, replication of vectors that function primarily for the replication of DNA or RNA, and expression vectors that function for transcription and/or translation of the DNA or RNA. Also included are vectors that provide more than one of the above

functions. An "expression vector" is a polynucleotide which, when introduced into an appropriate host cell, can be transcribed and translated into a polypeptide(s). An "expression system" usually connotes a suitable host cell comprised of an expression vector that can function to yield a desired expression product.

[00135] "Serum degradation resistance," as applied to a polypeptide, refers to the ability of the polypeptides to withstand degradation in blood or components thereof, which typically involves proteases in the serum or plasma. The serum degradation resistance can be measured by combining the protein with human (or mouse, rat, monkey, as appropriate) serum or plasma, typically for a range of days (e.g. 0.25, 0.5, 1, 2, 4, 8, 16 days), typically at about 37°C. The samples for these time points can be run on a Western blot assay and the protein is detected with an antibody. The antibody can be to a tag in the protein. If the protein shows a single band on the western, where the protein's size is identical to that of the injected protein, then no degradation has occurred. In this exemplary method, the time point where 50% of the protein is degraded, as judged by Western blots or equivalent techniques, is the serum degradation half-life or "serum half-life" of the protein.

[00136] The term "$t_{1/2}$" as used herein means the terminal half-life calculated as $\ln(2)/K_{el}$. K_{el} is the terminal elimination rate constant calculated by linear regression of the terminal linear portion of the log concentration vs. time curve. Half-life typically refers to the time required for half the quantity of an administered substance deposited in a living organism to be metabolized or eliminated by normal biological processes. The terms "$t_{1/2}$", "terminal half-life", "elimination half-life" and "circulating half-life" are used interchangeably herein.

[00137] "Active clearance" means the mechanisms by which a protein is removed from the circulation other than by filtration or coagulation, and which includes removal from the circulation mediated by cells, receptors, metabolism, or degradation of the protein.

[00138] "Apparent molecular weight factor" and "apparent molecular weight" are related terms referring to a measure of the relative increase or decrease in apparent molecular weight exhibited by a particular amino acid sequence. The apparent molecular weight is determined using size exclusion chromatography (SEC) or similar methods by comparing to globular protein standards, and is measured in "apparent kD" units. The apparent molecular weight factor is the ratio between the apparent molecular weight and the actual molecular weight; the latter predicted by adding, based on amino acid composition, the calculated molecular weight of each type of amino acid in the composition or by estimation from comparison to molecular weight standards in an SDS electrophoresis gel.

[00139] The terms "hydrodynamic radius" or "Stokes radius" is the effective radius (R_h in nm) of a molecule in a solution measured by assuming that it is a body moving through the solution and resisted by the solution's viscosity. In the embodiments of the invention, the hydrodynamic radius measurements of the XTEN fusion proteins correlate with the 'apparent molecular weight factor', which is a more intuitive measure. The "hydrodynamic radius" of a protein affects its rate of diffusion in aqueous solution as well as its ability to migrate in gels of macromolecules. The hydrodynamic radius of a protein is determined by its molecular weight as well as by its structure, including shape and

compactness. Methods for determining the hydrodynamic radius are well known in the art, such as by the use of size exclusion chromatography (SEC), as described in U.S. Patent Nos. 6,406,632 and 7,294,513. Most proteins have globular structure, which is the most compact three-dimensional structure a protein can have with the smallest hydrodynamic radius. Some proteins adopt a random and open, unstructured, or 'linear' conformation and as a result have a much larger hydrodynamic radius compared to typical globular proteins of similar molecular weight.

[00140] "Physiological conditions" refers to a set of conditions in a living host as well as *in vitro* conditions, including temperature, salt concentration, pH, that mimic those conditions of a living subject. A host of physiologically relevant conditions for use in *in vitro* assays have been established. Generally, a physiological buffer contains a physiological concentration of salt and is adjusted to a neutral pH ranging from about 6.5 to about 7.8, and preferably from about 7.0 to about 7.5. A variety of physiological buffers are listed in Sambrook et al. (2001). Physiologically relevant temperature ranges from about 25^{0}C to about 38^{0}C, and preferably from about 35^{0}C to about 37^{0}C.

[00141] A "reactive group" is a chemical structure that can be coupled to a second reactive group. Examples for reactive groups are amino groups, carboxyl groups, sulfhydryl groups, hydroxyl groups, aldehyde groups, azide groups. Some reactive groups can be activated to facilitate coupling with a second reactive group. Non-limiting examples for activation are the reaction of a carboxyl group with carbodiimide, the conversion of a carboxyl group into an activated ester, or the conversion of a carboxyl group into an azide function.

[00142] "Controlled release agent", "slow release agent", "depot formulation" and "sustained release agent" are used interchangeably to refer to an agent capable of extending the duration of release of a polypeptide of the invention relative to the duration of release when the polypeptide is administered in the absence of agent. Different embodiments of the present invention may have different release rates, resulting in different therapeutic amounts.

[00143] The terms "antigen", "target antigen" and "immunogen" are used interchangeably herein to refer to the structure or binding determinant that an antibody fragment or an antibody fragment-based therapeutic binds to or has specificity against.

[00144] The term "payload" as used herein refers to a protein or peptide sequence that has biological or therapeutic activity; the counterpart to the pharmacophore of small molecules. Examples of payloads include, but are not limited to, coagulation factors, cytokines, enzymes, hormones, and blood and growth factors.

[00145] The term "antagonist", as used herein, includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native polypeptide disclosed herein. Methods for identifying antagonists of a polypeptide may comprise contacting a native polypeptide with a candidate antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the native polypeptide. In the context of the present invention, antagonists may include proteins, nucleic acids, carbohydrates, antibodies or any other molecules that decrease the effect of a biologically active protein.

[00146] The term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native polypeptide disclosed herein. Suitable agonist molecules specifically include agonist antibodies or antibody fragments, fragments or amino acid sequence variants of native polypeptides, peptides, small organic molecules, etc. Methods for identifying agonists of a native polypeptide may comprise contacting a native polypeptide with a candidate agonist molecule and measuring a detectable change in one or more biological activities normally associated with the native polypeptide.

[00147] As used herein, "treat" or "treating," or "palliating" or "ameliorating" are used interchangeably and mean administering a drug or a biologic to achieve a therapeutic benefit, to cure or reduce the severity of an existing condition, or to achieve a prophylactic benefit, prevent or reduce the likelihood of onset or severity the occurrence of a condition. By therapeutic benefit is meant eradication or amelioration of the underlying condition being treated or one or more of the physiological symptoms associated with the underlying condition such that an improvement is observed in the subject, notwithstanding that the subject may still be afflicted with the underlying condition.

[00148] A "therapeutic effect" or "therapeutic benefit," as used herein, refers to a physiologic effect, including but not limited to the mitigation, amelioration, or prevention of disease in humans or other animals, or to otherwise enhance physical or mental wellbeing of humans or animals, resulting from administration of a fusion protein of the invention other than the ability to induce the production of an antibody against an antigenic epitope possessed by the biologically active protein. For prophylactic benefit, the compositions may be administered to a subject at risk of developing a particular disease, condition or symptom of the disease (e.g., a bleed in a diagnosed hemophilia A subject), or to a subject reporting one or more of the physiological symptoms of a disease, even though a diagnosis of this disease may not have been made.

[00149] The terms "therapeutically effective amount" and "therapeutically effective dose", as used herein, refer to an amount of a drug or a biologically active protein, either alone or as a part of a fusion protein composition, that is capable of having any detectable, beneficial effect on any symptom, aspect, measured parameter or characteristics of a disease state or condition when administered in one or repeated doses to a subject. Such effect need not be absolute to be beneficial. Determination of a therapeutically effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein.

[00150] The term "therapeutically effective dose regimen", as used herein, refers to a schedule for consecutively administered multiple doses (i.e., at least two or more) of a biologically active protein, either alone or as a part of a fusion protein composition, wherein the doses are given in therapeutically effective amounts to result in sustained beneficial effect on any symptom, aspect, measured parameter or characteristics of a disease state or condition.

l). GENERAL TECHNIQUES

[00151] The practice of the present invention employs, unless otherwise indicated, conventional techniques of immunology, biochemistry, chemistry, molecular biology, microbiology, cell biology,

genomics and recombinant DNA, which are within the skill of the art. *See* Sambrook, J. *et al.*, "Molecular Cloning: A Laboratory Manual," 3rd edition, Cold Spring Harbor Laboratory Press, 2001; "Current protocols in molecular biology", F. M. Ausubel, *et al.* eds.,1987; the series "Methods in Enzymology," Academic Press, San Diego, CA.; "PCR 2: a practical approach", M.J. MacPherson, B.D. Hames and G.R. Taylor eds., Oxford University Press, 1995; "Antibodies, a laboratory manual" Harlow, E. and Lane, D. eds., Cold Spring Harbor Laboratory,1988; "Goodman & Gilman's The Pharmacological Basis of Therapeutics," 11th Edition, McGraw-Hill, 2005; and Freshney, R.I., "Culture of Animal Cells: A Manual of Basic Technique," 4th edition, John Wiley & Sons, Somerset, NJ, 2000.

l). COAGULATION FACTOR VIII

[00152] The present invention relates, in part, to compositions comprising factor VIII coagulation factor (CF) linked to one or more extended recombinant proteins (XTEN), resulting in a CFXTEN fusion protein composition. As used herein, "CF" refers to factor VIII (FVIII) or mimetics, sequence variants and truncated versions of FVIII, as described below.

[00153] "Factor VIII" or "FVIII" or "FVIII protein" means a blood coagulation factor protein and species (including human, porcine, canine, rat or murine FVIII proteins) and sequence variants thereof that includes, but is not limited to the 2351 amino acid single-chain precursor protein (with a 19-amino acid hydrophobic signal peptide), the mature 2332 amino acid factor VIII cofactor protein of approximately 270-330 kDa with the domain structure A1-A2-B-A3-C1-C2, as well as the nonenzymatic "active" or cofactor form of FVIII (FVIIIa) that is a circulating heterodimer of two chains that form as a result of proteolytic cleavage after R1648 of a heavy chain form composed of A1-A2-B (in the range of 90-220 kD) of amino acids 1-1648 (numbered relative to the mature FVIII form) and a light chain A3-C1-C2 of 80 kDa of amino acids 1649-2232, each of which is depicted schematically in FIG. 1. Further, and as used herein, each of A1, A2 and the A3 domain encompasses acidic spacer regions; a1, a2, and a3 acidic regions, respectively. Thus, it will be understood that CFXTEN constructs described as having A1, A2, A3, B, C1 and C2 domains include the a1, a2 and a3 acidic regions. As used herein, "Factor VIII" or "FVIII" or "FVIII polypeptide" also includes variant forms, including proteins with substitutions, additions and/or deletions so long as the variant retains a desired biological activity such as procoagulant activity. Myriad functional FVIII variants have been constructed and can be used as recombinant FVIII proteins as described herein. See PCT Publication Nos. WO 2011/069164 A2, WO 2012/006623 A2, WO 2012/006635 A2, or WO 2012/006633 A2. A great many functional FVIII variants are known. In addition, hundreds of nonfunctional mutations in FVIII have been identified in hemophilia patients. See, e.g., Cutler et al., Hum. Mutat. 19:274-8 (2002). In addition, comparisons between FVIII from humans and other species have identified conserved residues that are likely to be required for function. See, e.g., Cameron et al., Thromb. Haemost. 79:317-22 (1998) and US 6,251,632.

[00154] In one embodiment, the human factor VIII domains are defined by the following amino acid residues: A1, residues Ala1-Arg372; A2, residues Ser373-Arg740; B, residues Ser741-Arg1648; A3, residues

Ser1649-Asn2019; C1, residues Lys2020-Asn2172; C2, residues Ser2173-Tyr2332. The A3-C1-C2 sequence includes residues Ser1649-Tyr2332. In another embodiment, residues Arg336-Arg372 is usually referred to as the a1 region, and the Arg372 is cleaved by thrombin. In certain embodiments, the a2 region is part of the A1 domain. In another embodiment, residues Glu1649-Arg1689, is referred to as the a3 acidic region. In certain embodiments, the a3 acidic region is a part of the A3 domain. In another embodiment, a native FVIII protein has the following formula: A1-a1-A2-a2-B-a3-A3-C1-C2, where A1, A2, and A3 are the structurally-related "A domains," B is the "B domain," C1 and C2 are the structurally-related "C domains," and a1, a2 and a3 are acidic spacer regions. In the foregoing formula and referring to the primary amino acid sequence position in FIG. 30, the A1 domain of human FVIII extends from Ala1 to about Arg336, the a1 spacer region extends from about Met337 to about Arg372, the A2 domain extends from about Ser373 to about Tyr719, the a2 spacer region extends from about Glu720 to about Arg740, the B domain extends from about Ser741 to about Arg 1648, the a3 spacer region extends from about Glu1649 to about Arg1689, the A3 domain extends from about Ser1690 to about Asn2019, the C1 domain extends from about Lys2020 to about Asn2172, and the C2 domain extends from about Ser2173 to Tyr2332 (Saenko et al., 2005, J Thromb Hemostasis, 1, 922-930). Other than specific proteolytic cleavage sites, designation of the locations of the boundaries between the domains and regions of FVIII can vary in different literature references. The boundaries noted herein are therefore designated as approximate by use of the term "about."

[00155] Such factor VIII include truncated sequences such as B-domain deleted "BDD" sequences in which a portion or the majority of the B domain sequence is deleted (such as BDD sequences disclosed or referenced in US Pat Nos. 6,818,439 and 7,632,921). An example of a BDD FVIII is REFACTO® or XYNTHA® (recombinant BDD FVIII), which comprises a first polypeptide corresponding to amino acids 1 to 743 of FIG. 30, fused to a second polypeptide corresponding to amino acids 1638 to 2332 of FIG. 30. Exemplary BDD FVIII constructs which can be used to produce recombinant proteins of the invention include, but are not limited to FVIII with a deletion of amino acids corresponding to amino acids 747-1638 of mature human FVIII (FIG. 30) (Hoeben R.C., et al. J. Biol. Chem. 265 (13): 7318-7323 (1990)), and FVIII with a deletion of amino acids corresponding to amino acids 771–1666 or amino acids 868-1562 of mature human FVIII (FIG. 30) (Meulien P., et al. Protein Eng. 2(4): 301-6 (1988)).

[00156] In addition, sequences that include heterologous amino acid insertions or substitutions (such as aspartic acid substituted for valine at position 75), or single chain FVIII (scFVIII) in which the heavy and light chains are covalently connected by a linker. As used herein, "FVIII" shall be any functional form of factor VIII molecule with the typical characteristics of blood coagulation factor VIII capable of correcting human factor VIII deficiencies when administered to such a subject, e.g., a subject with hemophilia A. FVIII or sequence variants have been isolated, characterized, and cloned, as described in U.S. Patent or Application Nos. 4,757,006; 4,965,199; 5,004,804; 5,198,349, 5,250,421; 5,919,766; 6,228,620; 6,818,439; 7,138,505; 7,632,921; and 20100081615.

[00157] Human factor VIII is encoded by a single-copy gene residing at the tip of the long arm of the X chromosome (q28). It comprises nearly 186,000 base pairs (bp) and constitutes approximately 0.1% of the X-chromosome (White, G.C. and Shoemaker, C.B., Blood (1989) 73:1-12). The human FVIII amino acid sequence was deduced from cDNA as shown in U.S. Pat. No. 4,965,199. Native mature human FVIII derived from the cDNA sequence (i.e., without the secretory signal peptide but prior to other post-translational processing) is presented as FIG. 3.

[00158] The DNA encoding the mature factor VIII mRNA is found in 26 separate exons ranging in size from 69 to 3,106 bp. The 25 intervening intron regions that separate the exons range in size from 207 to 32,400 bp. The complete gene consists of approximately 9 kb of exon and 177 kb of intron. The three repeat A domains have approximately 30% sequence homology. The B domain contains 19 of the approximately 25 predicted glycosylation sites, and the A3 domain is believed to contain a binding site for the von Willebrand factor. The tandem C domains follow the A3 domain and have approximately 37% homology to each other (White, G.C. and Shoemaker, C.B., Blood (1989) 73:1-12).

[00159] The B domain separates the A2 and A3 domains of native factor FVIII in the newly synthesized precursor single-chain molecule. The precise boundaries of the B domain have been variously reported as extending from amino acids 712 to 1648 of the precursor sequence (Wood et al., Nature (1984) 312:330-337) or amino acids 741–1648 (Pipe, SW, Haemophilia (2009) 15:1187–1196 and US Pat. No. 7,560,107) or amino acids 740-1689 (Toole, JJ. Proc. Natl. Acad. Sci. USA (1986) 83:5939-5942). As used herein, "B domain" means amino acids 741-1648 of mature factor VIII. As used herein, "FVIII B domain deletion" or "FVIII BDD" means a FVIII sequence with any, a fragment of, or all of amino acids 741 to 1648 deleted. In one embodiment, FVIII BDD variants retain remnant amino acids of the B domain from the N-terminal end ("B1" as used herein) and C-terminal end ("B2" as used herein). In one FVIII BDD variant, the B domain remnant amino acids are SFSQNPPVLKRHQR (SEQ ID NO: 1614). In one FVIII BDD variant, the B1 remnant is SFS and the B2 remnant is QNPPVLKRHQR (SEQ ID NO: 1615). In another FVIII BDD variant, the B1 remnant is SFSQN (SEQ ID NO: 1616) and the B2 remnant is PPVLKRHQR (SEQ ID NO: 1617). A "B-domain-deleted factor VIII," "FVIII BDD," or "BDD FVIII" may have the full or partial deletions disclosed in U.S. Pat. Nos. 6,316,226, 6,346,513, 7,041,635, 5,789,203, 6,060,447, 5,595,886, 6,228,620, 5,972,885, 6,048,720, 5,543,502, 5,610,278, 5,171,844, 5,112,950, 4,868,112, and 6,458,563. In some embodiments, a B-domain-deleted factor VIII sequence of the present invention comprises any one of the deletions disclosed at col. 4, line 4 to col. 5, line 28 and examples 1-5 of U.S. Pat. No. 6,316,226 (also in US 6,346,513). In another embodiment, a B-domain deleted factor VIII is the S743/Q1638 B-domain deleted factor VIII (SQ version factor VIII) (e.g., factor VIII having a deletion from amino acid 744 to amino acid 1637, e.g., factor VIII having amino acids 1-743 and amino acids 1638-2332 of full-length factor VIII). In some embodiments, a B-domain-deleted factor VIII of the present invention has a deletion disclosed at col. 2, lines 26-51 and examples 5-8 of U.S. Patent No. 5,789,203 (also US 6,060,447, US 5,595,886, and US

6,228,620). In some embodiments, a B-domain-deleted factor VIII has a deletion described in col. 1, lines 25 to col. 2, line 40 of US Patent No. 5,972,885; col. 6, lines 1-22 and example 1 of U.S. Patent no. 6,048,720; col. 2, lines 17-46 of U.S. Patent No. 5,543,502; col. 4, line 22 to col. 5, line 36 of U.S. Patent no. 5,171,844; col. 2, lines 55-68, figure 2, and example 1 of U.S. Patent No. 5,112,950; col. 2, line 2 to col. 19, line 21 and table 2 of U.S. Patent No. 4,868,112; col. 2, line 1 to col. 3, line 19, col. 3, line 40 to col. 4, line 67, col. 7, line 43 to col. 8, line 26, and col. 11, line 5 to col. 13, line 39 of U.S. Patent no. 7,041,635; or col. 4, lines 25-53, of U.S. Patent No. 6,458,563. In some embodiments, a B-domain-deleted factor VIII has a deletion of most of the B domain, but still contains amino-terminal sequences of the B domain that are essential for *in vivo* proteolytic processing of the primary translation product into two polypeptide chain, as disclosed in WO 91/09122. In some embodiments, a B-domain-deleted factor VIII is constructed with a deletion of amino acids 747-1638, *i.e.*, virtually a complete deletion of the B domain. Hoeben R.C., *et al. J. Biol. Chem.* 265 (13): 7318-7323 (1990). A B-domain-deleted factor VIII may also contain a deletion of amino acids 771–1666 or amino acids 868-1562 of factor VIII. Meulien P., *et al. Protein Eng.* 2(4): 301-6 (1988). Additional B domain deletions that are part of the invention include: deletion of amino acids 982 through 1562 or 760 through 1639 (Toole *et al., Proc. Natl. Acad. Sci. U.S.A.* (1986) 83, 5939-5942)), 797 through 1562 (Eaton, *et al. Biochemistry* (1986) 25:8343-8347)), 741 through 1646 (Kaufman (PCT published application No. WO 87/04187)), 747-1560 (Sarver, *et al., DNA* (1987) 6:553-564)), 741 though 1648 (Pasek (PCT application No.88/00831)), or 816 through 1598 or 741 through 1648 (Lagner (Behring Inst. Mitt. (1988) No 82:16-25, EP 295597)). Each of the foregoing deletions may be made in any factor VIII sequence utilized in the embodiments of the present invention.

[00160] Proteins involved in clotting include factor I, factor II, factor III, factor IV, factor V, factor VI, factor VII, factor VIII, factor IX, factor X, factor XI, factor XII, factor XIII, Protein C, and tissue factor (collectively or individually "clotting protein(s)"). The interaction of the major clotting proteins in the intrinsic and extrinsic clotting pathways is showed in FIG. 2. The majority of the clotting proteins are present in zymogen form, but when activated, exhibit a procoagulant protease activity in which they activate another of the clotting proteins, contributing to the intrinsic or extrinsic coagulation pathway and clot formation. In the intrinsic pathway of the coagulation cascade, FVIII associates with a complex of activated factor IX, factor X, calcium, and phospholipid. The factor VIII heterodimer has no enzymatic activity, but the heterodimer becomes active as a cofactor of the enzyme factor IXa after proteolytic activation by thrombin or factor Xa, with the activity of factor VIIIa characterized by its ability to form a membrane binding site for factors IXa and X in a conformation suitable for activation of the factor X by factor IXa. Upon cleavage by thrombin, activated FVIII (FVIIIa) dissociates from von Willebrand factor and binds to negatively charged phospholipid PL, and the resulting complex participates as a cofactor to factor IXa in the factor X activating (tenase) complex. Within the C2 domain and amino acid residues

1649 through 1689 in the A3 domain are von Willebrand factor (vWF) binding sites that act to complex with von Willebrand factor, the resulting circulating complex protects FVIII from rapid degradation in the blood (Weiss HJ, et al. Stabilization of factor VIII in plasma by the von Willebrand factor. Studies on posttransfusion and dissociated factor VIII and in patients with von Willebrand's disease. J Clin Invest (1977) 60:390).

[00161] Activated factor VIII is a heterotrimer comprised of the A1 domain and the A2 domain and the light chain including domains A3-C1-C2. The activation of factor IX is achieved by a two-step removal of the activation peptide (Ala 146-Arg 180) from the molecule (Bajaj et al., Human factor IX and factor IXa, in METHODS IN ENZYMOLOGY. 1993). The first cleavage is made at the Arg 145-Ala 146 site by either factor XIa or factor VIIa/tissue factor. The second, and rate limiting cleavage is made at Arg 180-Val 181. The activation removes 35 residues. Activated human factor IX exists as a heterodimer of the C-terminal heavy chain (28 kDa) and an N-terminal light chain (18 kDa), which are held together by one disulfide bridge attaching the enzyme to the Gla domain. Factor IXa in turn activates factor X in concert with activated factor VIII. Alternatively, factors IX and X can both be activated by factor VIIa complexed with lipidated tissue factor, generated via the extrinsic pathway. Factor Xa then participates in the final common pathway whereby prothrombin is converted to thrombin, and thrombin, in turn converts fibrinogen to fibrin to form the clot.

[00162] Defects in the coagulation process can lead to bleeding disorders (coagulopathies) in which the time taken for clot formation is prolonged. Such defects can be congenital or acquired. For example, hemophilia A and B are inherited diseases characterized by deficiencies in FVIII and FIX, respectively. Stated differently, biologically active factor VIII corrects the coagulation defect in plasma derived from individuals afflicted with hemophilia A. Recombinant FVIII has been shown to be effective and has been approved for the treatment of hemophilia A in adult and pediatric patients, and also is used to stop bleeding episodes or prevent bleeding associated with trauma and/or surgery. Current therapeutic uses of factor VIII can be problematic in the treatment of individuals exhibiting a deficiency in factor VIII, as well as those individuals with Von Willebrand's disease. In addition, individuals receiving factor VIII in replacement therapy frequently develop antibodies to these proteins that often reduce or eliminate the procoagulant activity of the bound FVIII. Continuing treatment is exceedingly difficult because of the presence of these antibodies that reduce or negate the efficacy of the treatment.

[00163] In one aspect, the invention contemplates inclusion of FVIII sequences in the CFXTEN fusion protein compositions that are identical to human FVIII, sequences that have homology to FVIII sequences, sequences that are natural, such as from humans, non-human primates, mammals (including domestic animals), or truncated version of FVIII; all of which retain at least a portion of the procoagulant activity of native FVIII and that are useful for preventing, treating, mediating, or ameliorating hemophilia A or bleeding episodes related to trauma, surgery, or deficiency of coagulation factor VIII. Sequences with homology to FVIII may be found by standard homology searching techniques, such as NCBI BLAST, or in public databases such as Chemical Abstracts Services Databases (e.g., the CAS Registry),

GenBank, The Universal Protein Resource (UniProt) and subscription provided databases such as GenSeq (e.g., Derwent).

[00164] In one embodiment, the FVIII incorporated into the subject CFXTEN compositions is a recombinant polypeptide with a sequence corresponding to a FVIII protein found in nature. In another embodiment, the FVIII is a non-natural FVIII sequence variant, fragment, homolog, or a mimetic of a natural sequence that retains at least a portion of the procoagulant activity of the corresponding native FVIII. In another embodiment, the FVIII is a truncated variant with all or a portion of the B domain deleted ("FVIII BDD"), which can be in either heterodimeric form or can remain as a single chain ("scFVIII"), the latter described in Meulien et al., Protein Eng. (1988) 2(4):301-306. Non-limiting examples of FVIII BDD are factor VIII sequences in which the amino acids are deleted between residue number 741 and residue number 1640 (numbered relative to native, mature FVIII), or between residue number 745 and residue number 1640, or between residue number 745 and residue number 1640, or between residue number 741 and residue number 1690, or between residue number 745 and residue number 1667, or between residue number 745 and residue number 1657, or between residue number 747 and residue number 1642, or between residue number 751 and residue number 1667.

[00165] In another embodiment, heterologous sequences are incorporated into the FVIII, which may include XTEN, as described more fully below. Table 1 provides a non-limiting list of amino acid sequences of FVIII that are encompassed by the CFXTEN fusion proteins of the invention. In some embodiments, FVIII incorporated into CFXTEN fusion proteins include proteins that have at least about 70% sequence identity, or alternatively 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to an amino acid sequence of comparable length selected from Table 1.

Table 1: FVIII amino acid sequences

Name (source)	Amino Acid Sequence	SEQ ID NO:
FVIII precursor polypeptide (human)	MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPK SFPFNTSVVYKKTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNM ASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKEN GPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLF AVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKS VYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLL FCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFD DDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNG PQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASR PYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPR CLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENR SWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWY ILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHN SDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNSRHPS TRQKQFNATTIPENDIEKTDPWFAHRTPMPKIQNVSSSDLLMLLRQSPTPHGLSLS DLQEAKYETFSDDPSPGAIDSNNSLSEMTHFRPQLHHSGDMVFTPESGLQLRLNE KLGTTAATELKKLDFKVSSTSNNLISTIPSDNLAAGTDNTSSLGPPSMPVHYDSQL DTTLFGKKSSPLTESGGPLSLSEENNDSKLLESGLMNSQESSWGKNVSSTESGRLF KGKRAHGPALLTKDNALFKVSISLLKTNKTSNNSATNRKTHIDGPSLLIENSPSVW QNILESDTEFKKVTPLIHDRMLMDKNATALRLNHMSNKTTSSKNMEMVQQKKE	1

Name (source)	Amino Acid Sequence	SEQ ID NO:
	GPIPPDAQNPDMSFFKMLFLPESARWIQRTHGKNSLNSGQGPSPKQLVSLGPEKSV EGQNFLSEKNKVVVGKGEFTKDVGLKEMVFPSSRNLFLTNLDNLHENNTHNQEK KIQEEIEKKETLIQENVVLPQIHTVTGTKNFMKNLFLLSTRQNVEGSYDGAYAPVL QDFRSLNDSTNRTKKHTAHFSKKGEEENLEGLGNQTKQIVEKYACTTRISPNTSQ QNFVTQRSKRALKQFRLPLEETELEKRIIVDDTSTQWSKNMKHLTPSTLTQIDYNE KEKGAITQSPLSDCLTRSHSIPQANRSPLPIAKVSSFPSIRPIYLTRVLFQDNSSHLPA ASYRKKDSGVQESSHFLQGAKKNNLSLAILTLEMTGDQREVGSLGTSATNSVTY KKVENTVLPKPDLPKTSGKVELLPKVHIYQKDLFPTETSNGSPGHLDLVEGSLLQ GTEGAIKWNEANRPGKVPFLRVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWK SQEKSPEKTAFKKKDTILSLNACESNHAIAAINEGQNKPEIEVTWAKQGRTERLCS QNPPVLKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQK KTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPL YRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPR KNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLL VCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPT FKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVR KKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNK CQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDL LAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVD SSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDA QITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVT GVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVN SLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY	
FVIII mature (human)	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT DHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNSRHPSTRQKQFNATTIPENDIEKTD PWFAHRTPMPKIQNVSSSDLLMLLRQSPTPHGLSLSDLQEAKYETFSDDPSPGAID SNNSLSEMTHFRPQLHHSGDMVFTPESGLQLRLNEKLGTTAATELKKLDFKVSST SNNLISTIPSDNLAAGTDNTSSLGPPSMPVHYDSQLDTTLFGKKSSPLTESGGPLSL SEENNDSKLLESGLMNSQESSWGKNVSSTESGRLFKGKRAHGPALLTKDNALFK VSISLLKTNKTSNNSATNRKTHIDGPSLLIENSPSVWQNILESDTEFKKVTPLIHDR MLMDKNATALRLNHMSNKTTSSKNMEMVQQKKEGPIPPDAQNPDMSFFKMLFL PESARWIQRTHGKNSLNSGQGPSPKQLVSLGPEKSVEGQNFLSEKNKVVVGKGEF TKDVGLKEMVFPSSRNLFLTNLDNLHENNTHNQEKKIQEEIEKKETLIQENVVLP QIHTVTGTKNFMKNLFLLSTRQNVEGSYDGAYAPVLQDFRSLNDSTNRTKKHTA HFSKKGEEENLEGLGNQTKQIVEKYACTTRISPNTSQQNFVTQRSKRALKQFRLPL EETELEKRIIVDDTSTQWSKNMKHLTPSTLTQIDYNEKEKGAITQSPLSDCLTRSHS IPQANRSPLPIAKVSSFPSIRPIYLTRVLFQDNSSHLPAASYRKKDSGVQESSHFLQG AKKNNLSLAILTLEMTGDQREVGSLGTSATNSVTYKKVENTVLPKPDLPKTSGKV ELLPKVHIYQKDLFPTETSNGSPGHLDLVEGSLLQGTEGAIKWNEANRPGKVPFL RVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWKSQEKSPEKTAFKKKDTILSLN ACESNHAIAAINEGQNKPEIEVTWAKQGRTERLCSQNPPVLKRHQREITRTTLQSD QEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSS SPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVE DNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHH MAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEF ALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGL VMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETV EMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASG	2

Name (source)	Amino Acid Sequence	SEQ ID NO:
	QYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSL YISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHP THYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKA RLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEF LISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVH QIALRMEVLGCEAQDLY	
FVIII (Canine)	MQVELYTCCFLCLLPFSLSATRKYYLGAVELSWDYMQSDLLSALHADTSFSSRVP GSLPLTTSVTYRKTVFVEFTDDLFNIAKPRPPWMGLLGPTIQAEVYDTVVIVLKN MASHPVSLHAVGVSYWKASEGAEYEDQTSQKEKEDDNVIPGESHTYVWQVLKE NGPMASDPPCLTYSYFSHVDLVKDLNSGLIGALLVCKEGSLAKERTQTLQEFVLL FAVFDEGKSWHSETNASLTQAEAQHELHTINGYVNRSLPGLTVCHKRSVYWHVI GMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTFLMDLGQFLLFCHIPSH QHDGMEAYVKVDSCPEEPQLRMKNNEDKDYDDGLYDSDMDVVSFDDDSSSPFI QIRSVAKKHPKTWVHYIAAEEEDWDYAPSGPTPNDRSHKNLYLNNGPQRIGKKY KKVRFVAYTDETFKTREAIQYESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGI NYVTPLHTGRLPKGVKHLKDMPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSS FINLERDLASGLIGPLLICYKESVDQRGNQMMSDKRNVILFSVFDENRSWYLTEN MQRFLPNADVVQPHDPEFQLSNIMHSINGYVFDNLQLSVCLHEVAYWYILSVGA QTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWVLGCHNSDFR NRGMTALLKVSSCNRNIDDYYEDTYEDIPTPLLNENNVIKPRSFSQNSRHPSTKEK QLKATTTPENDIEKIDLQSGERTQLIKAQSVSSSDLLMLLGQNPTPRGLFLSDLREA TDRADDHSRGAIERNKGPPEVASLRPELRHSEDREFTPEPELQLRLNENLGTNTTV ELKKLDLKISSSSDSLMTSPTIPSDKLAAATEKTGSLGPPNMSVHFNSHLGTIVFGN NSSHLIQSGVPLELSEEDNDSKLLEAPLMNIQESSLRENVLSMESNRLFKEERIRGP ASLIKDNALFKVNISSVKTNRAPVNLTTNRKTRVAIPTLLIENSTSVWQDIMLERN TEFKEVTSLIHNETFMDRNTTALGLNHVSNKTTLSKNVEMAHQKKEDPVPLRAE NPDLSSSKIPFLPDWIKTHGKNSLSSEQRPSPKQLTSLGSEKSVKDQNFLSEEKVVV GEDEFTKDTELQEIFPNNKSIFFANLANVQENDTYNQEKKSPEEIERKEKLTQENV ALPQAHTMIGTKNFLKNLFLLSTKQNVAGLEEQPYTPILQDTRSLNDSPHSEGIHM ANFSKIREEANLEGLGNQTNQMVERFPSTTRMSSNASQHVITQRGKRSLKQPRLS QGEIKFERKVIANDTSTQWSKNMNYLAQGTLTQIEYNEKEKRAITQSPLSDCSMR NHVTIQMNDSALPVAKESASPSVRHTDLTKIPSQHNSSHLPASACNYTFRERTSGV QEGSHFLQEAKRNNLSLAFVTLGITEGQGKFSSLGKSATNQPMYKKLENTVLLQP GLSETSDKVELLSQVHVDQEDSFPTKTSNDSPGHLDLMGKIFLQKTQGPVKMNK TNSPGKVPFLKWATESSEKIPSKLLGVLAWDNHYDTQIPSEEWKSQKKSQTNTAF KRKDTILPLGPCENNDSTAAINEGQDKPQREAMWAKQGEPGRLCSQNPPVSKHH QREITVTTLQPEEDKFEYDDTFSIEMKREDFDIYGDYENQGLRSFQKKTRHYFIAA VERLWDYGMSRSPHILRNRAQSGDVQQFKKVVFQEFTDGSFTQPLYRGELNEHL GLLGPYIRAEVEDNIVVTFKNQASRPYSFYSSLISYDEDEGQGAEPRRKFVNPNET KIYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLICRSNTLNPA HGRQVTVQEFALVFTIFDETKSWYFTENLERNCRAPCNVQKEDPTLKENFRFHAI NGYVKDTLPGLVMAQDQKVRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMA VYNLYPGVFETVEMLPSQVGIWRIECLIGEHLQAGMSTLFLVYSKKCQTPLGMAS GHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKDPFSWIKVDLLAPMIIHGI MTQGARQKFSSLYVSQFIIMYSLDGNKWHSYRGNSTGTLMVFFGNVDSSGIKHNI FNPPIIAQYIRLHPTHYSIRSTLRMELLGCDFNSCSMPLGMESKAISDAQITASSYLS SMLATWSPSQARLHLQGRTNAWRPQANNPKEWLQVDFRKTMKVTGITTQGVKS LLISMYVKEFLISSSQDGHNWTLFLQNGKVKVFQGNRDSSTPVRNRLEPPLVARY VR LHPQSWAHHIALRLEVLGCDTQQPA	
FVIII (Pig)	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT DHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL	4

Name (source)	Amino Acid Sequence	SEQ ID NO:
	IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNSRHPSTRQKQFNATTIPENDIEKTD PWFAHRTPMPKIQNVSSSDLLMLLRQSPTPHGLSLSDLQEAKYETFSDDPSPGAID SNNSLSEMTHFRPQLHHSGDMVFTPESGLQLRLNEKLGTTAATELKKLDFKVSST SNNLISTIPSDNLAAGTDNTSSLGPPSMPVHYDSQLDTTLFGKKSSPLTESGGPLSL SEENNDSKLLESGLMNSQESSWGKNVSSTESGRLFKGKRAHGPALLTKDNALFK VSISLLKTNKTSNNSATNRKTHIDGPSLLIENSPSVWQNILESDTEFKKVTPLIHDR MLMDKNATALRLNHMSNKTTSSKNMEMVQQKKEGPIPPDAQNPDMSFFKMLFL PESARWIQRTHGKNSLNSGQGPSPKQLVSLGPEKSVEGQNFLSEKNKVVVGKGEF TKDVGLKEMVFPSSRNLFLTNLDNLHENNTHNQEKKIQEEIEKKETLIQENVVLP QIHTVTGTKNFMKNLFLLSTRQNVEGSYDGAYAPVLQDFRSLNDSTNRTKKHTA HFSKKGEEENLEGLGNQTKQIVEKYACTTRISPNTSQQNFVTQRSKRALKQFRLPL EETELEKRIIVDDTSTQWSKNMKHLTPSTLTQIDYNEKEKGAITQSPLSDCLTRSHS IPQANRSPLPIAKVSSFPSIRPIYLTRVLFQDNSSHLPAASYRKKDSGVQESSHFLQG AKKNNLSLAILTLEMTGDQREVGSLGTSATNSVTYKKVENTVLPKPDLPKTSGKV ELLPKVHIYQKDLFPTETSNGSPGHLDLVEGSLLQGTEGAIKWNEANRPGKVPFL RVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWKSQEKSPEKTAFKKKDTILSLN ACESNHAIAAINEGQNKPEIEVTWAKQGRTERLCSQNPPVLKRHQREITRTTLQSD QEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSS SPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVE DNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHH MAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEF ALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGL VMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETV EMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASG QYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSL YISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHP THYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKA RLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEF LISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVH QIALRMEVLGCEAQDLY	
FVIII (Mouse)	AIRRYYLGAVELSWNYIQSDLLSVLHTDSRFLPRMSTSFPFNTSIMYKKTVFVEYK DQLFNIAKPRPPWMGLLGPTIWTEVHDTVVITLKNMASHPVSLHAVGVSYWKAS EGDEYEDQTSQMEKEDDKVFPGESHTYVWQVLKENGPMASDPPCLTYSYMSHV DLVKDLNSGLIGALLVCKEGSLSKERTQMLYQFVLLFAVFDEGKSWHSETNDSY TQSMDSASARDWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEIHSIF LEGHTFFVRNHRQASLEISPITFLTAQTLLIDLGQFLLFCHISSHKHDGMEAYVKV DSCPEESQWQKKNNNEEMEDYDDDLYSEMDMFTLDYDSSPFIQIRSVAKKYPKT WIHYISAEEEDWDYAPSVPTSDNGSYKSQYLSNGPHRIGRKYKKVRFIAYTDETF KTRETIQHESGLLGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVSPLHARRLPR GIKHVKDLPIHPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFINPERDLASGLIGP LLICYKESVDQRGNQMMSDKRNVILFSIFDENQSWYITENMQRFLPNAAKTQPQD PGFQASNIMHSINGYVFDSLELTVCLHEVAYWHILSVGAQTDFLSIFFSGYTFKHK MVYEDTLTLFPFSGETVFMSMENPGLWVLGCHNSDFRKRGMTALLKVSSCDKST SDYYEEIYEDIPTQLVNENNVIDPRSFFQNTNHPNTRKKKFKDSTIPKNDMEKIEPQ FEEIAEMLKVQSVSVSDMLMLLGQSHPTPHGLFLSDGQEAIYEAIHDDHSPNAIDS NEGPSKVTQLRPESHHSEKIVFTPQPGLQLRSNKSLETTIEVKWKKLGLQVSSLPS NLMTTTILSDNLKATFEKTDSSGFPDMPVHSSSKLSTTAFGKKAYSLVGSHVPLN ASEENSDSNILDSTLMYSQESLPRDNILSIENDRLLREKRFHGIALLTKDNTLFKDN VSLMKTNKTYNHSTTNEKLHTESPTSIENSTTDLQDAILKVNSEIQEVTALIHDGT LLGKNSTYLRLNHMLNRTTSTKNKDIFHRKDEDPIPQDEENTIMPFSKMLFLSESS NWFKKTNGNNSLNSEQEHSPKQLVYLMFKKYVKNQSFLSEKNKVTVEQDGFTK NIGLKDMAFPHNMSIFLTTLSNVHENGRHNQEKNIQEEIEKEALIEEKVVLPQVHE ATGSKNFLKDILILGTRQNISLYEVHVPVLQNITSINNSTNTVQIHMEHFFKRRKDK ETNSEGLVNKTREMVKNYPSQKNITTQRSKRALGQFRLSTQWLKTINCSTQCIIKQ IDHSKEMKKFITKSSLSDSSVIKSTTQTNSSDSHIVKTSAFPPIDLKRSPFQNKFSHV QASSYIYDFKTKSSRIQESNNFLKETKINNPSLAILPWNMFIDQGKFTSPGKSNTNS VTYKKRENIIFLKPTLPEESGKIELLPQVSIQEEEILPTETSHGSPGHLNLMKEVFLQ	5

Name (source)	Amino Acid Sequence	SEQ ID NO:
	KIQGPTKWNKAKRHGESIKGKTESSKNTRSKLLNHHAWDYHYAAQIPKDMWKS KEKSPEIISIKQEDTILSLRPHGNSHSIGANEKQNWPQRETTWVKQGQTQRTCSQIP PVLKRHQRELSAFQSEQEATDYDDAITIETIEDFDIYSEDIKQGPRSFQQKTRHYFI AAVERLWDYGMSTSHVLRNRYQSDNVPQFKKVVFQEFTDGSFSQPLYRGELNEH LGLLGPYIRAEVEDNIMVTFKNQASRPYSFYSSLISYKEDQRGEEPRRNFVKPNET KIYFWKVQHHMAPTEDEFDCKAWAYFSDVDLERDMHSGLIGPLLICHANTLNPA HGRQVSVQEFALLFTIFDETKSWYFTENVKRNCKTPCNFQMEDPTLKENYRFHAI NGYVMDTLPGLVMAQDQRIRWYLLSMGNNENIQSIHFSGHVFTVRKKEEYKMA VYNLYPGVFETLEMIPSRAGIWRVECLIGEHLQAGMSTLFLVYSKQCQIPLGMAS GSIRDFQITASGHYGQWAPNLARLHYSGSINAWSTKEPFSWIKVDLLAPMIVHGIK TQGARQKFSSLYISQFIIMYSLDGKKWLSYQGNSTGTLMVFFGNVDSSGIKHNSF NPPIIARYIRLHPTHSSIRSTLRMELMGCDLNSCSIPLGMESKVISDTQITASSYFTN MFATWSPSQARLHLQGRTNAWRPQVNDPKQWLQVDLQKTMKVTGIITQGVKSL FTSMFVKEFLISSSQDGHHWTQILYNGKVKVFQGNQDSSTPMMNSLDPPLLTRYL RIHPQIWEHQIALRLEILGCEAQQQY	
FVIII BDD variant (US Pat No. 7632921, SEQ ID NO: 3)	MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPK SFPFNTSVVYKKTLFVEFTVHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNM ASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKEN GPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLF AVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKS VYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLL FCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFD DDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNG PQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASR PYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPR CLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENR SWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWY ILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHN SDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVL KRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYF IAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNE HLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPN ETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTL NPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRF HAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYK MALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLG MASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMII HGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIK HNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASS YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQ GVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPL LTRYLRIHPQSWVHQIALRMEVLGCEAQDLY	6
FVIII BDD-2	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT VHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDY DDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLR NRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVT FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKD	7

Name (source)	Amino Acid Sequence	SEQ ID NO:
	EFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIF DETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQD QRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSK AGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQW APKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFII MYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIR STLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQ GRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQ DGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALR MEVLGCEAQDLY	
FVIII BDD-3 (G1648)	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT VHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQGEITRTTLQSDQEEIDY DDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLR NRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVT FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKD EFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIF DETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQD QRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSK AGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQW APKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFII MYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIR STLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQ GRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQ DGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALR MEVLGCEAQDLY	8
FVIII BDD-4	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT VHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQQSPRSFQKKTRHYFIAAVERLWDY GMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIR AEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKV QHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVT VQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMD TLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPG VFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQ ITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQ KFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYI	9

Name (source)	Amino Acid Sequence	SEQ ID NO:
	RLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWS PSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMY VKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQS WVHQIALRMEVLGCEAQDLY	
FVIII BDD-5	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT VHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK QSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFT DGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEED QRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDV HSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAP CNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIH FSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMS TLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKE PFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGT LMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLG MESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVD FQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGN QDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY	10
FVIII BDD-6	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT DHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNSRHPSTRQKQFNATTIPENDIEKTD TISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNR AQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFR NQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEF DCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDE TKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQR IRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAG IWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAP KLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMY SLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTL RMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRS NAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGH QWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEV LGCEAQDLY	11
FVIII BDD-7	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT VHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV	12

Name (source)	Amino Acid Sequence	SEQ ID NO:
	DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQSPRSFQKKTRHYFIAAVERLWDYG MSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRA EVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQ HHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTV QEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTL PGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVF ETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQIT ASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKF SSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRL HPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPS KARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVK EFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW VHQIALRMEVLGCEAQDLY	
FVIII BDD-8 precursor (US Pat. No. 6818439 SEQ ID NO: 47)	MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPK SFPFNTSVVYKKTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNM ASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKEN GPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLF AVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKS VYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLL FCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFD DDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNG PQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASR PYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPR CLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENR SWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWY ILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHN SDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVL KRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYF IAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNE HLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPN ETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTL NPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRF HAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYK MALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLG MASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMII HGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIK HNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASS YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQ GVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPL LTRYLRIHPQSWVHQIALRMEVLGCEAQDLY	13
FVIII BDD-9 mature (US Pat. No. 6818439)	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT DHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP	14

Name (source)	Amino Acid Sequence	SEQ ID NO:
	KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDY DDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLR NRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVT FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKD EFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIF DETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQD QRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSK AGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQW APKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFII MYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIR STLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQ GRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQ DGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALR MEVLGCEAQDLY	
FVIII BDD-10	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT DHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQAEITRTTLQSDQEEIDY DDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLR NRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVT FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKD EFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIF DETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQD QRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSK AGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQW APKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFII MYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIR STLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQ GRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQ DGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALR MEVLGCEAQDLY	15
FVIII BDD-11	ATRATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLF VEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSY WKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSY LSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETK NSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEV HSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAY VKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKK HPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMA YTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYS RRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDL ASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPA GVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSG YTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVS SCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQAEITRTTLQSDQ	16

Name (source)	Amino Acid Sequence	SEQ ID NO:
	EEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSS PHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVED NIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHM APTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFA LFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLV MAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVE MLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQ YGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYI SQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTH YSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARL HLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLIS SSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIA LRMEVLGCEAQDLY	
FVIII BDD-12	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT DHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDY DDTISVEMKKEDFDIFDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLR NRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVT FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKD EFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIF DETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQD QRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSK AGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQW APKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFII MYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIR STLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQ GRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQ DGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALR MEVLGCEAQDLY	17
FVIII BDD-13	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFT DHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKAS EGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHV DLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLM QDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFL EGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKV DSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKT WVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDE TFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLP KGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGL IGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQL EDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFK HKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDK NTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDY DDTISVEMKKEDFDIFDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLR NRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVT FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKD EFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIF DETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQD	18

Name (source)	Amino Acid Sequence	SEQ ID NO:
	QRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKA GIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPK LARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLD GKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMEL MGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRP QVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQ NGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY	

[00166] The present invention also contemplates CFXTEN comprising FVIII with various amino acid deletions, insertions and substitutions made in the FVIII sequences of Table 1 that retain procoagulant activity. Examples of conservative substitutions for amino acids in polypeptide sequences are shown in Table 2. In embodiments of the CFXTEN in which the sequence identity of the FVIII is less than 100% compared to a specific sequence disclosed herein, the invention contemplates substitution of any of the other 19 natural L-amino acids for a given amino acid residue of the given FVIII, which may be at any position within the sequence of the FVIII, including adjacent amino acid residues. If any one substitution results in an undesirable change in procoagulant activity, then one of the alternative amino acids can be employed and the construct protein evaluated by the methods described herein (e.g., the assays of Table 49), or using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Pat. No. 5,364,934, or using methods generally known in the art. In a preferred substitution, the FVIII component of the CFXTEN embodiments is modified by replacing the R1648 residue (numbered relative to the native mature form of FVIII) with glycine or alanine to prevent proteolytic processing to the heterodimer form. In another substitution, the FVIII component of the CFXTEN embodiments is modified by replacing the Y1680 residue (numbered relative to the native mature form of FVIII) with phenylalanine. In another embodiment, the FVIII component of the CFXTEN embodiments is modified by replacing the Y1680 residue (numbered relative to the native mature form of FVIII) with phenylalanine and the R1648 residue (numbered relative to the native mature form of FVIII) with glycine or alanine.

[00167] In one embodiment, the FVIII of the fusion protein composition has one or more amino acid substitutions designed to reduce the binding of FVIII inhibitors at epitopes recognized by the antibodies of Table 9, including but not limited to substitutions at Lys(377), Lys(466), Lys(380), Ser(488), Arg(489), Arg(490), Leu(491), Lys(493), Lys(496), His(497), Lys(499), Lys(512), Lys(523), Lys(556), Met (2199), Phe(2200), Leu(2252), Val(2223), and Lys(2227). In addition, variants can include, for instance, polypeptides wherein one or more amino acid residues are added or deleted at or near the N- or C-terminus of the full-length native amino acid sequence or of a domain of a FVIII so long as the variant retains some if not all of the procoagulant activity of the native peptide. The resulting FVIII sequences that retain at least a portion (e.g., at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or at least 95% or more) of the procoagulant activity in comparison to native circulating FVIII are considered useful for the fusion protein compositions of this invention. Examples of FVIII variants are known in the art, including those described in

US Patent and Application Nos. 6,316,226; 6,818,439; 7,632,921; 20080227691. In one embodiment, a FVIII sequence variant has an aspartic acid substituted for valine at amino acid position 75 (numbered relative to the native mature form of FVIII).

Table 2: Exemplary conservative amino acid substitutions

Original Residue	Exemplary Substitutions
Ala (A)	val; leu; ile
Arg (R)	lys; gln; asn
Asn (N)	gin; his; lys; arg
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
Gly (G)	Pro
His (H)	asn: gin: lys: arg
Ile (I)	leu; val; met; ala; phe: norleucine
Leu (L)	norleucine: ile: val; met; ala: phe
Lys (K)	arg: gin: asn
Met (M)	leu; phe; ile
Phe (F)	leu: val: ile; ala
Pro (P)	Gly
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr(Y)	Trp: phe: thr: ser
Val (V)	Ile; leu; met; phe; ala; norleucine

II). EXTENDED RECOMBINANT POLYPEPTIDES

[00168] In one aspect, the invention provides XTEN polypeptide compositions that are useful as fusion protein partner(s) to link to and/or incorporate within a FVIII polypeptide, resulting in a CFXTEN fusion protein. XTEN are generally polypeptides with non-naturally occurring, substantially non-repetitive sequences having a low degree of or no secondary or tertiary structure under physiologic conditions. XTEN typically have from about 36 to about 3000 amino acids of which the majority or the entirety are small hydrophilic amino acids. As used herein, "XTEN" specifically excludes whole antibodies or antibody fragments (e.g. single-chain antibodies and Fc fragments). XTEN polypeptides have utility as a fusion protein partners in that they serve various roles, conferring certain desirable pharmacokinetic, physicochemical, pharmacologic, and pharmaceutical properties when linked to a FVIII protein to a create a CFXTEN fusion protein. Such CFXTEN fusion protein compositions have enhanced properties compared to the corresponding FVIII not linked to XTEN, making them useful in the treatment of certain conditions related to FVIII deficiencies or bleeding disorders, as more fully described below.

[00169] The selection criteria for the XTEN to be fused to the FVIII proteins used to create the inventive fusion proteins compositions generally relate to attributes of physical/chemical properties and conformational structure of the XTEN that is, in turn, used to confer enhanced pharmaceutical, pharmacologic, and pharmacokinetic properties to the FVIII fusion proteins compositions. The unstructured characteristic and physical/chemical properties of the XTEN result, in part, from the overall amino acid composition disproportionately limited to 4-6 hydrophilic amino acids, the linking of the amino acids in a quantifiable non-repetitive design, and the length of the XTEN polypeptide. In an advantageous feature common to XTEN but uncommon to polypeptides, the properties of XTEN disclosed herein are not tied to absolute primary amino acid sequences, as evidenced by the diversity of the exemplary sequences of Table 4 that, within varying ranges of length, possess similar properties, many of which are documented in the Examples. The XTEN of the present invention may exhibit one or more, or all of the following advantageous properties: unstructured conformation, conformational flexibility, enhanced aqueous solubility, high degree of protease resistance, low immunogenicity, low binding to mammalian receptors, a defined degree of charge, and increased hydrodynamic (or Stokes) radii; properties that can make them particularly useful as fusion protein partners. Non-limiting examples of the enhanced properties that XTEN confer on the fusion proteins comprising FVIII fused to XTEN, compared to FVIII not linked to XTEN, include increases in the overall solubility and/or metabolic stability, reduced susceptibility to proteolysis, reduced immunogenicity, reduced rate of absorption when administered subcutaneously or intramuscularly, reduced binding to FVIII clearance receptors, reduced reactivity to anti-payload antibodies, enhanced interactions with substrate, and/or enhanced pharmacokinetic properties when administered to a subject. The enhanced pharmacokinetic properties of the CFXTEN compositions compared to FVIII not linked to XTEN include longer terminal half-life (e.g., two-fold, three-fold, four-fold or more), increased area under the curve (AUC) (e.g., 25%, 50%, 100% or more), lower volume of distribution, and enhanced absorption after subcutaneous or intramuscular injection (an advantage compared to commercially-available forms of FVIII that must be administered intravenously). In addition, it is believed that the CFXTEN compositions comprising cleavage sequences (described more fully, below) permit sustained release of biologically active FVIII, such that the administered CFXTEN acts as a depot. It is specifically contemplated that the inventive CFXTEN fusion proteins can exhibit one or more or any combination of the improved properties disclosed herein. As a result of these enhanced properties, it is believed that CFXTEN compositions permit less frequent dosing compared to FVIII not linked to XTEN when administered at comparable dosages. Such CFXTEN fusion protein compositions have utility to treat certain factor VIII-related conditions, as described herein.

[00170] A variety of methods and assays are known in the art for determining the physical/chemical properties of proteins such as the CFXTEN compositions comprising XTEN. Such properties include but are not limited to secondary or tertiary structure, solubility, protein aggregation, stability, absolute and apparent molecular weight, purity and uniformity, melting properties, contamination and water content. Methods to assay these properties include analytical centrifugation, EPR, HPLC-ion exchange, HPLC-

size exclusion, HPLC-reverse phase, light scattering, capillary electrophoresis, circular dichroism, differential scanning calorimetry, fluorescence, HPLC-ion exchange, HPLC-size exclusion, IR, NMR, Raman spectroscopy, refractometry, and UV/Visible spectroscopy. Additional methods are disclosed in Arnau, *et al.*, Prot Expr and Purif (2006) 48, 1-13.

[00171] The XTEN component(s) of the CFXTEN are designed to behave like denatured peptide sequences under physiological conditions, despite the extended length of the polymer. "Denatured" describes the state of a peptide in solution that is characterized by a large conformational freedom of the peptide backbone. Most peptides and proteins adopt a denatured conformation in the presence of high concentrations of denaturants or at elevated temperature. Peptides in denatured conformation have, for example, characteristic circular dichroism (CD) spectra and are characterized by a lack of long-range interactions as determined by NMR. "Denatured conformation" and "unstructured conformation" are used synonymously herein. In some embodiments, the invention provides XTEN sequences that, under physiologic conditions, are largely devoid of secondary structure. In other cases, the XTEN sequences are substantially devoid of secondary structure under physiologic conditions such that the XTEN can adopt random coil conformation. "Largely devoid," as used in this context, means that at least 50% of the XTEN amino acid residues of the XTEN sequence do not contribute to secondary structure as measured or determined by the means described herein. "Substantially devoid," as used in this context, means that at least about 60%, or about 70%, or about 80%, or about 90%, or about 95%, or at least about 99% of the XTEN amino acid residues of the XTEN sequence do not contribute to secondary structure, as measured or determined by the methods described herein.

[00172] A variety of methods have been established in the art to discern the presence or absence of secondary and tertiary structures in a given polypeptide. In particular, secondary structure can be measured spectrophotometrically, e.g., by circular dichroism spectroscopy in the "far-UV" spectral region (190-250 nm). Secondary structure elements, such as alpha-helix and beta-sheet, each give rise to a characteristic shape and magnitude of CD spectra, as does the lack of these structure elements. Secondary structure can also be predicted for a polypeptide sequence via certain computer programs or algorithms, such as the well-known Chou-Fasman algorithm (Chou, P. Y., *et al.* (1974) *Biochemistry*, 13: 222-45) and the Garnier-Osguthorpe-Robson ("GOR") algorithm (Garnier J, Gibrat JF, Robson B. (1996), GOR method for predicting protein secondary structure from amino acid sequence. Methods Enzymol 266:540-553), as described in US Patent Application Publication No. 20030228309A1. For a given sequence, the algorithms can predict whether there exists some or no secondary structure at all, expressed as the total and/or percentage of residues of the sequence that form, for example, alpha-helices or beta-sheets or the percentage of residues of the sequence predicted to result in random coil formation (which lacks secondary structure).

[00173] In one embodiment, the XTEN sequences used in the subject fusion protein compositions have an alpha-helix percentage ranging from 0% to less than about 5% as determined by the Chou-Fasman algorithm. In another embodiment, the XTEN sequences of the fusion protein compositions have a beta-sheet percentage ranging from 0% to less than about 5% as determined by the Chou-Fasman algorithm.

In some embodiments, the XTEN sequences of the fusion protein compositions have an alpha-helix percentage ranging from 0% to less than about 5% and a beta-sheet percentage ranging from 0% to less than about 5% as determined by the Chou-Fasman algorithm. In some embodiments, the XTEN sequences of the fusion protein compositions have an alpha-helix percentage less than about 2% and a beta-sheet percentage less than about 2%. The XTEN sequences of the fusion protein compositions have a high degree of random coil percentage, as determined by the GOR algorithm. In some embodiments, an XTEN sequence have at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, and most preferably at least about 99% random coil, as determined by the GOR algorithm. In some embodiments, the XTEN sequences of the fusion protein compositions have an alpha-helix percentage ranging from 0% to less than about 5% and a beta-sheet percentage ranging from 0% to less than about 5% as determined by the Chou-Fasman algorithm and at least about 90% random coil, as determined by the GOR algorithm. In other embodiments, the XTEN sequences of the fusion protein compositions have an alpha-helix percentage less than about 2% and a beta-sheet percentage less than about 2% at least about 90% random coil, as determined by the GOR algorithm.

1. Non-repetitive Sequences

[00174] It is contemplated that the XTEN sequences of the CFXTEN embodiments are substantially non repetitive. In general, repetitive amino acid sequences have a tendency to aggregate or form higher order structures, as exemplified by natural repetitive sequences such as collagens and leucine zippers. These repetitive amino acids may also tend to form contacts resulting in crystalline or pseudocrystaline structures. In contrast, the low tendency of non-repetitive sequences to aggregate enables the design of long-sequence XTENs with a relatively low frequency of charged amino acids that would otherwise be likely to aggregate if the sequences were repetitive. The non-repetitiveness of a subject XTEN can be observed by assessing one or more of the following features. In one embodiment, a "substantially non-repetitive" XTEN sequence has about 36, or at least 72, or at least 96, or at least 144, or at least 288, or at least 400, or at least 500, or at least 600, or at least 700, or at least 800, or at least 864, or at least 900, or at least 1000, or at least 2000, to about 3000 or more amino acid residues, or has a length ranging from about 36 to about 3000, about 100 to about 500, about 500 to about 1000, about 1000 to about 3000 amino acids and residues, in which no three contiguous amino acids in the sequence are identical amino acid types unless the amino acid is serine, in which case no more than three contiguous amino acids are serine residues. In another embodiment, as described more fully below, a "substantially non-repetitive" XTEN sequence comprises motifs of 9 to 14 amino acid residues wherein the motifs consist of 4 to 6 types of amino acids selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), and wherein the sequence of any two contiguous amino acid residues in any one motif is not repeated more than twice in the sequence motif.

[00175] The degree of repetitiveness of a polypeptide or a gene can be measured by computer programs or algorithms or by other means known in the art. According to the current invention, algorithms to be used in calculating the degree of repetitiveness of a particular polypeptide, such as an XTEN, are

disclosed herein, and examples of sequences analyzed by algorithms are provided (*see* Examples, below). In one aspect, the repetitiveness of a polypeptide of a predetermined length can be calculated (hereinafter "subsequence score") according to the formula given by Equation 1:

$$\text{Subsequence score} \quad \frac{\sum_{i=1}^{m} Count_i}{m} \qquad \text{I}$$

wherein: m = (amino acid length of polypeptide) – (amino acid length of subsequence) + 1; and $Count_i$ = cumulative number of occurrences of each unique subsequence within $sequence_i$

[00176] An algorithm termed "SegScore" was developed to apply the foregoing equation to quantitate repetitiveness of polypeptides, such as an XTEN, providing the subsequence score wherein sequences of a predetermined amino acid length "n" are analyzed for repetitiveness by determining the number of times (a "count") a unique subsequence of length "s" appears in the set length, divided by the absolute number of subsequences within the predetermined length of the sequence. FIG. 27 depicts a logic flowchart of the SegScore algorithm, while FIG. 28 portrays a schematic of how a subsequence score is derived for a fictitious XTEN with 11 amino acids and a subsequence length of 3 amino acid residues. For example, a predetermined polypeptide length of 200 amino acid residues has 192 overlapping 9 amino acid subsequences and 198 3-mer subsequences, but the subsequence score of any given polypeptide will depend on the absolute number of unique subsequences and how frequently each unique subsequence (meaning a different amino acid sequence) appears in the predetermined length of the sequence.

[00177] In the context of the present invention, "subsequence score" means the sum of occurrences of each unique 3-mer frame across 200 consecutive amino acids of the cumulative XTEN polypeptide divided by the absolute number of unique 3-mer subsequences within the 200 amino acid sequence. Examples of such subsequence scores derived from 200 consecutive amino acids of repetitive and non-repetitive polypeptides are presented in Example 45. In one embodiment, the invention provides a CFXTEN comprising one XTEN in which the XTEN has a subsequence score less than 12, more preferably less than 10, more preferably less than 9, more preferably less than 8, more preferably less than 7, more preferably less than 6, and most preferably less than 5. In another embodiment, the invention provides CFXTEN comprising at least two to about six XTEN in which 200 amino acids of the XTEN have a subsequence score of less than 10, more preferably less than 9, more preferably less than 8, more preferably less than 7, more preferably less than 6, and most preferably less than 5. In the embodiments of the CFXTEN fusion protein compositions described herein, an XTEN component of a fusion protein with a subsequence score of 10 or less (i.e., 9, 8, 7, etc.) is also substantially non-repetitive.

[00178] It is believed that the non-repetitive characteristic of XTEN of the present invention together with the particular types of amino acids that predominate in the XTEN, rather than the absolute primary

sequence, confers many of the enhanced physicochemical and biological properties of the CFXTEN fusion proteins. These enhanced properties include a higher degree of expression of the fusion protein in the host cell, greater genetic stability of the gene encoding XTEN, a greater degree of solubility, less tendency to aggregate, and enhanced pharmacokinetics of the resulting CFXTEN compared to fusion proteins comprising polypeptides having repetitive sequences. These enhanced properties permit more efficient manufacturing, lower cost of goods, and facilitate the formulation of XTEN-comprising pharmaceutical preparations containing extremely high protein concentrations, in some cases exceeding 100 mg/ml. Furthermore, the XTEN polypeptide sequences of the embodiments are designed to have a low degree of internal repetitiveness in order to reduce or substantially eliminate immunogenicity when administered to a mammal. Polypeptide sequences composed of short, repeated motifs largely limited to only three amino acids, such as glycine, serine and glutamate, may result in relatively high antibody titers when administered to a mammal despite the absence of predicted T-cell epitopes in these sequences. This may be caused by the repetitive nature of polypeptides, as it has been shown that immunogens with repeated epitopes, including protein aggregates, cross-linked immunogens, and repetitive carbohydrates are highly immunogenic and can, for example, result in the cross-linking of B-cell receptors causing B-cell activation. (Johansson, J., *et al.* (2007) Vaccine, 25 :1676-82 ; Yankai, Z., *et al.* (2006) Biochem Biophys Res Commun, 345 :1365-71 ; Hsu, C. T., *et al.* (2000) Cancer Res, 60:3701-5); Bachmann MF, *et al.* Eur J Immunol. (1995) 25(12):3445 3451).

2. Exemplary Sequence Motifs

[00179] The present invention encompasses XTEN used as fusion partners that comprise multiple units of shorter sequences, or motifs, in which the amino acid sequences of the motifs are non-repetitive. The non-repetitive property is met despite the use of a "building block" approach using a library of sequence motifs that are multimerized to create the XTEN sequences. Thus, while an XTEN sequence may consist of multiple units of as few as four different types of sequence motifs, because the motifs themselves generally consist of non-repetitive amino acid sequences, the overall XTEN sequence is designed to render the sequence substantially non-repetitive.

[00180] In one embodiment, an XTEN has a substantially non-repetitive sequence of greater than about 36 to about 3000, or about 100 to about 2000, or about 144 to about 1000 amino acid residues, or even longer wherein at least about 80%, or at least about 85%, or at least about 90%, or at least about 95%, or at least about 97%, or about 100% of the XTEN sequence consists of non-overlapping sequence motifs, and wherein each of the motifs has about 9 to 36 amino acid residues. In other embodiments, at least about 80%, or at least about 85%, or at least about 90%, or at least about 95%, or at least about 97%, or about 100% of the XTEN sequence consists of non-overlapping sequence motifs wherein each of the motifs has 9 to 14 amino acid residues. In still other embodiments, at least about 80%, or at least about 85%, or at least about 90%, or at least about 95%, or at least about 97%, or about 100% of the XTEN sequence consists of non-overlapping sequence motifs wherein each of the motifs has 12 amino acid residues. In these embodiments, it is preferred that the sequence motifs are composed of substantially (e.g., 90% or more) or exclusively small hydrophilic amino acids, such that the overall sequence has an

unstructured, flexible characteristic. Examples of amino acids that are included in XTEN are, e.g., arginine, lysine, threonine, alanine, asparagine, glutamine, aspartate, glutamate, serine, and glycine. As a result of testing variables such as codon optimization, assembly polynucleotides encoding sequence motifs, expression of protein, charge distribution and solubility of expressed protein, and secondary and tertiary structure, it was discovered that XTEN compositions with the enhanced characteristics disclosed herein mainly or exclusively include glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P) residues wherein the sequences are designed to be substantially non-repetitive. In one embodiment, XTEN sequences have predominately four to six types of amino acids selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) or proline (P) that are arranged in a substantially non-repetitive sequence that is greater than about 36 to about 3000, or about 100 to about 2000, or about 144 to about 1000 amino acid residues in length. In some embodiment, an XTEN sequence is made of 4, 5, or 6 types of amino acids selected from the group consisting of glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) or proline (P). In some embodiments, XTEN have sequences of greater than about 36 to about 1000, or about 100 to about 2000, or about 400 to about 3000 amino acid residues wherein at least about 80% of the sequence consists of non-overlapping sequence motifs wherein each of the motifs has 9 to 36 amino acid residues and wherein at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or 100% of each of the motifs consists of 4 to 6 types of amino acids selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), and wherein the content of any one amino acid type in the full-length XTEN does not exceed 30%. In other embodiments, at least about 90% of the XTEN sequence consists of non-overlapping sequence motifs wherein each of the motifs has 9 to 36 amino acid residues wherein the motifs consist of 4 to 6 types of amino acids selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), and wherein the content of any one amino acid type in the full-length XTEN does not exceed 40%, or about 30%, or 25%, or about 17%. In other embodiments, at least about 90% of the XTEN sequence consists of non-overlapping sequence motifs wherein each of the motifs has 12 amino acid residues consisting of 4 to 6 types of amino acids selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), and wherein the content of any one amino acid type in the full-length XTEN does not exceed 40%, or 30%, or about 25%. In yet other embodiments, at least about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% of the XTEN sequence consists of non-overlapping sequence motifs wherein each of the motifs has 12 amino acid residues consisting of glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P).

[00181] In still other embodiments, XTENs comprise substantially non-repetitive sequences of greater than about 36 to about 3000 amino acid residues wherein at least about 80%, or at least about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99% of the sequence consists of non-overlapping sequence motifs of 9 to 14 amino acid residues wherein the motifs consist of 4 to 6 types of amino acids selected from glycine (G), alanine

(A), serine (S), threonine (T), glutamate (E) and proline (P), and wherein the sequence of any two contiguous amino acid residues in any one motif is not repeated more than twice in the sequence motif. In other embodiments, at least about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99% of an XTEN sequence consists of non-overlapping sequence motifs of 12 amino acid residues wherein the motifs consist of four to six types of amino acids selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), and wherein the sequence of any two contiguous amino acid residues in any one sequence motif is not repeated more than twice in the sequence motif. In other embodiments, at least about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99% of an XTEN sequence consists of non-overlapping sequence motifs of 12 amino acid residues wherein the motifs consist of glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), and wherein the sequence of any two contiguous amino acid residues in any one sequence motif is not repeated more than twice in the sequence motif. In yet other embodiments, XTENs consist of 12 amino acid sequence motifs wherein the amino acids are selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), and wherein the sequence of any two contiguous amino acid residues in any one sequence motif is not repeated more than twice in the sequence motif, and wherein the content of any one amino acid type in the full-length XTEN does not exceed 30%. The foregoing embodiments are examples of substantially non repetitive XTEN sequences. Additional examples are detailed below.

[00182] In some embodiments, the invention provides CFXTEN compositions comprising one, or two, or three, or four, five, six or more non-repetitive XTEN sequence(s) of about 36 to about 1000 amino acid residues, or cumulatively about 100 to about 3000 amino acid residues wherein at least about 80%, or at least about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99% to about 100% of the sequence consists of multiple units of four or more non-overlapping sequence motifs selected from the amino acid sequences of Table 3, wherein the overall sequence remains substantially non-repetitive. In some embodiments, the XTEN comprises non-overlapping sequence motifs in which about 80%, or at least about 85%, or at least about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99% or about 100% of the sequence consists of multiple units of non-overlapping sequences selected from a single motif family selected from Table 3, resulting in a family sequence. As used herein, "family" means that the XTEN has motifs selected only from a single motif category from Table 3; i.e., AD, AE, AF, AG, AM, AQ, BC, or BD XTEN, and that any other amino acids in the XTEN not from a family motif are selected to achieve a needed property, such as to permit incorporation of a restriction site by the encoding nucleotides, incorporation of a cleavage sequence, or to achieve a better linkage to a FVIII coagulation factor component of the CFXTEN. In some embodiments of XTEN families, an XTEN sequence comprises multiple units of non-overlapping sequence motifs of the AD motif family, or of the AE motif family, or of the AF motif family, or of the AG motif family, or of the AM motif family, or of the AQ motif family, or of the BC family, or of the BD family, with the

resulting XTEN exhibiting the range of homology described above. In other embodiments, the XTEN comprises multiple units of motif sequences from two or more of the motif families of Table 3. These sequences can be selected to achieve desired physical/chemical characteristics, including such properties as net charge, hydrophilicity, lack of secondary structure, or lack of repetitiveness that are conferred by the amino acid composition of the motifs, described more fully below. In the embodiments hereinabove described in this paragraph, the motifs incorporated into the XTEN can be selected and assembled using the methods described herein to achieve an XTEN of about 36 to about 3000 amino acid residues.

Table 3: XTEN Sequence Motifs of 12 Amino Acids and Motif Families

Motif Family	MOTIF SEQUENCE	SEQ ID NO:
AD	GESPGGSSGSES	19
AD	GSEGSSGPGESS	20
AD	GSSESGSSEGGP	21
AD	GSGGEPSESGSS	22
AE, AM	GSPAGSPTSTEE	23
AE, AM, AQ	GSEPATSGSETP	24
AE, AM, AQ	GTSESATPESGP	25
AE, AM, AQ	GTSTEPSEGSAP	26
AF, AM	GSTSESPSGTAP	27
AF, AM	GTSTPESGSASP	28
AF, AM	GTSPSGESSTAP	29
AF, AM	GSTSSTAESPGP	30
AG, AM	GTPGSGTASSSP	31
AG, AM	GSSTPSGATGSP	32
AG, AM	GSSPSASTGTGP	33
AG, AM	GASPGTSSTGSP	34
AQ	GEPAGSPTSTSE	35
AQ	GTGEPSSTPASE	36
AQ	GSGPSTESAPTE	37
AQ	GSETPSGPSETA	38
AQ	GPSETSTSEPGA	39
AQ	GSPSEPTEGTSA	40
BC	GSGASEPTSTEP	41
BC	GSEPATSGTEPS	42
BC	GTSEPSTSEPGA	43
BC	GTSTEPSEPGSA	44
BD	GSTAGSETSTEA	45
BD	GSETATSGSETA	46
BD	GTSESATSESGA	47
BD	GTSTEASEGSAS	48

* Denotes individual motif sequences that, when used together in various permutations, results in a "family sequence"

[00183] In some embodiments of XTEN families, an XTEN sequence comprises multiple units of non-overlapping sequence motifs of the AD motif family, the AE motif family, or the AF motif family, or the AG motif family, or the AM motif family, or the AQ motif family, or the BC family, or the BD family, with the resulting XTEN exhibiting the range of homology described above. In other embodiments, the XTEN comprises multiple units of motif sequences from two or more of the motif families of Table 3,

selected to achieve desired physicochemical characteristics, including such properties as net charge, lack of secondary structure, or lack of repetitiveness that may be conferred by the amino acid composition of the motifs, described more fully below. In the embodiments hereinabove described in this paragraph, the motifs or portions of the motifs incorporated into the XTEN can be selected and assembled using the methods described herein to achieve an XTEN of about 36, about 42, about 72, about 144, about 288, about 576, about 864, about 1000, about 2000 to about 3000 amino acid residues, or any intermediate length. Non-limiting examples of XTEN family sequences useful for incorporation into the subject CFXTEN are presented in Table 4. It is intended that a specified sequence mentioned relative to Table 4 has that sequence set forth in Table 4, while a generalized reference to an AE144 sequence, for example, is intended to encompass any AE sequence having 144 amino acid residues; e.g., AE144_1A, AE144_2A, etc., or a generalized reference to an AG144 sequence, for example, is intended to encompass any AG sequence having 144 amino acid residues, e.g., AG144_1, AG144_2, AG144_A, AG144_B, AG144_C, etc.

Table 4: XTEN Polypeptides

XTEN Name	Amino Acid Sequence	SEQ ID NO:
AE42	GAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPASS	49
AE42_1	TEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGS	50
AE42_2	PAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSG	51
AE42_3	SEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSP	52
AG42_1	GAPSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGPSGP	53
AG42_2	GPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASP	54
AG42_3	SPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGA	55
AG42_4	SASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATG	56
AE48	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGS	57
AM48	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGS	58
AE144	GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP	59
AE144_1A	SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	60
AE144_2A	TSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPG	61
AE144_2B	TSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPG	62
AE144_3A	SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG	63
AE144_3B	SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPG	64
AE144_4A	TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	65
AE144_4B	TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG	66

XTEN Name	Amino Acid Sequence	SEQ ID NO:
AE144_5A	TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTS TEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAG SPTSTEEGSPAGSPTSTEEG	67
AE144_6B	TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSE PATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPG	68
AF144	GTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAESPGPGSTSESPSGTAPGS TSSTAESPGPGTSPSGESSTAPGTSTPESGSASPGSTSSTAESPGPGTSPSGESSTAPGTSPS GESSTAPGTSPSGESSTAP	69
AG144_1	SGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSA STGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSAST GTGPGSSPSASTGTGPGASP	70
AG144_2	PGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGP GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPG ASPGTSSTGSPGTPGSGTASSS	71
AG144_A	GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPG SSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGA SPGTSSTGSPGASPGTSSTGSP	72
AG144_B	GTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPG SSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGA SPGTSSTGSPGASPGTSSTGSP	73
AG144_C	GTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPG TPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSS TPSGATGSPGASPGTSSTGSP	74
AG144_F	GSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPG SSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSS TPSGATGSPGASPGTSSTGSP	75
AG144_3	GTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPG ASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGA SPGTSSTGSPGASPGTSSTGSP	76
AG144_4	GTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPG ASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTP GSGTASSSPGSSTPSGATGSP	77
AE288_1	GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGT STEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESA TPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSE GSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP	78
AE288_2	GSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGT STEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPAT SGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPT STEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAP	79
AG288_1	PGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSP GSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPG SSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSS PSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASP GTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGS	80
AG288_2	GSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPG ASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGA SPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSST PSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGASPG TSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSP	81
AF504	GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPG SXPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGA SPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPG SGTASSSPGSSTPSGATGSPGSXPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPS GATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTS STGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSST	82

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	GSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTG SPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGP GASPGTSSTGSP	
AF540	GSTSSTAESPGPGSTSSTAESPGPGSTSESPSGTAPGSTSSTAESPGPGSTSSTAESPGPGT STPESGSASPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSPS GESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPS GTAPGSTSESPSGTAPGTSTPESGSASPGSTSESPSGTAPGTSTPESGSASPGSTSSTAESP GPGSTSSTAESPGPGTSTPESGSASPGTSTPESGSASPGSTSESPSGTAPGTSTPESGSASP GTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSSTAESPGPGT STPESGSASPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSE SPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSPSGESSTAPGSTSSTAESPGPGTSPSGE SSTAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAP	83
AD576	GSSESGSSEGGPGSGGEPSESGSSGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPG SSESGSSEGGPGSSESGSSEGGPGESPGGSSGSESGSEGSSGPGESSGSSESGSSEGGPGSS ESGSSEGGPGSSESGSSEGGPGSGGEPSESGSSGESPGGSSGSESGESPGGSSGSESGSGG EPSESGSSGSSESGSSEGGPGSGGEPSESGSSGSGGEPSESGSSGSEGSSGPGESSGESPGG SSGSESGSGGEPSESGSSGSGGEPSESGSSGSGGEPSESGSSGSSESGSSEGGPGESPGGSS GSESGESPGGSSGSESGESPGGSSGSESGESPGGSSGSESGESPGGSSGSESGSSESGSSEG GPGSGGEPSESGSSGSEGSSGPGESSGSSESGSSEGGPGSGGEPSESGSSGSSESGSSEGG PGSGGEPSESGSSGESPGGSSGSESGESPGGSSGSESGSSESGSSEGGPGSGGEPSESGSS GSSESGSSEGGPGSGGEPSESGSSGSGGEPSESGSSGESPGGSSGSESGSEGSSGPGESSG SSESGSSEGGPGSEGSSGPGESS	84
AE576	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSE GSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPES GPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT SESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTST EPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAP	85
AF576	GSTSSTAESPGPGSTSSTAESPGPGSTSESPSGTAPGSTSSTAESPGPGSTSSTAESPGPGT STPESGSASPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSPS GESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPS GTAPGSTSESPSGTAPGTSTPESGSASPGSTSESPSGTAPGTSTPESGSASPGSTSSTAESP GPGSTSSTAESPGPGTSTPESGSASPGTSTPESGSASPGSTSESPSGTAPGTSTPESGSASP GTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSSTAESPGPGT STPESGSASPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSTSE SPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSPSGESSTAPGSTSSTAESPGPGTSPSGE SSTAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGSTSSTAESPGPGTSTPESGS ASPGTSTPESGSASP	86
AG576	PGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSP GSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPG ASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGA SPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASP GTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSA STGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTA SSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGT GPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTG PGSSPSASTGTGPGASPGTSSTGS	87
AE624	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSPTSTEE GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGT SESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP	88

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	GTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSE SATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAP	
AD836	GSSESGSSEGGPGSSESGSSEGGPGESPGGSSGSESGSGGEPSESGSSGESPGGSSGSESG ESPGGSSGSESGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPGESPGGSSGSESGES PGGSSGSESGESPGGSSGSESGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPGSSES GSSEGGPGSSESGSSEGGPGSSESGSSEGGPGSGGEPSESGSSGESPGGSSGSESGESPGG SSGSESGSGGEPSESGSSGSEGSSGPGESSGSSESGSSEGGPGSGGEPSESGSSGSEGSSGP GESSGSSESGSSEGGPGSGGEPSESGSSGESPGGSSGSESGSGGEPSESGSSGSGGEPSES GSSGSSESGSSEGGPGSGGEPSESGSSGSGGEPSESGSSGSEGSSGPGESSGESPGGSSGS ESGSEGSSGPGESSGSEGSSGPGESSGSGGEPSESGSSGSSESGSSEGGPGSSESGSSEGGP GESPGGSSGSESGSGGEPSESGSSGSEGSSGPGESSGESPGGSSGSESGSEGSSGPGSSES GSSEGGPGSGGEPSESGSSGSEGSSGPGESSGSEGSSGPGESSGSEGSSGPGESSGSGGEP SESGSSGSGGEPSESGSSGESPGGSSGSESGESPGGSSGSESGSGGEPSESGSSGSEGSSGP GESSGESPGGSSGSESGSSESGSSEGGPGSSESGSSEGGPGSSESGSSEGGPGSGGEPSES GSSGSSESGSSEGGPGESPGGSSGSESGSGGEPSESGSSGSESGSSEGGPGESPGGSSGS ESGSGGEPSESGSSGESPGGSSGSESGSGGEPSESGSS	89
AE864	GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSE SATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSE GSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPES GPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGT SESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTST EPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSG SETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGP GTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP	90
AF864	GSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGT STPESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGTSPS GESSTAPGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSPSGESSTAPGSTSSTA ESPGPGTSTPESGSASPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGS ASPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGTSPSGESSTAPGSTSSTAESPG PGTSPSGESSTAPGTSTPESGSASPGSTSSTAESPGPGSTSSTAESPGPGSTSSTAESPGPG STSSTAESPGPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGPXXXGAS ASGAPSTXXXXSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSESPSGTAPGSTSES PSGTAPGSTSESPSGTAPGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGSTSSTAE SPGPGTSPSGESSTAPGTSTPESGSASPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESST APGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSTSESPSGTAPGTSTPESGSASP GSTSSTAESPGPGSTSESPSGTAPGSTSESPSGTAPGTSPSGESSTAPGSTSSTAESPGPGT SPSGESSTAPGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGTSPSGESSTAPGSTSS TAESPGPGSTSSTAESPGPGTSPSGESSTAPGSSPSASTGTGPGSSTPSGATGSPGSSTPSG ATGSP	91
AG864_2	GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPG SSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGA SPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPG SGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPS GATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTS STGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSST GSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTG SPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGP GASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPG ASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSS TPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPS ASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSA	92

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	STGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTS STGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSST GSP	
AM875	GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGS TSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSE SATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEP SEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSE TPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTSTEPSEGSAP GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGA SASGAPSTGGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSE SPSGTAPGTSPSGESSTAPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATS GSETPGTSESATPESGPGSEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESS TAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSAS PGSTSESPSGTAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPG SSPSASTGTGPGASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSS TPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSAPGTSTE PSEGSAP	93
AE912	MAEPAGSPTSTEEGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSPAGSPTSTEE GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGT SESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSE GSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAP GTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGS EPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSE SATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGP GTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGT STEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP	94
AM923	MAEPAGSPTSTEEGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTSTEPSEGSAP GSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGSTSESPSGTAPGS TSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSESATPESGPGSPA GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTST EEGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTSTEPSEGSAPGTSTEPSEGSAP GSEPATSGSETPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGASASGAPSTGGT SESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSTSSTAESPGPGSTSESPSGTAPGTSP SGESSTAPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSEPATSGSETPGTSESA TPESGPGSEPATSGSETPGSTSSTAESPGPGSTSSTAESPGPGTSPSGESSTAPGSEPATSG SETPGSEPATSGSETPGTSTEPSEGSAPGSTSSTAESPGPGTSTPESGSASPGSTSESPSGT APGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSSTPSGATGSPGSSPSASTGTGP GASPGTSSTGSPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSSTPSGATGSPGS SPSASTGTGPGASPGTSSTGSPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP	95
AM1318	GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSTSSTAESPGPGTSTPESGSASPGS TSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGTSTPESGSASPGSEPATSGSETPGTSE SATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEP SEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSE TPGSPAGSPTSTEEGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTSTEPSEGSAP GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGP EPTGPAPSGGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSTSST AESPGPGSTSESPSGTAPGTSPSGESSTAPGSTSESPSGTAPGSTSESPSGTAPGTSPSGES STAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPES	96

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	GPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSPSGESSTAP GTSPSGESSTAPGTSPSGESSTAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGS SPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGAS PGTSSTGSPGASASGAPSTGGTSPSGESSTAPGSTSSTAESPGPGTSPSGESSTAPGTSESA TPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSSPSASTGTGPGSSTPSGATGSPGASPGTSS TGSPGTSTPESGSASPGTSPSGESSTAPGTSPSGESSTAPGTSESATPESGPGSEPATSGSE TPGTSTEPSEGSAPGSTSESPSGTAPGSTSESPSGTAPGTSTPESGSASPGSPAGSPTSTEE GTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGS STPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGSTSESPSGTAPGTSPSGESSTAPGSTS STAESPGPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSPAGSPTSTEEGSPAG SPTSTEEGTSTEPSEGSAP	
BC 864	GTSTEPSEPGSAGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGSEPATSGTEPSGS EPATSGTEPSGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGTST EPSEPGSAGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEPGSAGSEPAT SGTEPSGSEPATSGTEPSGTSEPSTSEPGAGSGASEPTSTEPGTSEPSTSEPGAGSEPATSG TEPSGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEPGSAGSGASEPTSTEPGSEPATSGTE PSGSEPATSGTEPSGSEPATSGTEPSGSEPATSGTEPSGTSTEPSEPGSAGSEPATSGTEPS GSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGS GASEPTSTEPGSEPATSGTEPSGSGASEPTSTEPGSEPATSGTEPSGSGASEPTSTEPGTST EPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGTSTEP SEPGSAGSEPATSGTEPSGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSEPGSAGTSTEPSE PGSAGTSTEPSEPGSAGTSTEPSEPGSAGTSEPSTSEPGAGSGASEPTSTEPGTSTEPSEPG SAGTSTEPSEPGSAGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGSEPATSGTEPS GSEPATSGTEPSGSEPATSGTEPSGSEPATSGTEPSGTSEPSTSEPGAGSEPATSGTEPSGS GASEPTSTEPGTSTEPSEPGSAGSEPATSGTEPSGSGASEPTSTEPGTSTEPSEPGSA	97
BD864	GSETATSGSETAGTSESATSESGAGSTAGSETSTEAGTSESATSESGAGSETATSGSETA GSETATSGSETAGTSTEASEGSASGTSTEASEGSASGTSESATSESGAGSETATSGSETA GTSTEASEGSASGSTAGSETSTEAGTSESATSESGAGTSESATSESGAGSETATSGSETA GTSESATSESGAGTSTEASEGSASGSETATSGSETAGSETATSGSETAGTSTEASEGSAS GSTAGSETSTEAGTSESATSESGAGTSTEASEGSASGSETATSGSETAGSTAGSETSTEA GSTAGSETSTEAGSETATSGSETAGTSESATSESGAGTSESATSESGAGSETATSGSETA GTSESATSESGAGTSESATSESGAGSETATSGSETAGSETATSGSETAGTSTEASEGSAS GSTAGSETSTEAGSETATSGSETAGTSESATSESGAGSTAGSETSTEAGSTAGSETSTEA GSTAGSETSTEAGTSTEASEGSASGSTAGSETSTEAGSTAGSETSTEAGTSTEASEGSAS GSTAGSETSTEAGSETATSGSETAGTSTEASEGSASGTSESATSESGAGSETATSGSETA GTSESATSESGAGTSESATSESGAGSETATSGSETAGTSESATSESGAGSETATSGSETA GTSTEASEGSASGTSTEASEGSASGSTAGSETSTEAGSTAGSETSTEAGSETATSGSETA GTSESATSESGAGTSESATSESGAGSETATSGSETAGSETATSGSETAGSETATSGSETA GTSTEASEGSASGTSESATSESGAGSETATSGSETAGSETATSGSETAGTSESATSESGA GTSESATSESGAGSETATSGSETA	98
AE948	GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGS PAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGSEP ATSGSETPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSEPAT SGSETPGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSEPATSG SETPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPES GPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSEPATSGSETP GTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGT SESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSE SATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSTEP SEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPT STEEGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSE TPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEE GSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGT STEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSE SATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSESA TPESGPGTSESATPESGP	99
AE1044	GSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGT STEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSE SATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTSTEP	100

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	SEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATP ESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGS APGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAP GTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGT SESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTST EPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGS PTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATP ESGPGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGTSESATPES GPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEE GTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGT STEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGSEP ATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP SEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPT STEEGTSESATPESGPGTSESATPESGPGTST	
AE1140	GSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGS EPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSE SATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGS PTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATP ESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEE GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGS PAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPA GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEP SEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPT STEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSE TPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAP GTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGS PAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEP ATSGSETPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEP SEGSAPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPT STEEGTSTEPSEGSAPGSPAGSPTSTEEGSPA	101
AE1236	GSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGT STEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSEP ATSGSETPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPT STEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETP GSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGT SESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESA TPESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESATP ESGPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSE TPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAP GSEPATSGSETPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGT SESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTST EPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESA TPESGPGSEPATSGSETPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSG SETPGTSTEPSEGSAPGTSTEPSEGSAPGSEP	102
AE1332	GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGT STEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGSEPAT SGSETPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSG SETPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGS EPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTST EPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGSPAGS PTSTEEGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSEPATSGSE	103

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	TPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGP GTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGS EPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTST EPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSEPAT SGSETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSESATP ESGPGTSESATPESGPGTSTEPSEGSAPGTST	
AE1428	GSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGT STEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPAT SGSETPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSE GSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETP GSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGS EPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSE SATPESGPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSEPAT SGSETPGTSESATPESGPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSPAGSPT STEEGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPES GPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETP GTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGT SESATPESGPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSE SATPESGPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEP SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSE GSAPGSPAGSPTSTEEGTSESATPESGPGSPA	104
AE1524	GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGS PAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPAT SGSETPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPT STEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESATPES GPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGSPAGSPTSTEE GSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGS EPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGSEP ATSGSETPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSPAGS PTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATP ESGPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSPAGSPTST EEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSESATPESGP GTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGT SESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSPA GSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESA TPESGPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSESATPESGPGSPA	105
AE1620	GSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGT SESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTST EPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSESA TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPES GPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEE GTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGT STEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSEP ATSGSETPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEP SEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSE GSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTST EEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGP GSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGS EPATSGSETPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESA TPESGPGSEPATSGSETPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGSPAGSPTSTEEGTST	106
AE1716	GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGS PAGSPTSTEEGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGTSE	107

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	SATPESGPGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGTSESA TPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSE GSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGS APGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEE GTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGS PAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSPAGSPTSTEEGTST EPSEGSAPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSEPAT SGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPT STEEGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGTSESATPES GPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEE GSPAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGS PAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPA GSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP SEGSAPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSEPATSG SETPGSPAGSPTSTEEGTSESATPESGPGTSE	
AE1812	GTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGT STEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSE SATPESGPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESA TPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGS APGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETP GSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS PAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTST EPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESA TPESGPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPES GPGSEPATSGSETPGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGP GTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGT SESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGTSESA TPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATP ESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEP	108
AE1908	GSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGS PAGSPTSTEEGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGTST EPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATP ESGPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS APGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSESATPESGP GTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGT STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSE SATPESGPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEP SEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSTEPSE GSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGS APGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP GTSESATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGT STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSE SATPESGPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSESATPESGPGTSTEP SEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGTSTEPSE GSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGS APGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETP GTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGS PAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTST EPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEP SEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATP ESGPGSPAGSPTSTEEGTSESATPESGPGSEP	109
AE2004 A	GTSTEPSEGSAPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGS PAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTST EPSEGSAPGSPAGSPTSTEEGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSTEP SEGSAPGTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGSEPATSG SETPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPES GPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEE GTSESATPESGPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGS PAGSPTSTEEGTSESATPESGPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSE SATPESGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEP SEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSPAGSPTSTEEGTSESATP	110

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	ESGPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPES GPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEE GTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGS PAGSPTSTEEGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSEP ATSGSETPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGSPAGSPT STEEGTSTEPSEGSAPGTSESATPESGPGTSE	
AG948	GSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPG TPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGSSPSASTGTGPGSS TPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASP GTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGASPG TSSTGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSG ATGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGASPGTSS TGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTAS SSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTG PGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTGP GSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPG SSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSS PSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTP SGATGSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSG TASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGA TGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTAS SSPGSSTPSGATGSPGSSTPSGATGSP	111
AG1044	GTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPG TPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGSS PSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGTPG SGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPGASPG TSSTGSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTS STGSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGAT GSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTG SPGTPGSGTASSSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSP GASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPG TPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSS TPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGASP GTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGSSTPS GATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSAS TGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGASPGTSS TGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGAT GSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSPSASTGT GPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGSST	112
AG1140	GASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPG SSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSS TPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPS ASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPS GATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGSSTPSG ATGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGA TGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGASPGTSST GSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATG SPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGP GASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPG TPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSS PSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPG SGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSTPS GATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGSSTPSG ATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSPSAST GTGPGTPGSGTASSSPGASPGTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGAT GSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGSST	113
AG1236	GSSPSASTGTGPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPG ASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGTP	114

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	GSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSPS ASTGTGPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSSPSA STGTGPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSG ATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGSSTPSGA TGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSST GSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATG SPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSP GSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPG SSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGA SPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGSSP SASTGTGPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGASPG TSSTGSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGTPGSG TASSSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGTPGSGT ASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTA SSSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASP	
AG1332	GSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPG SSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGSS TPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTP SGATGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSA STGTGPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTS STGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGAT GSPGASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATG SPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSP GSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPG TPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSS TPSGATGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGSSTP SGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPS GATGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTS STGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSST GSPGSSTPSGATGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPG SPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGP GSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPG	115
AG1428	GTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPG TPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSS TPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASP GTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGS GTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSAS TGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTA SSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGT GPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTG PGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSP GASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPG SSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGA SPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGASP GTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGTPGS GTASSSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSG ATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGSSPSAST GTGPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGAT GSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGASP	116
AG1524	GSSTPSGATGSPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPG TPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPGTPGSGTASSSPGTP GSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPGASP GTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGS GTASSSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSG ATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTA SSSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASS SPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGP GSSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPG TPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGSS	117

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	TPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGSSPS ASTGTGPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSG TASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGA TGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGAT GSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTG SPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSP GASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGTPG	
AG1620	GSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPG ASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGSS PSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGSSTP SGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGASPGT SSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSST GSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGT GPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGS PGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSP GASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPG TPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSS TPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSPS ASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGSSPSA STGTGPGSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGSSTPSG ATGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGSSPSAST GTGPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTAS SSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGSST	118
AG1716	GASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSPGSSPSASTGTGPGTPGSGTASSSPG SSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPGTPGSGTASSSPGSS PSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPG SGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSTPS GATGSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPGTPGSGT ASSSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSST GSPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASS SPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSPSASTGTGP GTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPG TPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPGTP GSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGASP GTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGS GTASSSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSG ATGSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPGTPGSGTASSSPGSSPSAST GTGPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSST GSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTG SPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPG	119
AG1812	GSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPG SSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSS PSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTP SGATGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGTPGSG TASSSPGSSPSASTGTGPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGTPGSGT ASSSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGAT GSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATG SPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSP GASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPG ASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGSSTPSGATGSPGSS TPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPG SGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPG TSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGTPGSG TASSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGA TGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGAT GSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTG SPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGASP	120
AG1908	GSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSPSASTGTGPG	121

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	SSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGTPGSGTASSSPGA SPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGSSP SASTGTGPGASPGTSSTGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGS GTASSSPGTPGSGTASSSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSAS TGTGPGSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSST GSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS SPGASPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGP GSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPG SSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSS TPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGTPG SGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPS GATGSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGSSPSAS TGTGPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSPSAST GTGPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGAT GSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSSP	
AG2004A	GSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPG SSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGA SPGTSSTGSPGSSTPSGATGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGSST PSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPG TSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGSSTPSGATGSPGASPGTS STGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTA SSSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASS SPGASPGTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSP GASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSTPSGATGSPG SSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGSS PSASTGTGPGTPGSGTASSSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASP GTSSTGSPGTPGSGTASSSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGTPGS GTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGSSPSASTGTGPGSSTPSG ATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGATGSPGASPGTSS TGSPGTPGSGTASSSPGTPGSGTASSSPGSSPSASTGTGPGASPGTSSTGSPGSSTPSGAT GSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATG SPGTPGSGTASSSPGSSPSASTGTGPGSSPSASTGTGPGASP	122
AE72B	SPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSE PATSGSETPG	123
AE72C	TSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTS TEPSEGSAPG	124
AE108A	TEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSA PGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTS	125
AE108B	GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGS EPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAP	126
AE144A	STEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGS PTSTEEGSPAGSPTSTEEGS	127
AE144B	SEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSP AGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAG SPTSTEEGTSTEPSEGSAPG	128
AE180A	TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTS TEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSET PGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATS	129
AE216A	PESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESG PGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPG TSTEPSEGSAPGSEPATSGSETPGTSESAT	130
AE252A	ESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPES GPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAP GTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGS PAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTST EPSE	131

XTEN Name	Amino Acid Sequence	SEQ ID NO:
AE288A	TPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSG SETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSE TPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGP GTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGT STEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESA	132
AE324A	PESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPG SPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTS ESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTE PSEGSAPGSEPATS	133
AE360A	PESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTS TEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPG SEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTS ESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAG SPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESAT	134
AE396A	PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTS TEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESG PGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSP AGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPA TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS	135
AE432A	EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTE EGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPG SEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSP AGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAG SPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATS GSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATS	136
AE468A	EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSA PGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPG TSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESAT PESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEG SAPGTSTEPSEGSAPGSEPATSGSETPGTSESAT	137
AE504A	EGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTS TEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSET PGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEG SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPS EGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGS ETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESG PGTSTEPS	138
AE540A	TPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSE GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGP GSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGS PAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEP ATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGS PTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPT STEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSE TPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEP	139
AE576A	TPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATP ESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS APGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETP	140

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	GTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS PAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSE SATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEP SEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPT STEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPES GPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAP GSEPATSGSETPGTSESA	
AE612A	GSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTS TEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSA PGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEG TSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAG SPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSA PGTSESATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPG SPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESAT	141
AE648A	PESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEG SAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSET PGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPG TSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSE PATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSES ATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPS EGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPE SGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESG PGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPG TSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTS TEPSEGSAPGSEPATSGSETPGTSESAT	142
AE684A	EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG PGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSA PGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEG SPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSE PATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPA TS	143
AE720A	TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPS EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG PGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSP TSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSA PGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEG SPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSE PATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTE	144
AE756A	TSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPS EGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG PGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPG TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSP	145

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	TSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPE SGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSA PGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEG SPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESGPGSE PATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGSEPA TSGSETPGTSES	
AE792A	EGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPE SGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSA PGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPG TSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSE PATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTE PSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESAT PESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEG SAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESG PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPG TSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTS ESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSES ATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPS EGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPS	146
AE828A	PESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPE SGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSA PGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPG TSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTS ESATPESGPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSES ATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPS EGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGS ETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESG PGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPG TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTS TEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAG SPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESAT PESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEG SAPGSEPATSGSETPGTSESAT	147
AG72A	GPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGS PGTPGSGTASS	148
AG72B	GSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPG TPGSGTASSSP	149
AG72C	SPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSST PSGATGSPGA	150
AG108A	SASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASPG TSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASP	151
AG108B	PGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSP GSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSS	152
AG144A	PGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGP GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPG ASPGTSSTGSPGTPGSGTASSS	153
AG144B	PSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPS ASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSA STGTGPGASPGTSSTGSPGASP	154
AG180A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSAS TGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSS TGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGS	155
AG216A	TGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSS TGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSST GSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATG SPGSSPSASTGTGPGSSPSASTGTGPGSSTPSG	156
AG252A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSAS TGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSS TGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTAS	157

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	SSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGS PGASPG	
AG288A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSAS TGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSS TGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTAS SSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGS PGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGS	158
AG324A	TSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTS STGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTA SSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTG SPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSP GASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPG TPGSGTASSSPGSSTP	159
AG360A	TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGAT GSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGT GPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTG PGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSP GSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPG	160
AG396A	GATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGT ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTA SSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATG SPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSP GSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPG SSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSS TPSGATGSPGSSTPSGATGSPGASPGT	161
AG432A	GATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSG ATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTA SSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASS SPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSP GSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPG SSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSS TPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTP S	162
AG468A	TSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTS STGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGAT GSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTG SPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSP GTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPG SSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASP GTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPG	163
AG504A	TSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTS STGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGAT GSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTG SPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSP GTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPG SSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASP GTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGS GTASSSPGSSTP	164
AG540A	TSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTS STGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTA SSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTG SPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSP GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPG TPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGA SPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASP GTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSA	165

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	STGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGASPG	
AG576A	TSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSAS TGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGA TGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGAT GSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASS SPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSP GSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPG SSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSS TPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTP SGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPS GATGSPGSSTPSGATGSPGASPG	166
AG612A	STGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGAT GSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTG SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSP GASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPG ASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSS TPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASP GTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPG TSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSG TASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSS TGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTS	167
AG648A	GTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSG ATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSS TGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSST GSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTG SPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGP GASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPG ASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGS TPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPS ASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSA STGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTS STGSPGSSPSASTGTGPGTPGSGTASSSPGSSTP	168
AG684A	TSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSG ATGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTA SSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASS SPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSP GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPG SSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGA SPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPG SGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSA STGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSAST GTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTG TGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATG SPGASPG	169
AG720A	TSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSG ATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSPGASPGTSS TGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSST GSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTG SPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTASSSP GSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPG ASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGA SPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPG SGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSA STGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSAST GTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTG TGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATG SPGASPG	170
AG756A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSAS TGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSS TGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTAS	171

XTEN Name	Amino Acid Sequence	SEQ ID NO:
	SSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGS PGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSP GASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPG ASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASP GTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPG TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSG ATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTG TGPGASPGTSSTGSPGASPG	
AG792A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSAS TGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSS TGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTAS SSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGS PGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSP GASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPG ASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASP GTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPG TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSG ATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTG TGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPG	172
AG828A	TSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSAS TGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSS TGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSPGTPGSGTAS SSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGS PGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSP GASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGASPGTSSTGSPG ASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSPGTP GSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGASP GTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGASPG TSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSG ATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSAST GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTG TGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPG	173
AG288_DE	GTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPG ASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGA SPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSST PSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPS ASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP	1699

[00184] In other embodiments, the CFXTEN composition comprises one or more non-repetitive XTEN sequences of lengths ranging from about 36 to about 3000 amino acid residues, wherein at least about 80%, or at least about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99% to about 100% of the sequence consists of non-overlapping 36 amino acid sequence motifs selected from one or more of the polypeptide sequences of Tables 13-17, either as a family sequence, or where motifs are selected from two or more families of motifs.

[00185] In those embodiments wherein the XTEN component of the CFXTEN fusion protein has less than 100% of its amino acids consisting of 4, 5, or 6 types of amino acid selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), or less than 100% of the sequence

consisting of the sequence motifs from Table 3 or the XTEN sequences of Tables 4, and 13-17, the other amino acid residues of the XTEN are selected from any of the other 14 natural L-amino acids, but are preferentially selected from hydrophilic amino acids such that the XTEN sequence contains at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least about 99% hydrophilic amino acids. The XTEN amino acids that are not glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P) are either interspersed throughout the XTEN sequence, are located within or between the sequence motifs, or are concentrated in one or more short stretches of the XTEN sequence, e.g., to create a linker between the XTEN and the FVIII components. In such cases where the XTEN component of the CFXTEN comprises amino acids other than glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P), it is preferred that less than about 2% or less than about 1% of the amino acids be hydrophobic residues such that the resulting sequences generally lack secondary structure, e.g., not having more than 2% alpha helices or 2% beta-sheets, as determined by the methods disclosed herein. Hydrophobic residues that are less favored in construction of XTEN include tryptophan, phenylalanine, tyrosine, leucine, isoleucine, valine, and methionine. Additionally, one can design the XTEN sequences to contain less than 5% or less than 4% or less than 3% or less than 2% or less than 1% or none of the following amino acids: cysteine (to avoid disulfide formation and oxidation), methionine (to avoid oxidation), asparagine and glutamine (to avoid desamidation). Thus, in some embodiments, the XTEN component of the CFXTEN fusion protein comprising other amino acids in addition to glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P) have a sequence with less than 5% of the residues contributing to alpha-helices and beta-sheets as measured by the Chou-Fasman algorithm and have at least 90%, or at least about 95% or more random coil formation as measured by the GOR algorithm.

3. Length of Sequence

[00186] In another aspect, the invention provides XTEN of varying lengths for incorporation into CFXTEN compositions wherein the length of the XTEN sequence(s) are chosen based on the property or function to be achieved in the fusion protein. Depending on the intended property or function, the CFXTEN compositions comprise short or intermediate length XTEN located internal to the FVIII sequence or between FVIII domains and/or longer XTEN sequences that can serve as carriers, located in the fusion proteins as described herein. While not intended to be limiting, the XTEN or fragments of XTEN include short segments of about 6 to about 99 amino acid residues, intermediate lengths of about 100 to about 399 amino acid residues, and longer lengths of about 400 to about 1000 and up to about 3000 amino acid residues. Thus, the XTEN for incorporation into the subject CFXTEN encompass XTEN or fragments of XTEN with lengths of about 6, or about 12, or about 36, or about 40, or about 42, or about 72 or about 96, or about 144, or about 288, or about 400, or about 500, or about 576, or about 600, or about 700, or about 800, or about 864, or about 900, or about 1000, or about 1500, or about 2000, or about 2500, or up to about 3000 amino acid residues in length. Alternatively, the XTEN sequences can be about 6 to about 50, about 50 to about 100, about 100 to 150, about 150 to 250, about 250 to 400, about 400 to about 500, about 500 to about 900, about 900 to 1500, about 1500 to 2000, or about 2000 to

about 3000 amino acid residues in length. The precise length of an XTEN incorporated into the subject CFXTEN can vary without adversely affecting the activity of a CFXTEN composition. In one embodiment, one or more of the XTEN used in the CFXTEN disclosed herein has 36 amino acids, 42 amino acids, 144 amino acids, 288 amino acids, 576 amino acids, or 864 amino acids in length and may be selected from one of the XTEN family sequences; i.e., AD, AE, AF, AG, AM, AQ, BC or BD. In another embodiment, two or more of the XTEN used in the CFXTEN disclosed herein has 36 amino acids, 42 amino acids, 144 amino acids, 288 amino acids, 576 amino acids, or 864 amino acids in length and may be selected from two of the XTEN family sequences; i.e., AD, AE, AF, AG, AM, AQ, BC or BD, with combinations of AE and AG family sequences preferred. In some embodiments, CFXTEN comprising one or more of the XTEN used herein contain XTEN selected from any one of the sequences in Table 4, which may be linked to the FVIII component directly or via spacer sequences disclosed herein.

[00187] In particular CFXTEN configuration designs, where the XTEN serve as a flexible linker, or are inserted in external loops or unordered regions of the FVIII sequence to increase the bulk, flexibility, or hydrophilicity of the region, or are designed to interfere with clearance receptors for FVIII to enhance pharmacokinetic properties, or to interfere with binding of FVIII inhibitors or other anti-FVIII antibodies, or where a short or intermediate length of XTEN is used to facilitate tissue penetration or to vary the strength of interactions of the CFXTEN fusion protein with its target, or where it is desirable to distribute the cumulative length of XTEN in segments of short or intermediate length at multiple locations within the FVIII sequence, the invention contemplates CFXTEN compositions with one, two, three, four, five or more short or intermediate XTEN sequences inserted between or within one or more FVIII domains or within external loops, or at other sites in the FVIII sequence such as, but not limited to, locations at or proximal to the insertion sites identified in Table 5, Table 6, Table 7, Table 8, and Table 9 or as illustrated in FIGS. 8-9. In one embodiment of the foregoing, the CFXTEN fusion protein contains multiple XTEN segments, e.g., at least two, or at least three, or at least four, or at least five or more XTEN segments in which the XTEN segments can be identical or they can be different and wherein the CFXTEN retains at least 10%, 20%, 30%, 40%, 50%, 60%, 70% or more of the procoagulant activity of native FVIII when assayed by one of the assays disclosed herein. In other particular CFXTEN configuration designs, where the XTEN serves as a carrier to increase the bulk of the fusion protein, or to vary the strength of interactions of the CFXTEN fusion protein with its target, or to enhance the pharmacokinetic properties of the fusion protein, the invention contemplates CFXTEN compositions with one or more intermediate or longer length XTEN sequences inserted at the C-terminus, within the B domain (or the residual of the BDD sequence) between or within one or more FVIII domains, within external loops, or at other sites in the FVIII sequence such as, but not limited to, insertion sites identified in Table 5, Table 6, Table 7, Table 8, and Table 9 or as illustrated in FIGS. 8-9. However, it is believed that the incorporation of multiple XTEN of short to intermediate lengths into CFXTEN compositions confers enhanced properties on the fusion proteins compared to CFXTEN fusion proteins with the same number of amino acids in fewer but longer length XTEN, yet still results in compositions with

procoagulant activity and extended half-life; the rationale of which is detailed herein regarding the derived radii of multiple XTEN.

[00188] In the embodiments wherein the CFXTEN fusion proteins comprise multiple XTEN sequences, the cumulative length of the total residues in the XTEN sequences is greater than about 100 to about 3000, or about 200 to about 2000, or about 400 to about 1000 amino acid residues and the XTEN can be identical or they can be different in sequence, net charge, or in length. In one embodiment of CFXTEN comprising multiple XTEN, the individual XTEN sequences each exhibit at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to a motif or an XTEN selected from Tables 3, 4, and 13-17 or a fragment thereof, when optimally aligned with a sequence of comparable length.

[00189] As described more fully below, methods are disclosed in which the CFXTEN are designed by selecting the length of the XTEN and its site of incorporation within the CFXTEN to confer a target half-life, retention of procoagulant activity, reduced binding to FVIII inhibitors or an enhanced physicochemical property (e.g., stability or solubility) of a CFXTEN fusion protein, encoding constructs are created and expressed and the recombinant CFXTEN fusion proteins are isolated and recovered. In general, XTEN cumulative lengths longer that about 400 residues incorporated into the CFXTEN compositions result in longer half life compared to shorter cumulative lengths, e.g., shorter than about 280 residues. In one embodiment, CFXTEN fusion proteins designs are contemplated that comprise at least a single XTEN as a carrier, with a long sequence length of at least about 400, or at least about 600, or at least about 800, or at least about 900, or at least about 1000 or more amino acids. In another embodiment, multiple XTEN are incorporated into the fusion protein to achieve cumulative lengths of at least about 400, or at least about 600, or at least about 800, or at least about 900, or at least about 1000 or more amino acids, wherein the XTEN can be identical or they can be different in sequence or length. As used herein, "cumulative length" is intended to encompass the total length, in amino acid residues, when more than one XTEN is incorporated into the CFXTEN fusion protein. Both of the foregoing embodiments are designed to confer increased bioavailability and/or increased terminal half-life after administration to a subject compared to CFXTEN comprising shorter cumulative XTEN lengths, yet still result in a procoagulant activity and hemostasis effect. When administered subcutaneously or intramuscularly, the C_{max} is reduced but the area under the curve (AUC) is increased in comparison to a comparable dose of a CFXTEN with shorter cumulative length XTEN or FVIII not linked to XTEN, thereby contributing to the ability to maintain effective levels of the CFXTEN composition for a longer period of time and permitting increased periods of 2, 4, 7, 10, 14 or 21 days between dosing, as described more fully below. Thus, the XTEN confers the property of a depot to the administered CFXTEN, in addition to the other physicochemical properties described herein.

[00190] When XTEN are used as a carrier, the invention takes advantage of the discovery that increasing the length of the non-repetitive, unstructured polypeptides enhances the unstructured nature of the XTENs and correspondingly enhances the physical/chemical and pharmacokinetic properties of

fusion proteins comprising the XTEN carrier. As described more fully in the Examples, proportional increases in the length of the XTEN, even if created by a repeated order of single family sequence motifs (e.g., the four AE motifs of Table 3), result in a sequence with a higher percentage (e.g., 90% or more) of random coil formation, as determined by GOR algorithm, or reduced content of alpha-helices or beta-sheets (e.g., less than 2%), as determined by Chou-Fasman algorithm, compared to shorter XTEN lengths. In addition, increasing the length of the unstructured polypeptide fusion partner, as described in the Examples, results in a fusion protein with a disproportionate increase in terminal half-life (e.g., as much as 50, 100, 200 or more hours) compared to fusion proteins with unstructured polypeptide partners with shorter sequence lengths. The enhanced pharmacokinetic properties of the CFXTEN in comparison to FVIII not linked to XTEN are described more fully, below.

[00191] In another aspect, the invention provides methods to create XTEN of short or intermediate lengths from longer "donor" XTEN sequences, wherein the longer donor XTEN sequence is truncated at the N-terminus, or the C-terminus, or a fragment is created from the interior of a donor sequence, thereby resulting in a short or intermediate length XTEN. In non-limiting examples, as schematically depicted in FIG. 16A-C, an AG sequence of 864 amino acid residues can be truncated to yield an AG sequence with 144 residues, an AG sequence with 288 residues, an AG sequence with 576 residues, or other intermediate lengths, while the AE sequence of 864 residues (as depicted in FIG. 16D, E) can be truncated to yield multiple AE sequences of 144 residues, an AE sequence with 288 or 576 residues or other shorter or intermediate lengths. It is specifically contemplated that such an approach can be utilized with any of the XTEN embodiments described herein or with any of the sequences listed in Tables 4 or 13-17 to result in XTEN of a desired length. In preferred embodiments, the CFXTEN comprising multiple XTEN have XTEN exhibiting at least about 80%, or at least about 90%, or at least about 91%, or at least about 92%, or at least about 93%, or at least about 94%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99%, or 100% sequence identity to sequences selected from AE42_1, AE42_2, AE42_3, AG42_1, AG42_2, AG42_3, AG42_4, AE144_1A, AE144_2A, AE144_2B, AE144_3A, AE144_3B, AE144_4A, AE144_4B, AE144_5A, AE144_6B, AG144_1, AG144_2, AG144_A, AG144_B, AG144_C, AG144_F, AG144_3, AG144_4, AE288_1, AE288_2, AG288_1, AG288_2, and AG288_DE.

4. Net charge

[00192] In other embodiments, the unstructured characteristic of an XTEN polypeptide can be enhanced by incorporation of amino acid residues with a net charge and/or reduction of the overall percentage (e.g. less than 5%, or 4%, or 3%, or 2%, or 1%) of hydrophobic amino acids in the XTEN sequence. The overall net charge and net charge density is controlled by modifying the content of charged amino acids in the XTEN sequences, either positive or negative, with the net charge typically represented as the percentage of amino acids in the polypeptide contributing to a charged state beyond those residues that are cancelled by a residue with an opposite charge. In some embodiments, the net charge density of the XTEN of the compositions may be above +0.1 or below -0.1 charges/residue. By "net charge density" of a protein or peptide herein is meant the net charge divided by the total number of amino acids in the

protein or propeptide. In other embodiments, the net charge of an XTEN can be about 0%, about 1%, about 2%, about 3%, about 4%, about 5%, about 6%, about 7%, about 8%, about 9%, about 10% about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 19%, or about 20% or more. Based on the net charge, some XTENs have an isoelectric point (pI) of 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, or even 6.5. In preferred embodiments, the XTEN will have an isoelectric point between 1.5 and 4.5 and carry a net negative charge under physiologic conditions.

[00193] Since most tissues and surfaces in a human or animal have a net negative charge, in some embodiments the XTEN sequences are designed to have a net negative charge to minimize non-specific interactions between the XTEN containing compositions and various surfaces such as blood vessels, healthy tissues, or various receptors. Not to be bound by a particular theory, an XTEN can adopt open conformations due to electrostatic repulsion between individual amino acids of the XTEN polypeptide that individually carry a net negative charge and that are distributed across the sequence of the XTEN polypeptide. In some embodiments, the XTEN sequence is designed with at least 90% or 95% of the charged residues separated by other residues such as serine, alanine, threonine, proline or glycine, which leads to a more uniform distribution of charge, better expression or purification behavior. Such a distribution of net negative charge in the extended sequence lengths of XTEN can lead to an unstructured conformation that, in turn, can result in an effective increase in hydrodynamic radius. In preferred embodiments, the negative charge of the subject XTEN is conferred by incorporation of glutamic acid residues. Generally, the glutamic residues are spaced uniformly across the XTEN sequence. In some cases, the XTEN can contain about 10-80, or about 15-60, or about 20-50 glutamic residues per 20kDa of XTEN that can result in an XTEN with charged residues that would have very similar pKa, which can increase the charge homogeneity of the product and sharpen its isoelectric point, enhance the physicochemical properties of the resulting CFXTEN fusion protein for, and hence, simplifying purification procedures. For example, where an XTEN with a negative charge is desired, the XTEN can be selected solely from an AE family sequence, which has approximately a 17% net charge due to incorporated glutamic acid, or can include varying proportions of glutamic acid-containing motifs of Table 3 to provide the desired degree of net charge. Non-limiting examples of AE XTEN include, but are not limited to the 36, 42, 144, 288, 576, 624, 864, and 912 AE family sequences of Tables 4 and 14 or fragments thereof. In one embodiment, an XTEN sequence of Tables 4, or 13-17 can be modified to include additional glutamic acid residues to achieve the desired net negative charge. Accordingly, in one embodiment the invention provides XTEN in which the XTEN sequences contain about 1%, 2%, 4%, 8%, 10%, 15%, 17%, 20%, 25%, or even about 30% glutamic acid. In one embodiment, the invention contemplates incorporation of up to 5% aspartic acid residues into XTEN in addition to glutamic acid in order to achieve a net negative charge.

[00194] In other embodiments, where no net charge is desired, the XTEN can be selected from, for example, AG XTEN components, such as the AG motifs of Table 3, or those AM motifs of Table 3 that have no net charge. Non-limiting examples of AG XTEN include, but are not limited to 36, 42, 144, 288, 576, and 864 AG family sequences of Tables 4 and 16, or fragments thereof. In another

embodiment, the XTEN can comprise varying proportions of AE and AG motifs (in order to have a net charge that is deemed optimal for a given use or to maintain a given physicochemical property.

[00195] Not to be bound by a particular theory, the XTEN of the CFXTEN compositions with the higher net charge are expected to have less non-specific interactions with various negatively-charged surfaces such as blood vessels, tissues, or various receptors, which would further contribute to reduced active clearance. Conversely, it is believed that the XTEN of the CFXTEN compositions with a low (or no) net charge would have a higher degree of interaction with surfaces that can potentiate the activity of the associated coagulation factor, given the known contribution of cell (e.g., platelets) and vascular surfaces to the coagulation process and the intensity of activation of coagulation factors (Zhou, R., et al., Biomaterials (2005) 26(16):2965-2973; London, F., et al. Biochemistry (2000) 39(32):9850–9858).

[00196] The XTEN of the compositions of the present invention generally have no or a low content of positively charged amino acids. In some embodiments, the XTEN may have less than about 10% amino acid residues with a positive charge, or less than about 7%, or less than about 5%, or less than about 2%, or less than about 1% amino acid residues with a positive charge. However, the invention contemplates constructs where a limited number of amino acids with a positive charge, such as lysine, are incorporated into XTEN to permit conjugation between the epsilon amine of the lysine and a reactive group on a peptide, a linker bridge, or a reactive group on a drug or small molecule to be conjugated to the XTEN backbone. In one embodiment of the foregoing, the XTEN of the subject CFXTEN has between about 1 to about 100 lysine residues, or about 1 to about 70 lysine residues, or about 1 to about 50 lysine residues, or about 1 to about 30 lysine residues, or about 1 to about 20 lysine residues, or about 1 to about 10 lysine residues, or about 1 to about 5 lysine residues, or alternatively only a single lysine residue. Using the foregoing lysine-containing XTEN, fusion proteins can be constructed that comprise XTEN, a FVIII coagulation factor, plus a chemotherapeutic agent or other coagulation factor or cofactor useful in the treatment of coagulopathy conditions, wherein the maximum number of molecules of the agent incorporated into the XTEN component is determined by the numbers of lysines or other amino acids with reactive side chains (e.g., cysteine) incorporated into the XTEN.

[00197] As hydrophobic amino acids impart structure to a polypeptide, the invention provides that the content of hydrophobic amino acids in the XTEN will typically be less than 5%, or less than 2%, or less than 1% hydrophobic amino acid content. In one embodiment, the amino acid content of methionine and tryptophan in the XTEN component of a CFXTEN fusion protein is typically less than 5%, or less than 2%, and most preferably less than 1%. In another embodiment, the XTEN of the subject CFXTEN compositions will have a sequence that has less than 10% amino acid residues with a positive charge, or less than about 7%, or less that about 5%, or less than about 2% amino acid residues with a positive charge, the sum of methionine and tryptophan residues will be less than 2%, and the sum of asparagine and glutamine residues will be less than 5% of the total XTEN sequence.

5. Low immunogenicity

[00198] In another aspect, the XTEN sequences provided herein have a low degree of immunogenicity or are substantially non-immunogenic. Several factors can contribute to the low immunogenicity of

XTEN, e.g., the non-repetitive sequence, the unstructured conformation, the high degree of solubility, the low degree or lack of self-aggregation, the low degree or lack of proteolytic sites within the sequence, and the low degree or lack of epitopes in the XTEN sequence.

[00199] Conformational epitopes are formed by regions of the protein surface that are composed of multiple discontinuous amino acid sequences of the protein antigen. The precise folding of the protein brings these sequences into a well-defined, stable spatial configurations, or epitopes, that can be recognized as "foreign" by the host humoral immune system, resulting in the production of antibodies to the protein or the activation of a cell-mediated immune response. In the latter case, the immune response to a protein in an individual is heavily influenced by T-cell epitope recognition that is a function of the peptide binding specificity of that individual's HLA-DR allotype. Engagement of a MHC Class II peptide complex by a cognate T-cell receptor on the surface of the T-cell, together with the cross-binding of certain other co-receptors such as the CD4 molecule, can induce an activated state within the T-cell. Activation leads to the release of cytokines further activating other lymphocytes such as B cells to produce antibodies or activating T killer cells as a full cellular immune response.

[00200] The ability of a peptide to bind a given MHC Class II molecule for presentation on the surface of an APC (antigen presenting cell) is dependent on a number of factors; most notably its primary sequence. In one embodiment, a lower degree of immunogenicity is achieved by designing XTEN sequences that resist antigen processing in antigen presenting cells, and/or choosing sequences that do not bind MHC receptors well. The invention provides CFXTEN fusion proteins with substantially non-repetitive XTEN polypeptides designed to reduce binding with MHC II receptors, as well as avoiding formation of epitopes for T-cell receptor or antibody binding, resulting in a low degree of immunogenicity. Avoidance of immunogenicity can attribute to, at least in part, a result of the conformational flexibility of XTEN sequences; i.e., the lack of secondary structure due to the selection and order of amino acid residues. For example, of particular interest are sequences having a low tendency to adapt compactly folded conformations in aqueous solution or under physiologic conditions that could result in conformational epitopes. The administration of fusion proteins comprising XTEN, using conventional therapeutic practices and dosing, would generally not result in the formation of neutralizing antibodies to the XTEN sequence, and also reduce the immunogenicity of the FVIII fusion partner in the CFXTEN compositions.

[00201] In one embodiment, the XTEN sequences utilized in the subject fusion proteins can be substantially free of epitopes recognized by human T cells. The elimination of such epitopes for the purpose of generating less immunogenic proteins has been disclosed previously; see for example WO 98/52976, WO 02/079232, and WO00/3317. Assays for human T cell epitopes have been described (Stickler, M., *et al.* (2003) *J Immunol Methods* 281:95-108). Of particular interest are peptide sequences that can be oligomerized without generating T cell epitopes or non-human sequences. This is achieved by testing direct repeats of these sequences for the presence of T-cell epitopes and for the occurrence of 6 to 15-mer and, in particular, 9-mer sequences that are not human, and then altering the design of the XTEN sequence to eliminate or disrupt the epitope sequence.

In some embodiments, the XTEN sequences are substantially non-immunogenic by the restriction of the numbers of epitopes of the XTEN predicted to bind MHC receptors. With a reduction in the numbers of epitopes capable of binding to MHC receptors, there is a concomitant reduction in the potential for T cell activation as well as T cell helper function, reduced B cell activation or upregulation and reduced antibody production. The low degree of predicted T-cell epitopes can be determined by epitope prediction algorithms such as, e.g., TEPITOPE (Sturniolo, T., *et al.* (1999) Nat Biotechnol, 17: 555-61), as shown in Example 46. The TEPITOPE score of a given peptide frame within a protein is the log of the K_d (dissociation constant, affinity, off-rate) of the binding of that peptide frame to multiple of the most common human MHC alleles, as disclosed in Sturniolo, T. *et al.* (1999) *Nature Biotechnology* 17:555). The score ranges over at least 20 logs, from about 10 to about -10 (corresponding to binding constraints of $10e^{10}$ K_d to $10e^{-10}$ K_d), and can be reduced by avoiding hydrophobic amino acids that serve as anchor residues during peptide display on MHC, such as M, I, L, V, F. In some embodiments, an XTEN component incorporated into a CFXTEN does not have a predicted T-cell epitope at a TEPITOPE threshold score of about -5, or -6, or -7, or -8, or -9, or at a TEPITOPE score of -10. As used herein, a score of "-9" is a more stringent TEPITOPE threshold than a score of -5.

[00202] In another embodiment, the inventive XTEN sequences, including those incorporated into the subject CFXTEN fusion proteins, are rendered substantially non-immunogenic by the restriction of known proteolytic sites from the sequence of the XTEN, reducing the processing of XTEN into small peptides that can bind to MHC II receptors. In another embodiment, the XTEN sequence is rendered substantially non-immunogenic by the use a sequence that is substantially devoid of secondary structure, conferring resistance to many proteases due to the high entropy of the structure. Accordingly, the reduced TEPITOPE score and elimination of known proteolytic sites from the XTEN render the XTEN compositions, including the XTEN of the CFXTEN fusion protein compositions, substantially unable to be bound by mammalian receptors, including those of the immune system or active clearance receptors that target FVIII. In one embodiment, an XTEN of a CFXTEN fusion protein can have >100 nM K_d binding to a mammalian receptor, or greater than 500 nM K_d, or greater than 1 μM K_d towards a mammalian cell surface receptor or circulating polypeptide receptor.

[00203] Additionally, the non-repetitive sequence and corresponding lack of epitopes of XTEN limit the ability of B cells to bind to or be activated by XTEN. A repetitive sequence is recognized and can form multivalent contacts with even a few B cells and, as a consequence of the cross-linking of multiple T-cell independent receptors, can stimulate B cell proliferation and antibody production. In contrast, while an XTEN can make contacts with many different B cells over its extended sequence, each individual B cell may only make one or a small number of contacts with an individual XTEN due to the lack of repetitiveness of the sequence. Not being to be bound by any theory, XTENs typically have a much lower tendency to stimulate proliferation of B cells and thus an immune response. In one embodiment, the CFXTEN have reduced immunogenicity as compared to the corresponding FVIII that is not fused to an XTEN. In one embodiment, the administration of up to three parenteral doses of a CFXTEN to a mammal result in detectable anti-CFXTEN IgG at a serum dilution of 1:100 but not at a

dilution of 1:1000. In another embodiment, the administration of up to three parenteral doses of a CFXTEN to a mammal result in detectable anti-FVIII IgG at a serum dilution of 1:100 but not at a dilution of 1:1000. In another embodiment, the administration of up to three parenteral doses of a CFXTEN to a mammal result in detectable anti-XTEN IgG at a serum dilution of 1:100 but not at a dilution of 1:1000. In the foregoing embodiments, the mammal can be a mouse, a rat, a rabbit, or a cynomolgus monkey.

[00204] An additional feature of XTENs with non-repetitive sequences relative to sequences with a high degree of repetitiveness is non-repetitive XTENs form weaker contacts with antibodies. Antibodies are multivalent molecules. For instance, IgGs have two identical binding sites and IgMs contain 10 identical binding sites. Thus antibodies against repetitive sequences can form multivalent contacts with such repetitive sequences with high avidity, which can affect the potency and/or elimination of such repetitive sequences. In contrast, antibodies against non-repetitive XTENs may yield monovalent interactions, resulting in less likelihood of immune clearance such that the CFXTEN compositions can remain in circulation for an increased period of time. In addition, it is believed, as schematically portrayed in FIG. 6, the flexible unstructured nature of XTEN provides steric shielding of FVIII regions proximal to the XTEN site of insertion and providing steric hindrance to binding by FVIII inhibitors.

[00205] In another aspect, a subject XTEN useful as a fusion partner has a high hydrodynamic radius; a property that in some embodiments confers a corresponding increased apparent molecular weight to the CFXTEN fusion protein incorporating the XTEN, while in other embodiments enhances steric hindrance to FVIII inhibitors and to anti-FVIII antibodies, reducing their ability to bind to CFXTEN. As detailed in Example 26, the linking of XTEN to therapeutic protein sequences results in CFXTEN compositions that can have increased hydrodynamic radii, increased apparent molecular weight, and increased apparent molecular weight factor compared to a therapeutic protein not linked to an XTEN. For example, in therapeutic applications in which prolonged half-life is desired, compositions in which an XTEN with a high hydrodynamic radius is incorporated into a fusion protein comprising a therapeutic protein can effectively enlarge the hydrodynamic radius of the composition beyond the glomerular pore size of approximately 3-5 nm (corresponding to an apparent molecular weight of about 70 kDa) (Caliceti. 2003. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. Adv Drug Deliv Rev 55:1261-1277), resulting in reduced renal clearance of circulating proteins with a corresponding increase in terminal half-life and other enhanced pharmacokinetic properties. The hydrodynamic radius of a protein is conferred by its molecular weight as well as by its structure, including shape or compactness. Not to be bound by a particular theory, the XTEN can adopt open conformations due to electrostatic repulsion between individual charges of the peptide or the inherent flexibility imparted by the particular amino acids in the sequence that lack potential to confer secondary structure. The open, extended and unstructured conformation of the XTEN polypeptide can have a greater proportional hydrodynamic radius compared to polypeptides of a comparable sequence length and/or molecular weight that have secondary and/or tertiary structure, such as typical globular proteins. Methods for determining the hydrodynamic radius are well known in the art, such as by the use of size

exclusion chromatography (SEC), as described in U.S. Patent Nos. 6,406,632 and 7,294,513. Example 26 demonstrates that increases in XTEN length result in proportional increase in the hydrodynamic radius, apparent molecular weight, and/or apparent molecular weight factor, and thus permit the tailoring of CFXTEN to desired cut-off values of apparent molecular weights or hydrodynamic radii. Accordingly, in certain embodiments, the CFXTEN fusion protein can be configured with an XTEN such that the fusion protein can have a hydrodynamic radius of at least about 5 nm, or at least about 8 nm, or at least about 10 nm, or about 12 nm, or about 15 nm, or about 20 nm, or about 30 nm or more. In the foregoing embodiments, the large hydrodynamic radius conferred by the XTEN in a CFXTEN fusion protein can lead to reduced clearance of the resulting fusion protein, an increase in terminal half-life, and an increase in mean residence time.

[00206] Generally, the actual molecular weight of the mature form of FVIII component is about 265 kDa, while in the case of a FVIII BDD, it is about 165 kDa. The actual molecular weight of a CFXTEN fusion protein for comprising a FVIII BDD plus one or more XTEN ranges from about 200 to about 270 kDa, depending on the length of the XTEN components. As described in the Examples, when the molecular weights of the CFXTEN fusion proteins are derived from size exclusion chromatography analyses, the open conformation of the XTEN due to the low degree of secondary structure results in an increase in the apparent molecular weight of the fusion proteins into which they are incorporated. In some embodiments, the CFXTEN comprising a FVIII and at least one or more XTEN exhibits an apparent molecular weight of at least about 400 kD, or at least about 500 kD, or at least about 700 kD, or at least about 1000 kD, or at least about 1400 kD, or at least about 1600 kD, or at least about 1800kD, or at least about 2000 kD. Accordingly, the CFXTEN fusion proteins comprising one or more XTEN exhibit an apparent molecular weight that is about 1.3-fold greater, or about 2-fold greater, or about 3-fold greater or about 4-fold greater, or about 8-fold greater, or about 10-fold greater, or about 12-fold greater, or about 15-fold greater than the actual molecular weight of the fusion protein. In one embodiment, the isolated CFXTEN fusion protein of any of the embodiments disclosed herein exhibit an apparent molecular weight factor under physiologic conditions that is greater than about 1.3, or about 2, or about 3, or about 4, or about 5, or about 6, or about 7, or about 8, or about 10, or greater than about 15. In another embodiment, the CFXTEN fusion protein has, under physiologic conditions, an apparent molecular weight factor that is about 3 to about 20, or is about 5 to about 15, or is about 8 to about 12, or is about 9 to about 10 relative to the actual molecular weight of the fusion protein. It is believed that the increased apparent molecular weight of the subject CFXTEN compositions enhances the pharmacokinetic properties of the fusion proteins by a combination of factors, which include reduced active clearance, reduced binding by FVIII inhibitors, and reduced loss in capillary and venous bleeding.

IV). CFXTEN COMPOSITIONS

[00207] The present invention provides compositions comprising fusion proteins having factor VIII linked to one or more XTEN sequences, wherein the fusion protein acts to replace or augment the amount of existing FVIII in the intrinsic or contact activated coagulation pathway when administered

into a subject. The invention addresses a long-felt need in increasing the terminal half-life of exogenously administered factor VIII to a subject in need thereof. One way to increase the circulation half-life of a therapeutic protein is to ensure that renal clearance or metabolism of the protein is reduced. Another way to increase the terminal half-life is to reduce the active clearance of the therapeutic protein, whether mediated by receptors, active metabolism of the protein, or other endogenous mechanisms. Both may be achieved by conjugating the protein to a polymer, which, on one hand, is capable of conferring an increased molecular size (or hydrodynamic radius) to the protein and, hence, reduced renal clearance, and, on the other hand, interferes with binding of the protein to clearance receptors or other proteins that contribute to metabolism or clearance. Thus, certain objects of the present invention include, but are not limited to, providing improved FVIII molecules with a longer circulation or terminal half-life, decreasing the number or frequency of necessary administrations of FVIII compositions, retaining at least a portion of the activity compared to native coagulation factor VIII, and/or enhancing the ability to treat coagulation deficiencies and uncontrolled bleedings more efficiently, more effectively, more economically, and/or with greater safety compared to presently available factor VIII preparations.

[00208] Accordingly, the present invention provides recombinant factor VIII fusion protein compositions comprising an FVIII covalently linked to one or more extended recombinant polypeptides ("XTEN"), resulting in a CFXTEN fusion protein composition. The term "CFXTEN", as used herein, is meant to encompass fusion polypeptides that comprise at least one payload region comprising a FVIII or a portion of a FVIII that is capable of procoagulant activity associated with a FVIII coagulation factor and at least one other region comprising one or more XTEN polypeptides that may be interspersed within the payload region and/or attached to the terminus. In one embodiment, the FVIII is native FVIII. In another embodiment, the FVIII is a sequence variant, fragment, homolog, or mimetic of a natural sequence that retains at least a portion of the procoagulant activity of native FVIII, as disclosed herein. Non-limiting examples of FVIII suitable for inclusion in the compositions include the sequences of Table 1 or sequences having at least 80%, or at least 90%, or at least 91%, or at least 92%, or at least 93%, or at least 94%, or at least 95%, or at least 96%, or at least 97%, or at least 98%, or at least 99% sequence identity to a sequence of Table 1. In a preferred embodiment, the FVIII is a B-domain deleted (BDD) FVIII sequence variant, such as those BDD sequences from Table 1 or other such sequences known in the art. In another preferred embodiment, the CFXTEN comprises a B-domain deleted (BDD) FVIII sequence variant expressed with the native 19 amino acid signal sequence, which is cleaved during the maturation of the protein.

[00209] The compositions of the invention include fusion proteins that are useful, when administered to a subject in need thereof, for mediating or preventing or ameliorating a condition associated with factor VIII deficiencies or defects in endogenously produced FVIII, or bleeding disorders associated with trauma, surgery, factor VIII deficiencies or defects. Of particular interest are CFXTEN fusion protein compositions for which an increase in a pharmacokinetic parameter, increased solubility, increased stability, or some other enhanced pharmaceutical property compared to native FVIII is sought, or for which increasing the terminal half-life would improve efficacy, safety, or result in reduced dosing

frequency and/or improve patient management. The CFXTEN fusion proteins of the embodiments disclosed herein exhibit one or more or any combination of the improved properties and/or the embodiments as detailed herein. In some embodiments, the CFXTEN fusion composition remains at a level above a threshold value of at least 0.01-0.05, or 0.05 to 0.1, or 0.1 to 0.4 IU/ml when administered to a subject, for a longer period of time when compared to a FVIII not linked to XTEN and administered at a comparable dose to a subject in need thereof (e.g., a subject such as a human or mouse or monkey with hemophilia A).

[00210] The FVIII of the subject compositions, particularly those disclosed in Table 1, together with their corresponding nucleic acid and amino acid sequences, are available in public databases such as Chemical Abstracts Services Databases (e.g., the CAS Registry), GenBank, The Universal Protein Resource (UniProt), subscription provided databases such as GenSeq (e.g., Derwent), as well as in the patent and primary literature. Polynucleotide sequences applicable for expressing the subject CFXTEN sequences may be a wild type polynucleotide sequence encoding a given FVIII (e.g., either full length or mature), or in some instances the sequence may be a variant of the wild type polynucleotide sequence (e.g., a polynucleotide which encodes the wild type biologically active protein, wherein the DNA sequence of the polynucleotide has been optimized, for example, for expression in a particular species, or a polynucleotide encoding a variant of the wild type protein, such as a site directed mutant or an allelic variant. It is well within the ability of the skilled artisan to use a wild type or consensus cDNA sequence or a codon-optimized variant of a FVIII to create CFXTEN constructs contemplated by the invention using methods known in the art and/or in conjunction with the guidance and methods provided herein, and described more fully in the Examples.

[00211] In one embodiment, a CFXTEN fusion protein comprises a single FVIII molecule exhibiting at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least about 99%, or 100% sequence identity to a sequence of Table 1 linked to a single XTEN (e.g., an XTEN as described above) including, but not limited to sequences of the AE or AG family with 42, 144, 288, 576, or 864 amino acids, as set forth in Table 4. In another embodiment, the CFXTEN comprises a single FVIII linked to two XTEN, wherein the XTEN may be identical or they may be different. In another embodiment, the CFXTEN fusion protein comprises a single FVIII molecule linked to one, two, three, four, five, six or more XTEN sequences, in which the FVIII is a sequence that has at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least about 99%, or 100% sequence identity compared to a protein sequence selected from Table 1, when optimally aligned, and the one or more XTEN are each having at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least about 99%, or 100% sequence identity compared to one or more sequences selected from any one of Tables 3, 4, and 13-17, when optimally aligned. In the foregoing embodiment, where the CFXTEN has two or more XTEN, the XTEN may be identical or they may be different sequences. In yet another embodiment, the CFXTEN fusion protein comprises a single FVIII

exhibiting at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least about 99%, or 100% sequence identity compared to sequences of comparable length selected from Table 1, when optimally aligned, with the portions interspersed with and linked by three, four, five, six or more XTEN sequences that may be identical or may be different and wherein each has at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least about 99%, or 100% sequence identity compared to sequences selected from any one of Tables 3, 4, and 13-17, or fragments thereof, when optimally aligned. In yet another embodiment, the invention provides a CFXTEN fusion protein comprising a sequence with at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least about 99%, or 100% sequence identity to a sequence from Table 21, when optimally aligned.

1. CFXTEN Fusion Protein Configurations

[00212] The invention provides CFXTEN fusion protein compositions with the CF and XTEN components linked in specific N- to C-terminus configurations.

[00213] In one embodiment of the CFXTEN composition, the invention provides a fusion protein of formula I:

$(XTEN)_x$ CF $(XTEN)_y$ I

wherein independently for each occurrence, CF is a factor VIII as defined herein, including sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity with sequenced from Table 1; x is either 0 or 1 and y is either 0 or 1 wherein $x+y \geq 1$; and XTEN is an extended recombinant polypeptide as described herein, including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4. Accordingly, the CFXTEN fusion composition can have XTEN-CF, XTEN-CF-XTEN, or CF-XTEN configurations.

[00214] In another embodiment of the CFXTEN composition, the invention provides a fusion protein of formula II:

$(XTEN)_x$-$(S)_x$-(CF)-$(XTEN)_y$ II

wherein independently for each occurrence, CF is a factor VIII as defined herein, including sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 1; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; x is either 0 or 1 and y is either 0 or 1 wherein $x+y \geq 1$; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4.

[00215] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein, wherein the fusion protein is of formula III:

$(XTEN)_x$-$(S)_x$-(CF)-$(S)_y$-$(XTEN)_y$ III

wherein independently for each occurrence, CF is a factor VIII as defined herein, including sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequence set for in Table 1; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; x is either 0 or 1 and y is either 0 or 1 wherein $x+y \geq 1$; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4.

[00216] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula IV:

$(A1)$-$(XTEN)_u$-$(A2)$-$(XTEN)_v$-(B)-$(XTEN)_w$-$(A3)$-$(XTEN)_x$-$(C1)$-$(XTEN)_y$-$(C2)$-$(XTEN)_z$ IV

wherein independently for each occurrence, A1 is an A1 domain of FVIII; A2 is an A2 domain of FVIII; A3 is an A3 domain of FVIII; B is a B domain of FVIII which can be a fragment or a splice variant of the B domain; C1 is a C1 domain of FVIII; C2 is a C2 domain of FVIII; v is either 0 or 1; w is either 0 or 1; x is either 0 or 1; y is either 0 or 1; y is either 0 or 1 with the proviso that $u + v + x + y+z \geq 1$; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4.

[00217] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula V:

$(XTEN)_t$-$(S)_a$ -$(A1)$-$(S)_b$-$(XTEN)_u$-$(S)_b$-$(A2)$-$(S)_c$-$(XTEN)_v$-$(S)_c$-(B)-$(S)_d$-$(XTEN)_w$-$(S)_d$-$(A3)$-$(S)_e$-$(XTEN)_x$-$(S)_e$-$(C1)$-$(S)_f$-$(XTEN)_y$-$(S)_f$-$(C2)$-$(S)_g$-$(XTEN)_z$ V

wherein independently for each occurrence, A1 is an A1 domain of FVIII; A2 is an A2 domain of FVIII; A3 is an A3 domain of FVIII; B is a B domain of FVIII which can be a fragment or a splice variant of the B domain; C1 is a C1 domain of FVIII; C2 is a C2 domain of FVIII; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; a is either 0 or 1; b is either 0 or 1; c is either 0 or 1; d is either 0 or 1; e is either 0 or 1; f is either 0 or 1; g is either 0 or 1; t is either 0 or 1; u is either 0 or 1; v is either 0 or 1; w is 0 or 1, x is either 0 or 1; y is either 0 or 1; z is either 0 or 1 with the proviso that $t + u + v + w+ x + y + z \geq 1$; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity

to sequences set forth in Table 4. In another embodiment of formula V, the spacer sequence is glycine or a sequence selected from Tables 11 and 12.

[00218] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula VI:

$(XTEN)_u$-$(S)_a$-(A1)-$(S)_b$-$(XTEN)_v$-$(S)_b$-(A2)-$(S)_c$-$(XTEN)_w$-$(S)_c$-(A3)-$(S)_d$-$(XTEN)_x$-$(S)_d$-(C1)-$(S)_e$-$(XTEN)_y$-$(S)_e$-(C2)-$(S)_f$-$(XTEN)_z$ VI

wherein independently for each occurrence, A1 is an A1 domain of FVIII; A2 is an A2 domain of FVIII; A3 is an A3 domain of FVIII; C1 is a C1 domain of FVIII; C2 is a C2 domain of FVIII; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; a is either 0 or 1; b is either 0 or 1; c is either 0 or 1; d is either 0 or 1; e is either 0 or 1; f is either 0 or 1; u is either 0 or 1; v is either 0 or 1; w is 0 or 1, x is either 0 or 1; y is either 0 or 1; z is either 0 or 1 with the proviso that $u + v + w + x + y + z \geq 1$; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4. In another embodiment of formula V, the spacer sequence is glycine or a sequence selected from Tables 11 and 12.

[00219] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula VII:

(SP)-$(XTEN)_x$-$(CS)_x$-$(S)_x$-(FVIII_1-745)-$(S)_y$-$(XTEN)_y$-$(S)_y$-(FVIII_1635-2332)-$(S)_z$-$(CS)_z$-$(XTEN)_z$ VII

wherein independently for each occurrence, SP is a signal peptide, preferably with sequence MQIELSTCFFLCLLRFCFS (SEQ ID NO: 1611), CS is a cleavage sequence listed in Table 12, S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include amino acids compatible with restrictions sites, "FVIII_1-745" is residues 1-745 of Factor FVIII and "FVIII_1635-2332" is residues 1635-2332 of FVIII, x is either 0 or 1, y is either 0 or 1, and z is either 0 or 1, wherein $x+y+z > 2$; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity sequences set forth in Table 4. In one embodiment of formula VII, the spacer sequence is GPEGPS (SEQ ID NO: 1612). In another embodiment of formula V, the spacer sequence is glycine or a sequence selected from Tables 11 and 12.

[00220] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula VIII:

(A1)-$(S)_a$-$(XTEN)_v$-$(S)_a$-(A2)-(B1)-$(S)_b$-$(XTEN)_w$-$(S)_b$-(B2)-(A3)-$(S)_c$-$(XTEN)_x$-$(S)_c$-(C1)-$(S)_d$-$(XTEN)_y$-$(S)_d$-(C2)-$(S)_e$-$(XTEN)_z$ VIII

wherein independently for each occurrence, A1 is an A1 domain of FVIII; A2 is an A2 domain of FVIII; B1 is a fragment of the B domain that can have from residue 741 to 743-750 of FVIII or alternatively

from about residue 741 to about residues 745 of FVIII; B2 is a fragment of the B domain that can have from residues 1635-1686 to 1689 of FVIII or alternatively from about residue 1640 to about residues 1689 of FVIII; A3 is an A3 domain of FVIII; C1 is a C1 domain of FVIII; C2 is a C2 domain of FVIII; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; a is either 0 or 1; b is either 0 or 1; c is either 0 or 1; d is either 0 or 1; e is either 0 or 1; f is either 0 or 1; u is either 0 or 1; v is either 0 or 1; w is 0 or 1, x is either 0 or 1; y is either 0 or 1; z is either 0 or 1 with the proviso that u + v + w+ x + y + z $\geq$1; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity to sequences set forth in Table 4. In one embodiment of formula VIII, the spacer sequence is GPEGPS (SEQ ID NO: 1612). In another embodiment of formula V, the spacer sequence is glycine or a sequence selected from Tables 11 and 12.

[00221] In another embodiment of the CFXTEN composition, the invention provides a recombinant factor VIII fusion protein of formula IX:

$(A1_N)$-$(S)_a$-$(XTEN)_t$-$(S)_b$-$(A1_C)$-$(A2_N)$-$(S)_c$-$(XTEN)_u$-$(S)_d$-$(A2_C)$-(B_N)-$(S)_e$-$(XTEN)_v$-$(S)_f$-(B_C)-$(A3_N)$-$(S)_g$-$(XTEN)_w$-$(S)_h$-$(A3_C)$-$(C1_N)$-$(S)_i$-$(XTEN)_x$-$(S)_j$-$(C1_C)$-$(C2_N)$-$(S)_k$-$(XTEN)_y$-$(S)_l$-$(C2_C)$-$(S)_m$-$(XTEN)_z$

IX

wherein independently for each occurrence, $A1_N$ is a fragment of the A1 domain from at least residue number 1 (numbered relative to native, mature FVIII) to no more than residue number 371, $A1_c$ is a fragment of the A1 domain from at least residue number 2 to no more than residue number 372; $A2_N$ is a fragment of the A2 domain from at least residue number 373 to no more than residue number 739, $A2_c$ is a fragment of the A2 domain from at least residue number 374 to no more than residue number 740; B_N is a fragment of the B domain from at least residue number 741 to no more than residue number 1647, B_c is a fragment of the B domain from at least residue number 742 to no more than residue number 1648; $A3_N$ is a fragment of the A3 domain from at least residue number 1649 to no more than residue number 2019, $A3_c$ is a fragment of the A3 domain from at least residue number 1650 to no more than residue number 2019; $C1_N$ is a fragment of the C1 domain from at least residue number 2020 to no more than residue number 2171, $C1_c$ is a fragment of the C1 domain from at least residue number 2021 to no more than residue number 2172; $C2_N$ is a fragment of the C2 domain from at least residue number 2173 to no more than residue number 2331, $C2_c$ is a fragment of the C2 domain from at least residue number 2174 to no more than residue number 2332; S is a spacer sequence having between 1 to about 50 amino acid residues that can optionally include a cleavage sequence or amino acids compatible with restrictions sites; a is either 0 or 1; b is either 0 or 1; c is either 0 or 1; d is either 0 or 1; e is either 0 or 1; f is either 0 or 1; g is either 0 or 1; h is either 0 or 1; i is either 0 or 1; j is either 0 or 1; k is either 0 or 1; l is either 0 or 1; m is either 0 or 1; t is either 0 or 1; u is either 0 or 1; v is either 0 or 1; w is 0 or 1, x is either 0 or 1; y is either 0 or 1; z is either 0 or 1 with the proviso that t + u + v + w+ x + y + z $\geq$1; and XTEN is an extended recombinant polypeptide as described herein including, but not limited to sequences having at

least 90% identity to sequences set forth in Table 4. In one embodiment of formula IX, the spacer sequence is GPEGPS (SEQ ID NO: 1612). In another embodiment of formula IX, the spacer sequence is glycine or a sequence selected from Tables 11 and 12.

[00222] The embodiments of formulae IV-VIII encompass CFXTEN configurations wherein one or more XTEN of lengths ranging from about 6 amino acids to ≥1000 amino acids (e.g., sequences selected from any one of Tables 3, 4, and 13-17 or fragments thereof, or sequences exhibiting at least about 90-99% or more sequence identity thereto) are inserted and linked between adjoining domains of the factor VIII or are linked to the N- or C-terminus of the FVIII. In other embodiments of formulae V-VIII, the invention further provides configurations wherein the XTEN are linked to FVIII domains via spacer sequences which can optionally comprise amino acids compatible with restrictions sites or can include cleavage sequences (e.g., the sequences of Tables 11 and 12, described more fully below) such that the XTEN encoding sequence can be, in the case of a restriction site, integrated into a CFXTEN construct and, in the case of a cleavage sequence, the XTEN can be released from the fusion protein by the action of a protease appropriate for the cleavage sequence.

[00223] The embodiments of formulae VI -VIII differ from those of formula V in that the FVIII component of formulae VI-VIII are only the B-domain deleted forms ("FVIII BDD") of factor VIII that retain short residual sequences of the B-domain, non-limiting examples of sequences of which are provided in Table 1, wherein one or more XTEN or fragments of XTEN of lengths ranging from about 6 amino acids to ≥1000 amino acids (e.g., sequences selected from any one of Tables 3, 4, and 13-17) are inserted and linked between adjoining domains of the factor VIII and/or between the remnants of the B domain residues, such as those of Table 8. The embodiment of formula IX generally differs from those of the other formulae in that the one or more XTEN are each inserted within domains of FVIII rather than between domains, and/or has an XTEN linked to the C-terminus of the FVIII (or is linked via a spacer sequence to the C-terminus of the FVIII).

[00224] In some embodiments of a CFXTEN, the fusion protein comprises a B-domain deleted form of FVIII wherein the B-domain deletion starts from a first position at about amino acid residue number 745 and ends at a second position at amino acid residue number 1635 to about 1690 with reference to the full-length human factor VIII sequence and an XTEN links the first position and the second position of the B-domain deletion. In one embodiment of the foregoing, the first position and the second position of the B-domain deletion are selected from the positions of Table 8. In another embodiment of the foregoing, at least one XTEN links the first and second position wherein the at least one XTEN links factor VIII amino acid residue 745 and amino acid residue 1640, or amino acid residue 741 and amino acid residue 1640, or amino acid residue 741 and amino acid residue 1690, or amino acid residue 745 and amino acid residue 1667, or amino acid residue 745 and amino acid residue 1657, or amino acid residue 745 and amino acid residue 1657, or amino acid residue 747 and amino acid residue 1642, or amino acid residue 751 and amino acid residue 1667. In one embodiment of the CFXTEN, wherein the factor VIII comprises an XTEN linking a first position and a second position of a B-domain deletion described in the embodiments of this paragraph, the XTEN is a sequence having at least 80%, or at least about 90%, or at

least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or 100% sequence identity compared to a sequence of comparable length selected from any one of Table 4, Table 13, Table 14, Table 15, Table 16, and Table 17, when optimally aligned, wherein the CFXTEN retains at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% of the procoagulant activity of native FVIII.

[00225] The invention contemplates all possible permutations of insertions of XTEN between or within the domains of FVIII or at or proximal to the insertion points of Table 5, Table 6, Table 7, Table 8, and Table 9 or those illustrated in FIGS. 8-9, with optional linking of an additional XTEN to the N- or C-terminus of the FVIII, optionally linked via an additional cleavage sequence selected from Table 12, resulting in a CFXTEN composition; non-limiting examples of which are portrayed in FIGS. 5 and 12. In one embodiment, the CFXTEN comprises a FVIII BDD sequence of Table 1 in which one or more XTEN that each has at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% or more sequence identity compared to a sequence from any one of Tables 3, 4, and 13-17 or fragments thereof are inserted between any two of the residual B domain amino acids of the FVIII BDD sequence, resulting in a single chain FVIII fusion protein, wherein the CFXTEN retains at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% of the procoagulant activity of native FVIII. In the foregoing embodiment, the CFXTEN can have an additional XTEN sequence of any one of Tables 4, and 13-17 linked to the N- or C-terminus of the fusion protein. In another embodiment, a CFXTEN comprises at least a first XTEN inserted at a site set forth in Table 8, wherein the CFXTEN retains at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% of the procoagulant activity of native FVIII. In one embodiment of a fusion protein of formula VII, the CFXTEN comprises a FVIII BDD sequence of Table 1 in which two or more XTEN that each has at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99%, or 100% sequence identity compared to a sequence from any one of Tables 3, 4, and 13-17 or fragments thereof are linked to a FVIII-BDD sequence in which at least one XTEN is inserted from about 3 to about 20 amino acid residues to the C-terminus side of the FVIII cleavage site amino acid R740 and from about 3 to about 20 amino acid residues to the N-terminus side of the FVIII cleavage site amino acid R1689 of the residual B domain amino acids of the FVIII BDD sequence, resulting in a single chain FVIII fusion protein, and one or two XTEN are linked by a cleavage sequence to the N- and/or C-terminus of the FVIII-BDD sequence, wherein the CFXTEN exhibits at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% of the procoagulant activity of native FVIII after release of the XTEN by cleavage of the cleavage sequences.

[00226] In one embodiment, the A3 domain comprises an a3 acidic region or a portion thereof. In another embodiment, at least one XTEN is inserted within the a3 acidic region or the portion thereof, N-terminus of the a3 acidic region or the portion thereof, C-terminus of the a3 acidic region or the portion

thereof, or a combination thereof. In certain embodiments, at least one XTEN is inserted within the C2 domain, N-terminus of C2 domain, C-terminus of C2 domain, or a combination thereof. In still other embodiments, the Factor VIII comprises all or portion of B domain. In yet other embodiments, at least one XTEN is inserted within all or a portion of B domain, N-terminus of B domain, C-terminus of B domain, or a combination thereof.

2. CFXTEN Fusion Protein Configurations with Internal XTEN

[00227] In another aspect, the invention provides CFXTEN configured with one or more XTEN sequences located internal to the FVIII sequence. In one embodiment, invention provides CFXTEN configured with one or more XTEN sequences located internal to the FVIII sequence to confer properties such as, but not limited to, increased stability, increased resistance to proteases, increased resistance to clearance mechanisms including but not limiting to interaction with clearance receptors or FVIII inhibitors, and increased hydrophilicity, compared to FVIII without the incorporated XTEN.

[00228] The invention contemplates that different configurations or sequence variants of FVIII can be utilized as the platform into which one or more XTEN are inserted. These configurations include, but are not limited to, native FVIII, FVIII BDD, and single chain FVIII (scFVIII), and variants of those configurations. In the case of scFVIII, the invention provides CFXTEN that can be constructed by replacing one or multiple amino acids of the processing site of FVIII. In one embodiment, the scFVIII utilized in the CFXTEN is created by replacing the R1648 in the FVIII sequence RHQREITR (SEQ ID NO: 1698) with glycine or alanine to prevent proteolytic processing to the heterodimer form. It is specifically contemplated that any of the CFXTEN embodiments disclosed herein with a 1648 FVIII residue can have a glycine or alanine substitution for the arginine at position 1648. In some embodiments, the invention provides CFXTEN comprising scFVIII wherein parts of the sequence surrounding the R1648 processing site are replaced with XTEN, as illustrated in FIGS. 10A and 10B. In one embodiment, at least about 60%, or about 70%, or about 80%, or about 90%, or about 95%, or about 97% or more of the B-domain is replaced with an XTEN sequence disclosed herein, including one or more of the R740, R1648, or R1689 cleavage sites. In another embodiment, the CFXTEN has the FVIII sequence of the B-domain between the FXIa cleavage sites at R740 and R1689 (with at least 1-5 adjacent B-domain amino acids also retained between the cut site and the start of the XTEN to permit the protease to access the cut site) replaced with XTEN. In another embodiment, the CFXTEN has the FVIII sequence of the B-domain between the FXIa cleavage site at N745 and P1640 replaced with XTEN. In other embodiments, the invention provides CFXTEN FVIII BDD sequence variants in which portions of the B-domain are deleted but only one of the FXI R740 or R1689 activation sites (and 1-5 adjacent amino acids of the B-domain) are left within the construct, wherein the XTEN remains attached at one end to either the light or heavy chain after cleavage by FXIa, as illustrated in FIG. 5B and 5D. In one embodiment of the foregoing, the CFXTEN comprises a FVIII BDD sequence in which the amino acids between N745 to P1640 or between S743 to Q1638 or between P747 to V1642 or between N745 and Q1656 or between N745 and S1657 or between N745 and T1667 or between N745 and Q1686 or between R747 and V1642 or between T751 and T1667 are deleted and an XTEN sequence is linked

between these amino acids, connecting the heavy and light chains, and can further comprise additional XTEN inserted either in external surface loops, between FVIII domains, or at the N- or C-termini of the FVIII BDD sequence, such as one or more insertion sites from Table 5, Table 6, Table 7, Table 8, and Table 9 or those illustrated in FIGS. 8-9. In another embodiment of the foregoing, the CFXTEN comprises a FVIII BDD sequence in which the amino acids between K713 to Q1686 or between residues 741 and 1648 are deleted and an XTEN linked between the two amino acids, and additional XTEN can be inserted either in surface loops, between FVIII domains, or at the N- or C-termini of the FVIII BDD sequence, including but not limited to one or more insertion sites from Table 5, Table 6, Table 7, Table 8, and Table 9 or those illustrated in FIGS. 8-9. In some embodiments such CFXTEN sequences can have one or more XTEN exhibiting at least about 80%, or at least about 90%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99%, or 100% sequence identity to an XTEN sequence from any one of Tables 4 and 13-17.

[00229] The invention contemplates other CFXTEN with internal XTEN in various configurations; schematics of exemplary configurations are illustrated in FIGS. 5 and 10. The regions suitable for XTEN insertion sites include the known domain boundaries of FVIII, exon boundaries, known surface (external) loops and solvent accessible surface area sites identified by X-ray crystallography analysis, and structure models derived from molecular dynamic simulations of FVIII, regions with a low degree of order (assessed by programs described in FIGS. 7 legend), regions of low homology/lack of conservation across different species, and hydrophilic regions. In another embodiment, XTEN insertion sites were selected based on FVIII putative clearance receptor binding sites. In another embodiment, CFXTEN comprises XTEN inserted at locations not within close proximity to mutations implicated in hemophilia A listed in the Haemophilia A Mutation, Search, Test and Resource Site (HAMSTeRS) database were eliminated (Kemball-Cook G, et al. The factor VIII Structure and Mutation Resource Site: HAMSTeRS version 4. Nucleic Acids Res. (1998) 26(1):216-219). In another embodiment, potential sites for XTEN insertion include residues within FVIII epitopes that are capable of being bound by anti-FVIII antibodies occurring in sensitized hemophiliacs and that do not otherwise serve as protein interactive sites. Regions and/or sites that are considered for exclusion as XTEN insertion sites include residues/regions of factor VIII that are important in various interactions including other clotting proteins, residues surrounding each arginine activating/inactivating cleavage site acted on by the proteases thrombin, factor Xa, activated protein C, residues surrounding the signal peptide processing site (residue 1) if the construct contains the signal peptide, regions known to interact with other proteins such as FIXa, FX/FXa, thrombin, activated protein C, protein S cofactor to Protein C, von Willebrand factor, sites known to interact with phospholipid cofactors in coagulation, residues involved in domain interactions, residues coordinating Ca^{++} or Cu^{++} ions, cysteine residues involved in S-S intramolecular bonds, documented amino acid insertion and point mutation sites in FVIII produced in hemophilia A subjects affecting procoagulant activity, and mutation sites in FVIII made in a research lab that affect procoagulant activity. Sites considered for either insertion (to prolong half-life) or for exclusion (needed to remove spent FVIIIa or

FXa) include regions known to interact with heparin sulfate proteoglycan (HSPG) or low-density lipoprotein receptor-related protein (LPR).

[00230] By analysis of the foregoing criteria, as described in Example 34, different insertion sites or ranges of insertions sites across the FVIII BDD sequence have been identified and/or confirmed as candidates for insertion of XTEN, non-limiting examples of which are listed in Table 5, Table 6, Table 7, Table 8, and Table 9 and are shown schematically in FIGS. 8 and 9. In one embodiment, CFXTEN comprise XTEN insertions between the individual domains of FVIII, i.e., between the A1 and A2, or between the A2 and the B, or between the B and the A3, or between the A3 and the C1, or between the C1 and the C2 domains. In another embodiment, CFXTEN comprises XTEN inserted within the B domain or between remnant residues of the BDD sequence. In another embodiment, CFXTEN comprises XTEN inserted at known exon boundaries of the encoding FVIII gene as exons represent evolutionary conserved sequence modules that have a high probability of functioning in the context of other protein sequences. In another embodiment, CFXTEN comprise XTEN inserted within surface loops identified by the x-ray structure of FVIII. In another embodiment, CFXTEN comprise XTEN inserted within regions of low order identified as having low or no detected electron density by X-ray structure analysis. In another embodiment, CFXTEN comprise XTEN inserted within regions of low order, predicted by structure prediction algorithms such as, but not limited to FoldIndex, RONN, and Kyte & Doolitlle algorithms. In another embodiment, CFXTEN comprise XTEN inserted within sequence areas of high frequency of hydrophilic amino acids. In another embodiment, CFXTEN comprise XTEN inserted within epitopes capable of being bound by naturally-occurring anti-FVIII antibodies in sensitized hemophiliacs. In another embodiment, CFXTEN comprise XTEN inserted within sequence areas of low sequence conservation and/or differences in sequence segment length across FVIII sequences from different species. In another embodiment, CFXTEN comprise XTEN linked to the N-terminus and/or C-terminus. In another embodiment, the invention provides CFXTEN configurations with inserted XTEN selected from two or more of the criteria from the embodiments listed above. In another embodiment, the invention provides CFXTEN configurations with at least one, alternatively at least two, alternatively at least three, alternatively at least four, alternatively at least five or more XTEN inserted into a factor VIII sequence wherein the points of insertion are at or proximal to the N- or C-terminus side of the at least one, two, three, four, or five, or six or more amino acids selected from the insertion residue amino acids of Table 5, Table 6, Table 7, Table 8, and Table 9 or those illustrated in FIGS. 8-9, or alternatively within one, or within two, or within three, or within four, or within five, or within six amino acids of the insertion residue amino acids from Table 5, Table 6, Table 7, Table 8, and Table 9, or within the various spans of the insertion residue amino acids schematically portrayed for an exemplary FVIII BDD sequence in FIG. 9.

[00231] As described above, the one or more internally-located XTEN or a fragment of XTEN can have a sequence length of 6 to 1000 or more amino acid residues. In some embodiments, wherein the CFXTEN have one or two or three or four or five or more XTEN sequences internal to the FVIII, the XTEN sequences can be identical or can be different. In one embodiment, each internally-located

XTEN has at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to comparable lengths or fragments of XTEN or motifs selected from any one of Tables 3, 4, and 13-17, when optimally aligned. In another embodiment, the invention provides a CFXTEN configured with one or more XTEN inserted internal to a FVIII BDD sequence with at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to a sequence of Table 1, wherein the insertions are located at the insertion points or range of insertion points indicated in Table 5, Table 6, Table 7, Table 8, and Table 9, FIG. 8 or within the range of insertions as illustrated in FIG. 9. It will be understood by those of skill in the art that an XTEN inserted within the FVIII sequence at an insertion point of Table 5, Table 6, Table 7, Table 8, and Table 9 is linked by its N- and C-termini to flanking FVIII amino acids (or via a linking spacer or cleavage sequences, as described above), while an XTEN linked to the N- or C-terminus of FVIII would only be linked to a single FVIII amino acid (or to a linking spacer or cleavage sequence amino acid, as described above). By way of example only, variations of CFXTEN with three internal XTEN could have: XTEN (as described herein) incorporated between FVIII BDD residues 741 and 1640, residues 18 and 19, and residues 1656 and 1657; or XTEN incorporated between FVIII BDD residues 741 and 1640, residues 1900 and 1901, and at the C-terminus at residue 2332; or XTEN incorporated between FVIII BDD residues 26 and 27, residues 1656 and 1657, and residues 1900 and 1901; or XTEN incorporated between FVIII BDD residues 741 and 1640, residues 1900 and 1901, and at the C-terminus at residue 2332.

[00232] In evaluating the CFXTEN fusion proteins with XTEN inserted in the locations from Table 5, it was discovered that insertions in certain regions of the FVIII sequence resulted in CFXTEN with good expression and retention of procoagulant activity. Accordingly, in preferred embodiments, the invention provides CFXTEN fusion proteins configured with one, or two, or three, or four, or five, or six or more XTEN, each having at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to an XTEN selected from any one of Tables 4, and 13-17 inserted internal or linked to a FVIII BDD sequence with at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to a sequence of Table 1, wherein the insertions are located at an insertion point within one, or two, or three, or four, or five, or six or more ranges set forth in Table 7. In the foregoing embodiments, the CFXTEN fusion proteins with the XTEN insertions retain at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% or more of the procoagulant activity compared to the corresponding FVIII not linked to XTEN.

[00233] In evaluating the CFXTEN fusion proteins with XTEN inserted in one or more locations from Table 5, it was surprisingly discovered that a high percentage of fusion proteins with the XTEN insertions retained procoagulant activity, as described in Example 25. Accordingly, the invention provides CFXTEN fusion proteins configured with one, two, three, four, five, six or more XTEN wherein

the resulting fusion protein exhibits at least about 10%, or 20%, or 30%, or 40%, or 50%, or 60%, or 70%, or 80%, or 90% or more of the procoagulant activity compared to the corresponding FVIII not linked to XTEN when assayed by a coagulation assay described herein. In a preferred embodiment, the invention provides CFXTEN fusion proteins comprising one, or two, or three, or four, or five, or six or more XTEN, each having at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to an XTEN selected from any one of Tables 4, and 13-17 linked to a FVIII BDD sequence with at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to a sequence of Table 1, wherein the insertions are located at one or more insertion points selected from Table 5, Table 6, Table 7, Table 8, and Table 9, and wherein the resulting fusion protein exhibits at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70% or more procoagulant activity compared to the corresponding FVIII not linked to XTEN, when assayed *in vitro* by an assay described herein (e.g., a chromogenic assay). As the subject CFXTEN fusion proteins typically exhibit increased terminal half-life compared to native FVIII, it will be appreciated by one of skill in the art that a CFXTEN with lower procoagulant activity relative to an equimolar amount of native FVIII would nevertheless be acceptable when administered as a therapeutic composition to a subject in need therof. In another embodiment, the CFXTEN fusion proteins comprising one, or two, or three, or four, or five or more XTEN, each having at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to an XTEN selected from any one of Tables 4, and 13-17 linked to a FVIII BDD sequence with at least about 80% sequence identity, or alternatively 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity compared to a sequence of Table 1, wherein the insertions are located at one or more insertion points or the range of insertion points selected from Table 5, Table 6, Table 7, Table 8, and Table 9, wherein the resulting fusion protein exhibits at least about 0.5 IU/ml, or at least about 0.75 IU/ml, or at least about 1.0 IU/ml, or at least about 1.5 IU/ml, or at least about 2.0 IU/ml, or at least about 2.5 IU/ml, or at least about 3 IU/ml, or at least about 4 IU/ml, or at least about 5 IU/ml, or at least about 7 IU/ml, or at least about 10 IU/ml, or at least about 20 IU/ml, or at least about 30 IU/ml FVIII activity when expressed in cell culture medium and assayed in a chromogenic assay, wherein the culture and expression are according to methods described herein; e.g., the methods of Example 25.

[00234] It is believed that the discovery of the insertions sites wherein the FVIII retains at least a portion of its procoagulant activity would also permit the insertion of other peptides and polypeptides with either unstructured or structured characteristics that are associated with the prolongation of half-life when fused to a FVIII protein in one or more of those same sites. Non-limiting examples include albumin, albumin fragments, Fc fragments of immunoglobulins, the β subunit of the C-terminal peptide (CTP) of human chorionic gonadotropin, a HAP sequence, a transferrin, the PAS polypeptides of U.S.

Pat Application No. 20100292130, polyglycine linkers, polyserine linkers, peptides and short polypeptides of 6-40 amino acids of two types of amino acids selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P) with varying degrees of secondary structure from less than 50% to greater than 50%, amongst others, would be suitable for insertion in the identified active insertions sites of FVIII.

[00235] In the fusion protein embodiments described herein, the CFXTEN fusion protein can further comprise one or more cleavage sequence from Table 12 or other sequences known in the art, the cleavage sequence being located between or within 6 amino acid residues of the intersection of the FVIII and the XTEN sequences, which may include two cleavage sequences in a given internal XTEN sequence. In one embodiment, the CFXTEN comprising cleavage sequences has two identical cleavage sequences, each located at or near the respective ends of one or more internal XTEN such that the XTEN is released from the fusion protein when cleaved by the protease that binds to and cleaves that sequence. The sequences that can be cleaved are described more fully below and exemplary sequences are provided in Table 12.

Table 5: Insertion locations for XTEN linked to the FVIII BDD sequence

No.	XTEN Insertion Point	Insertion Residue	FVIII BDD Downstream Sequence	FVIII Domain
1	0	(N-terminus)	ATR	A1
2	3	R	RYY	A1
3	17	M	QSD	A1
4	18	Q	SDL	A1
5	22	G	ELP	A1
6	24	L	PVD	A1
7	26	V	DAR	A1
8	28	A	RFP	A1
9	32	P	RVP	A1
10	38	F	PFN	A1
11	40	F	NTS	A1
12	41	N	TSV	A1
13	60	N	IAK	A1
14	61	I	AKP	A1
15	65	R	PPW	A1
16	81	Y	DTV	A1
17	111	G	AEY	A1
18	116	D	QTS	A1
19	119	S	QRE	A1
20	120	Q	REK	A1
21	128	V	FPG	A1
22	129	F	PGG	A1
23	130	P	GGS	A1
24	182	G	SLA	A1
25	185	A	KEK	A1
26	188	K	TQT	A1
27	205	G	KSW	A1
28	210	S	ETK	A1
29	211	E	TKN	A1
30	216	L	MQD	A1
31	220	R	DAA	A1
32	222	A	ASA	A1

No.	XTEN Insertion Point	Insertion Residue	FVIII BDD Downstream Sequence	FVIII Domain
33	223	A	SAR	A1
34	224	S	ARA	A1
35	230	K	MHT	A1
36	243	P	GLI	A1
37	244	G	LIG	A1
38	250	R	KSV	A1
39	318	D	GME	A1
40	333	P	QLR	A1
42	334	Q	LRM	A1
43	336	R	MKN	a1
44	339	N	NEE	a1
45	345	D	YDD	a1
46	357	V	VRF	a1
47	367	S	FIQ	a1
48	370	S	RPY	a1
49	375	A	KKH	A2
50	376	K	KHP	A2
51	378	H	PKT	A2
52	399	V	LAP	A2
53	403	D	DRS	A2
54	405	R	SYK	A2
55	409	S	QYL	A2
56	416	P	QRI	A2
57	434	E	TFK	A2
58	438	T	REA	A2
59	441	A	IQH	A2
60	442	I	QHE	A2
61	463	I	IFK	A2
62	487	Y	SRR	A2
63	490	R	LPK	A2
64	492	P	KGV	A2
65	493	K	GVK	A2
66	494	G	VKH	A2
67	500	D	FPI	A2
68	506	G	EIF	A2
69	518	E	DGP	A2
70	556	K	ESV	A2
71	565	Q	IMS	A2
72	566	I	MSD	A2
73	598	P	AGV	A2
74	599	A	GVQ	A2
75	603	L	EDP	A2
76	616	S	ING	A2
77	686	G	LWI	A2
78	713	K	NTG	A2
79	719	Y	EDS	A2
80	730	L	LSK	A2
81	733	K	NNA	A2
82	745	N	PPV	B
83	1640	P	PVL	B
84	1652	R	TTL	B
85	1656	Q	SDQ	A3
86	1685	N	QSP	A3

No.	XTEN Insertion Point	Insertion Residue	FVIII BDD Downstream Sequence	FVIII Domain
87	1711	M	SSS	A3
88	1713	S	SPH	A3
89	1720	N	RAQ	A3
90	1724	S	GSV	A3
91	1725	G	SVP	A3
92	1726	S	VPQ	A3
93	1741	G	SFT	A3
94	1744	T	QPL	A3
95	1749	R	GEL	A3
96	1773	V	TFR	A3
97	1792	Y	EED	A3
98	1793	E	EDQ	A3
99	1796	Q	RQG	A3
100	1798	Q	GAE	A3
101	1799	G	AEP	A3
102	1802	P	RKN	A3
103	1803	R	KNF	A3
104	1807	V	KPN	A3
105	1808	K	PNE	A3
106	1827	K	DEF	A3
107	1844	E	KDV	A3
108	1861	N	TLN	A3
109	1863	L	NPA	A3
110	1896	E	RNC	A3
111	1900	R	APC	A3
112	1904	N	IQM	A3
113	1905	I	QME	A3
114	1910	P	TFK	A3
115	1920	A	ING	A3
116	1937	D	QRI	A3
117	1981	G	VFE	A3
118	2019	N	KCQ	A3
119	2020	K	CQT	C1
120	2044	G	QWA	C1
121	2068	F	SWI	C1
122	2073	V	DLL	C1
123	2090	R	QKF	C1
124	2092	K	FSS	C1
125	2093	F	SSL	C1
126	2111	K	WQT	C1
127	2115	Y	RGN	C1
128	2120	T	GTL	C1
129	2125	V	FFG	C1
130	2171	L	NSC	C1
131	2173	S	CSM	C2
132	2188	A	QIT	C2
133	2223	V	NNP	C2
134	2224	N	NPK	C2
135	2227	K	EWL	C2
136	2268	G	HQW	C2
137	2277	N	GKV	C2
138	2278	G	KVK	C2
139	2290	F	TPV	C2
140	2332	Y	C terminus of FVIII	CT

Indicates an insertion point for XTEN based on the amino acid number of mature full-length human FVIII, wherein the insertion could be either on the N- or C-terminal side of the indicated amino acid

Downstream sequence in FVIII BDD with 746-1639 deletion

Table 6. Exemplary insertion locations for XTEN linked to a FVIII polypeptide

No.	XTEN Insertion Point	Insertion Residue	FVIII BDD Downstream Sequence	FVIII Domain	Distance from insertion residue
9	32	P	RVP	A1	-3, +6
31	220	R	DAA	A1	-
34	224	S	ARA	A1	+5
43	336	R	MKN	a1	-1, +6
44	339	N	NEE	a1	-4, +5
52	399	V	LAP	A2	-6, +3
56	416	P	QRI	A2	+6
75	603	L	EDP	A2	_6, +6
85	1656	Q	SDQ	B	-3, +6
87	1711	M	SSS	A3	-6, +1
91	1725	G	SVP	A3	+6
113	1905	I	QME	A3	+6
114	1910	P	TFK	A3	-5, +6

Distance from insertion residue refers to the relative number of amino acids away from the N-terminus (negative numbers) or C-terminus (positive numbers) of the designated insertion residue (residue "0") where an insertion may be made. The designation "-x" refers to an insertion site which is x amino acids away on the N-terminal side of the designated insertion residue. Similarly, the designation "+x" refers to an insertion site which is x amino acids away on the C-terminal side of the designated insertion residue.

For example, "-1, +2" indicates that the insertion is made at the N-terminus or C-terminus of amino acid residues denoted -1, 0, +1 or +2.

Table 7. Further exemplary insertion locations for XTEN linked to a FVIII polypeptide

No.	XTEN Insertion Point Range	First Insertion Residue	FVIII Domain
3	18-32	Q	A1
8	40	F	A1
18	211-224	E	A1
27	336-403	R	A1, A2
43	599	A	A2
47	745-1640	N	B
50	1656-1728	Q	B, A3
57	1796-1804	R	A3
65	1900-1912	R	A3
81	2171-2332	L	C1, C2

indicates range of insertion sites numbered relative to the amino acid number of mature human FVIII

Table 8. Exemplary XTEN insertion locations within B-domain deleted variants of a FVIII polypeptide

XTEN Insertion Point Range	First Insertion Residue	Second Insertion Residue
740-1640	R	P
740-1690	R	S
741-1648	S	R
743-1638	S	Q
745-1638	N	Q
745-1640	N	P
745-1656	N	Q
745-1657	N	S
745-1667	N	T
745-1686	N	Q
747-1642	R	V
751-1667	T	T

indicates the amino acids linked within the B-domain deleted variant and adjacent A3 domain, with the amino acids numbered relative to the amino acid number of mature human FVIII

indicates the amino acids linked by an XTEN inserted in the BDD-FVIII

Table 9. Exemplary insertion locations for XTEN linked to a FVIII polypeptide resulting in procoagulant activity

No.	XTEN Insertion Point	Insertion Residue	FVIII BDD Downstream Sequence	FVIII Domain
2	3	R	RYY	A1
4	18	Q	SDL	A1
5	22	G	ELP	A1
7	26	V	DAR	A1
9	32	P	RVP	A1
11	40	F	NTS	A1
18	116	D	QTS	A1
19	119	S	QRE	A1
26	188	K	TQT	A1
29	211	E	TKN	A1
30	216	L	MQD	A1
31	220	R	DAA	A1
34	224	S	ARA	A1
35	230	K	MHT	A1
40	333	P	QLR	A1
43	336	R	MKN	a1
44	339	N	NEE	a1
52	399	V	LAP	A2
53	403	D	DRS	A2
55	409	S	QYL	A2
56	416	P	QRI	A2
60	442	I	QHE	A2
62	487	Y	SRR	A2
63	490	R	LPK	A2
66	494	G	VKH	A2
69	518	E	DGP	A2
74	599	A	GVQ	A2
75	603	L	EDP	A2

No.	XTEN Insertion Point	Insertion Residue	FVIII BDD Downstream Sequence	FVIII Domain
78	713	K	NTG	A2
82	745	N	PPV	B
85	1656	Q	SDQ	A3
87	1711	M	SSS	A3
89	1720	N	RAQ	A3
91	1725	G	SVP	A3
99	1796	Q	RQG	A3
102	1802	P	RKN	A3
110	1896	E	RNC	A3
111	1900	R	APC	A3
112	1904	N	IQM	A3
113	1905	I	QME	A3
114	1910	P	TFK	A3
121	2068	F	SWI	C1
130	2171	L	NSC	C1
135	2227	K	EWL	C2
137	2277	N	GKV	C2
140	2332	Y	C terminus of FVIII	C2

Downstream sequence in FVIII BDD with 746-1639 deletion

[00236] In another aspect, the invention provides libraries of components and methods to create the libraries derived from nucleotides encoding FVIII segments, XTEN, and FVIII segments linked to XTEN that are useful in the preparation of genes encoding the subject CFXTEN. In a first step, a library of genes encoding FVIII and XTEN inserted into the various single sites at or within 1-6 amino acids of an insertion site identified in Table 5 or illustrated in FIGS. 8-9 are created, expressed, and the CFXTEN recovered and evaluated for activity and pharmacokinetics as illustrated in FIG. 15. Those CFXTEN showing enhanced properties are then used to create genes encoding a FVIII segment and the insertion site plus an XTEN, with components from each enhanced insertion represented in the library, as illustrated in FIG. 11. In one embodiment, the library components are assembled using standard recombinant techniques in combinatorial fashion, as illustrated in FIG. 11, resulting in permutations of CFXTEN with multiple internal and N- and C-terminus XTEN, that can include the insertion sites of or proximal to those Table 5, Table 6, Table 7, Table 8 and Table 9, or as illustrated in FIGS. 8-9. The resulting constructs would then be evaluated for activity and enhanced pharmacokinetics, and those candidates resulting in CFXTEN with enhanced properties, e.g., reduced active clearance, resistance to proteases, reduced immunogenicity, and enhance pharmacokinetics, compared to FVIII not linked to XTEN, are evaluated further.

3. XTEN Permissive Loops

[00237] As described in detail elsewhere herein and as illustrated in FIGS.33-36, the inventors have recognized that each FVIII "A" domain comprise at least two "XTEN permissive loops" into which XTEN sequences can be inserted without eliminating procoagulant activity of the recombinant protein, or the ability of the recombinant proteins to be expressed in vivo or in vitro in a host cell. The inventors have identified the XTEN permissive loops as regions with, among other attributes, high surface or

solvent exposure and high conformational flexibility. The A1 domain comprises an XTEN permissive loop-1 (A1-1) region and an XTEN permissive loop-2 (A1-2) region, the A2 domain comprises an XTEN permissive loop-1 (A2-1) region and an XTEN permissive loop-2 (A2-2) region, the A3 domain comprises an XTEN permissive loop-1 (A3-1) region and an XTEN permissive loop-2 (A3-2) region..

[00238] In certain aspects a recombinant FVIII protein as described above comprises at least one XTEN sequence inserted into at least one of the XTEN permissive loops A1-1, A1-2, A2-1, A2-2, A3-1, or A3-2, wherein the recombinant FVIII protein has procoagulant activity and can be expressed in vivo or in vitro in a host cell. In certain aspects a recombinant FVIII protein as described above comprises at least two XTEN sequences inserted into FVIII, e.g., into two different XTEN permissive loops A1-1, A1-2, A2-1, A2-2, A3-1, or A3-2, wherein the recombinant FVIII protein has procoagulant activity and can be expressed in vivo or in vitro in a host cell. Alternatively, a recombinant FVIII protein as described above can comprise two or more XTEN sequences inserted into a single XTEN permissive loop either with our without XTEN sequences inserted into other XTEN permissive loops, wherein the recombinant FVIII protein has procoagulant activity and can be expressed in vivo or in vitro in a host cell. In certain aspects a recombinant FVIII protein as described above can comprise at least one XTEN sequence inserted into at least one of the XTEN permissive loops as described above, and can further comprise one or more XTEN sequences inserted into a3, wherein the recombinant FVIII protein has procoagulant activity and can be expressed in vivo or in vitro in a host cell. In certain aspects, a recombinant FVIII protein of the invention can comprise three, four, five, six or more XTEN sequences inserted into one or more XTEN permissive loops or into a3, wherein the recombinant FVIII protein has procoagulant activity and can be expressed in vivo or in vitro in a host cell.

[00239] In certain aspects a recombinant FVIII protein as described above comprises at least one XTEN sequence inserted into a3, wherein the recombinant FVIII protein has procoagulant activity and can be expressed in vivo or in vitro in a host cell. In certain aspects a recombinant FVIII protein of the invention comprises at least one XTEN sequence inserted into a3, and further comprises one or more XTEN sequences inserted into one or more XTEN permissive loops as described above, wherein the recombinant FVIII protein has procoagulant activity and can be expressed in vivo or in vitro in a host cell.

[00240] The inventors have recognized that a recombinant FVIII protein of the invention comprises at least two XTEN permissive loops in each of the FVIII A domain regions which allows for insertion of an XTEN sequence while having procoagulant activity and still being able to be expressed in vivo or in vitro by a host cell. Various crystal structures of FVIII have been determined, of varying degrees of resolution. These structures of FVIII and FVIIIa, determined by X-ray crystallography and molecular dynamic simulation, were used to generate models of accessible surface area and conformational flexibility for FVIII. For example, the crystal structure of human FVIII has been determined by Shen et al. Blood 111: 1240-1247 (2008) and Ngo et al. Structure 16: 597-606 (2008). The data for these structures is available from the Protein Data Bank (pdb.org) under Accession Numbers 2R7E and 3CDZ, respectively.

[00241] The predicted secondary structure of the heavy and light chains of human FVIII according to the Shen et al. crystal structure is reproduced in FIGS. 37A and 37B. The various beta strands predicted from the Shen et al. crystal structure are numbered consecutively in FIGS. 8A and 8B. In certain embodiments, the XTEN permissive loops A1-1, A1-2, A2-1, A2-2, A3-1, and A3-2 are contained within surface-exposed, flexible loop structures in the A domains of FVIII. A1-1 is located between beta strand 1 and beta strand 2, A1-2 is located between beta strand 11 and beta strand 12, A2-1 is located between beta strand 22 and beta strand 23, A2-2 is located between beta strand 32 and beta strand 33, A3-1 is located between beta strand 38 and beta strand 39 and A3-2 is located between beta strand 45 and beta strand 46, according to the secondary structure of mature FVIII stored as Accession Number 2R7E of the PDB database (PDB:2R7E) and as shown in FIGS. 8A and 8B. The secondary structure of PDB Accession Number 2R7E shown in FIGS. 8A and 8B corresponds to the standardized secondary structure assignment according to the DSSP program (Kabsch and Sander, Biopolymers, 22:2577-2637 (1983)). The DSSP secondary structure of the mature FVIII stored as PDB Accession Number 2R7E can be accessed at the DSSP database, available at the world wide web site swift.cmbi.ru.nl/gv/dssp/ (last accessed February 9, 2012) (Joosten et al., 39(Suppl. 1): D411-D419 (2010)).

[00242] In certain aspects, a surface-exposed, flexible loop structure comprising A1-1 corresponds to a region in native mature human FVIII from about amino acid 15 to about amino acid 45 of FIG. 30. In certain aspects, A1-1 corresponds to a region in native mature human FVIII from about amino acid 18 to about amino acid 41 of FIG. 30. In certain aspects, the surface-exposed, flexible loop structure comprising A1-2 corresponds to a region in native mature human FVIII from about amino acid 201 to about amino acid 232 of FIG. 30. In certain aspects, A1-2 corresponds to a region in native mature human FVIII from about amino acid 218 to about amino acid 229 of FIG. 30. In certain aspects, the surface-exposed, flexible loop structure comprising A2-1 corresponds to a region in native mature human FVIII from about amino acid 395 to about amino acid 421 of FIG. 30. In certain aspects, A2-1 corresponds to a region in native mature human FVIII from about amino acid 397 to about amino acid 418 of FIG. 30. In certain aspects, the surface-exposed, flexible loop structure comprising A2-2 corresponds to a region in native mature human FVIII from about amino acid 577 to about amino acid 635 of FIG. 30. In certain aspects, A2-2 corresponds to a region in native mature human FVIII from about amino acid 595 to about amino acid 607 of FIG. 30. In certain aspects, the surface-exposed, flexible loop structure comprising A3-1 corresponds to a region in native mature human FVIII from about amino acid 1705 to about amino acid 1732 of FIG. 30. In certain aspects, A3-1 corresponds to a region in native mature human FVIII from about amino acid 1711 to about amino acid 1725 of FIG. 30. In certain aspects, the surface-exposed, flexible loop structure comprising A3-2 corresponds to a region in native mature human FVIII from about amino acid 1884 to about amino acid 1917 of FIG. 3. In certain aspects, A3-2 corresponds to a region in native mature human FVIII from about amino acid 1899 to about amino acid 1911 of FIG. 30.

[00243] In certain aspects a recombinant FVIII protein of the invention comprises one or more XTEN sequences inserted into one or more XTEN permissive loops of FVIII, or into the a3 region, wherein the

recombinant FVIII protein has procoagulant activity and can be expressed in vivo or in vitro in a host cell. XTEN sequences to be inserted include those that increase the in vivo half-life or the in vivo or in vitro stability of FVIII.

[00244] In certain aspects, a recombinant FVIII protein of the invention comprises an XTEN sequences inserted immediately downstream of one or more amino acids corresponding to one or more amino acids in mature native human FVIII including, but not limited to: amino acid 18 of FIG. 30, amino acid 26 of FIG. 30, amino acid 40 of FIG. 30, , amino acid 220 of FIG. 30, amino acid 224 of FIG. 30, amino acid 399 of FIG. 30, amino acid 403 of FIG. 30, amino acid 599 of FIG. 30, amino acid 603 of FIG. 30, amino acid 1711 of FIG. 30, amino acid 1720 of FIG. 30, amino acid 1725 of FIG. 30, amino acid 1900 of FIG. 30, amino acid 1905 of FIG. 30, amino acid 1910 of FIG. 30, or any combination thereof, including corresponding insertions in BDD-variants of FVIII described herein.

[00245] In certain aspects, a recombinant FVIII protein of the invention comprises at least one XTEN sequence inserted into the a3 region of FVIII, either alone or in combination with one or more XTEN sequences being inserted into the XTEN permissive loops of the A domains (e.g., A1-1, A1-2, A2-1, A2-2, A3-1, or A3-2 as described above), wherein the recombinant FVIII protein has procoagulant activity and can be expressed in vivo or in vitro in a host cell. In certain aspects, at least one XTEN sequence is inserted into the a3 region immediately downstream of an amino acid which corresponds to amino acid 1656 of FIG. 30. In certain aspects, a recombinant FVIII protein of the invention comprises an XTEN sequence inserted into the a3 region as described, and further includes one or more XTEN sequences inserted immediately downstream of one or more amino acids corresponding to one or more amino acids in mature native human FVIII including, but not limited to: amino acid 18 of FIG. 30, amino acid 26 of FIG. 30, amino acid 40 of FIG. 30, amino acid 220 of FIG. 30, amino acid 224 of FIG. 30, amino acid 399 of FIG. 30, amino acid 403 of FIG. 30, amino acid 599 of FIG. 30, amino acid 603 of FIG. 30, amino acid 1711 of FIG. 30, amino acid 1720 of FIG. 30, amino acid 1725 of FIG. 30, amino acid 1900 of FIG. 30, amino acid 1905 of FIG. 30, amino acid 1910 of FIG. 30, or any combination thereof.

[00246] It will be understood by one of skill in the art that the foregoing aspects of permissive loops of a native FVIII protein into which a heterologous protein can be inserted are also applicable to the B-domain deleted FVIII variants described herein; e.g., sequences set forth in Table 1. In practicing the present invention, it will be understood that a BDD-FVIII sequence of Table 1 can be substituted for the recombinant FVIII protein of the various embodiments described above, and it is believed that the resulting constructs will similarly retain procoagulant activity.

4. Interference with FVIII binding agents

[00247] It is an object of the present invention to provide procoagulant CFXTEN fusion protein compositions for use in human patients suffering from coagulopathies, such as haemophilia A, who have native or acquired antibodies, inhibitors, or other proteins or molecules that bind to FVIII that affect the activity or half-life of CFXTEN fusion proteins, wherein the CFXTEN retain a greater amount of procoagulant activity compared to the corresponding FVIII not linked to XTEN. As used herein, "FVIII binding agent" means any molecule capable of binding to native FVIII or to a recombinant factor VIII

fusion protein of the invention comprising factor VIII or a fragment thereof, whether native, derived, or produced recombinantly. It is specifically contemplated that FVIII binding agent includes anti-FVIII antibodies and FVIII inhibitors, amongst other proteins capable of specifically binding to FVIII. In one aspect, the invention provides procoagulant CFXTEN fusion proteins that exhibit reduced binding to an anti-FVIII antibody or FVIII inhibitor that interferes with the procoagulant activity of FVIII. As used herein, "anti-FVIII antibody" or "anti-factor VIII antibody" means an antibody capable of binding FVIII or a FVIII component of a CFXTEN of the invention, said antibody including but not limited to the antibodies of Table 10 or polyclonal antibody from a hemophilia A patient with FVIII inhibitors. The term antibody includes monoclonal antibodies, polyclonal antibodies, antibody fragments and antibody fragment clones. As used herein, "FVIII inhibitor" or "anti-FVIII inhibitor antibody" means an antibody capable of binding FVIII or a FVIII component of a CFXTEN of the invention and that reduces by any means the procoagulant activity of FVIII or the FVIII component of a CFXTEN. In another aspect, the invention provides CFXTEN fusion proteins that retain procoagulant activity in the presence of a FVIII inhibitor. In another aspect, the invention provides CFXTEN fusion proteins comprising FVIII that exhibit increased terminal half-life in the presence of a FVIII binding agent compared to the FVIII not linked to XTEN.

[00248] The majority of inhibitory antibodies to human factor VIII act by binding to epitopes located in the A2 domain or the C2 domain of factor VIII, disrupting specific functions associated with these domains, (U.S. Patent No. 6,770,744; Fulcher et al. Localization of human factor FVIII inhibitor epitopes to two polypeptide fragments. Proc. Natl. Acad. Sci. USA (1985) 82:7728-7732; Scandella et al. Epitope mapping of human factor VIII inhibitor antibodies by deletion analysis of fVIII fragments expressed in Escherichia coli. Proc. Natl. Acad. Sci. USA (1988) 85:6152-6156). While 68% percent of inhibitory antibodies are reported to be directed against the A2 and/or C2 domain, 3% act against the A1 domain and 46% against the a3 acidic region (Lavigne-Lissalde, G., et al. Characteristics, mechanisms of action, and epitope mapping of anti-factor VIII antibodies. Clin Rev Allergy Immunol (2009) 37:67-79). For example, certain heavy chain-specific inhibitors react with the 18.3-kD amino-terminal segment of the A2 domain (Scandella D, et al. 1988); Lollar P et al. Inhibition of human factor VIIIa by anti-A2 subunit antibodies. J Clin Invest 1994;93:2497). FVIII contains a phospholipid binding site in the C2 domain between amino acids 2302 and 2332, and there is also a von Willebrand factor binding site in the C2 domain that acts in conjunction with amino acids 1649-1689 in the A3 domain. The C2 domain also has epitopes that, when bound by inhibitors, block the activation of FVIII by thrombin or factor Xa. Inhibitors binding specifically to the light chain recognize epitopes in the A3 domain or a major antigenic region in the C2 domain and can result in reduced procoagulant activity by preventing the binding of FVIII to phospholipid or reducing the dissociation rate of FVIII from von Willebrand factor (Gilles JG, et al. Anti-factor VIII antibodies of hemophiliac patients are frequently directed towards nonfunctional determinants and do not exhibit isotypic restriction. Blood (1993) 82:2452; Shima M, et al. A factor VIII neutralizing monoclonal antibody and a human inhibitor alloantibody recognizing epitopes in the C2 domain inhibit factor VIII binding to von Willebrand factor and to phosphatidylserine. Thromb Haemost

(1993) 69:240). Non-limiting examples of monoclonal FVIII inhibitors are listed in Table 9. In patients with high-titer inhibitors, there is an increased risk of developing recurrent bleeding in particular joints, which may ultimately result in decreased quality of life, disability, or death from excessive blood loss (U.S. Pat. Application No. 20120065077; Zhang et al., Clinic. Rev. Allerg. Immunol., 37:114-124 (2009); Gouw and van den Berg, Semin. Thromb. Hemost., 35:723-734 (2009))

[00249] While not intending to be bound by any particular theory, it is believed that the unstructured characteristic of the XTEN incorporated into the CFXTEN fusion proteins permits the XTEN to adopt conformations that result in steric hindrance to inhibitors that would otherwise bind to FVIII epitopes. As illustrated in FIG. 6, as the incorporated XTEN assumes various random coil conformations, it spatially covers regions of the FVIII component of the fusion protein and sterically interferes with the ability of an inhibitor to bind to a FVIII epitope.

[00250] In one embodiment, the invention provides CFXTEN exhibiting procoagulant activity and reduced binding in the presence of an antibody binding to the C2 domain of factor VIII compared to the corresponding factor VIII not linked to XTEN and/or to native FVIII. In another embodiment, the invention provides CFXTEN exhibiting procoagulant activity and reduced binding in the presence of an antibody binding to the A2 domain of Factor VIII compared to the corresponding factor VIII not linked to XTEN or to native FVIII. In another embodiment, the invention provides CFXTEN exhibiting procoagulant activity and reduced binding in the presence of antibodies binding to the A2 and the C2 domain of Factor VIII, compared to the corresponding factor VIII not linked to XTEN or to native FVIII. In one embodiment, the invention provides CFXTEN exhibiting procoagulant activity and reduced binding, compared to the corresponding FVIII not linked to XTEN, in the presence of an antibody selected from the group consisting of the antibodies of Table 10. In one embodiment, the CFXTEN fusion protein exhibits reduced binding to the antibody GMA8021. In another embodiment, the CFXTEN fusion protein exhibits reduced binding to the antibody GMA8008. In another embodiment, the CFXTEN fusion protein exhibits reduced binding to the antibody ESH4. In another embodiment, the CFXTEN fusion protein exhibits reduced binding to the antibody ESH8. In another embodiment, the CFXTEN fusion protein exhibits reduced binding to the antibody B02C11. In another embodiment, the CFXTEN fusion protein exhibits reduced binding and a greater degree of procoagulant activity, compared to the corresponding FVIII not linked to XTEN, in the presence of plasma from a hemophilia A subject with polyclonal antibody FVIII inhibitors, wherein the greater degree of procoagulant activity is determined by an in vitro assay such as a Bethesda assay or other assay described herein.

[00251] The CFXTEN exhibiting reduced binding by FVIII inhibitors can have one, or two, or three, or four, or five, or six or more individual XTEN, embodiments of which are disclosed herein. In the foregoing embodiments of this paragraph, a CFXTEN exhibits at least 5%, or 10%, or 15%, or 20%, or 30%, or 40%, or 50%, or 60%, or 70% or less binding to the antibody when assessed *in vitro* in an assay capable of assaying the binding of an antibody to FVIII, such as assays described herein below or those known in the art. Alternatively, the reduced binding of the subject CFXTEN to the FVIII-binding antibodies can be assessed by retention of a higher degree of procoagulant activity in the presence of the

antibody compared to FVIII not linked to XTEN, as described in the Examples. Thus, in the embodiments pertaining to reduced binding by FVIII inhibitors described herein, a CFXTEN exhibits, when reacted with the anti-FVIII antibody, at least 5%, or 10%, or 15%, or 20%, or 30%, or 40%, or 50%, or 60%, or 70%, or 80%, or 100%, or 200%, or 300%, or 400%, or 500% or more activity in a coagulation assay (such as described herein below) compared to the corresponding FVIII not linked to XTEN and reacted with the antibody. In the foregoing, the anti-FVIII antibody can be an antibody from Table 9 or a circulating anti-FVIII antibody from a hemophilia A subject. In another embodiment, the invention provides CFXTEN in which the assayed fusion protein, when assayed utilizing the Bethesda assay and an anti-FVIII antibody selected from Table 10 or a polyclonal anti-FVIII antibody preparation such as, but not limited to, plasma from a hemophilia A subject with FVIII inhibitors, results in a Bethesda titer with at least about 2, 4, 6, 8, 10, 12, 15, 20, 30, 40, 50, 60, 70, 80, 100, or 200 fewer Bethesda units compared to a FVIII not linked to XTEN and assayed under comparable conditions. In another embodiment, the invention provides CFXTEN in which the assayed fusion protein results in less than 50%, or less than 40%, or less than 30%, or less than 25%, or less than 20%, or less than 15%, or less than 14%, or less than 13%, or less than 12%, or less than 11%, or less than 10% of the Bethesda Units compared to a FVIII not linked to XTEN when assayed under comparable conditions utilizing the Bethesda assay and a polyclonal anti-FVIII antibody preparation such as, but not limited to, plasma from a hemophilia A subject with FVIII inhibitors.

Table 10: Anti-factor VIII antibodies

Antibody Designation	Epitope	Inhibitor Titer BU/mg	Reference
BO2C11	C2 Domain Met2199/Phe2200	20000	U.S. 6,770,744 Blood (2007) 110:4234–4242
NMC VIII-5	C2 Domain Glu2181-Val2243		U.S. 6,770,744
ESH2	Light Chain		ADI
ESH4	Light Chain 2303-2332	39	U.S. 6,770,744 Blood (2007) 110:4234–4242
ESH8	C2 Domain 2248-2285	10000	U.S. 6,770,744 Blood (2007) 110:4234–4242
RHD5 (LMBP 6165CB)	C1 Domain		WO 2005/016455 US Pat. Application 20090263380
LE2E9	C1 Domain		US Pat. Application 20090263380 Blood (2000) 95:156-163
I54	C2 Domain	1300	Blood (2007) 110:4234–4242
F85	C2 Domain	6	Blood (2007) 110:4234–4242
F100	C2 Domain	5	Blood (2007) 110:4234–4242
F137	C2 Domain	6	Blood (2007) 110:4234–4242
I89	C2 Domain	1900	Blood (2007) 110:4234–4242
I117	C2 Domain	1800	Blood (2007) 110:4234–4242
I109	C2 Domain Met2199/Phe2200	1500	Blood (2007) 110:4234–4242
1B5	C2 Domain	930	Blood (2007) 110:4234–4242
3C6	C2 Domain	71	Blood (2007) 110:4234–4242
3D12	C2 Domain Phe2196	2600	Blood (2007) 110:4234–4242

Antibody Designation	Epitope	Inhibitor Titer BU/mg	Reference
D102	C2 Domain	3800	Blood (2007) 110:4234–4242
3G6	C2 Domain	25000	Blood (2007) 110:4234–4242
2–77	C2 Domain	25000	Blood (2007) 110:4234–4242
B45	C2 Domain	21000	Blood (2007) 110:4234–4242
B9	C2 Domain	31000	Blood (2007) 110:4234–4242
B11	C2 Domain	3300	Blood (2007) 110:4234–4242
B75	C2 Domain	Indeterminate	Blood (2007) 110:4234–4242
D105	C2 Domain Val2223/Lys2227	0.8	Blood (2007) 110:4234–4242
F77	C2 Domain	26000	Blood (2007) 110:4234–4242
F178	C2 Domain	18000	Blood (2007) 110:4234–4242
F67	C2 Domain	21000	Blood (2007) 110:4234–4242
G99	C2 Domain Val2223/Lys2227	15000	Blood (2007) 110:4234–4242
G86	C2 Domain	4300	Blood (2007) 110:4234–4242
I14	C2 Domain	44000	Blood (2007) 110:4234–4242
I55	C2 Domain	10000	Blood (2007) 110:4234–4242
2–117	C2 Domain	>0.4	Blood (2007) 110:4234–4242
GMA012	A2 domain 497-510; 584-593		GMA
GMA8001	A3 Domain	156	GMA
GMA8002	A1 Domain	<1	GMA
GMA8003	C2 Domain		GMA
GMA8004	A1 Domain		GMA
GMA8005	A1A3/A1 Domain		GMA
GMA8006	C2 Domain		GMA
GMA8008	C2 Domain	1047	GMA
GMA8009	A2 Domain	7923	GMA
GMA8010	LC Domain		GMA
GMA8011	C1 Domain	97	GMA
GMA8012	A1A3 Domain	204	GMA
GMA8013	A3C2 Domain	30	GMA
GMA8014	C2 Domain	7799	GMA
GMA8015	A2 Domain	17079	GMA
GMA8016	A2 Domain	<1	GMA
GMA8017	A2 Domain	334	GMA
GMA8018	LC Domain	242	GMA
GMA8019	CR-LC Domain		GMA
GMA8020	A1A3 Domain	196	GMA
GMA8021	A2 Domain	33928	GMA
4A4	A2 Domain	40000	J Thromb Haemost (2009) 7:658-664
3E6	C2 Domain	41	Blood (2007) 110:4234–4242

American Diagnostica Inc. internet site, URL located on the World Wide Web at americandiagnostica.com/html/Product_Detail.asp?idCategory=5&idSubCategory=104&idpro=ESH-8 as it existed on January 12, 2012

Green Mountain Antibodies internet site, URL located on the World Wide Web at greenmoab.com/product_details/16316/21582.html as it existed on January 12, 2012

[00252] Assays For Inhibitor and Antibody Binding

[00253] The fusion proteins of the invention may be assayed to confirm reduced binding by FVIII inhibitors using methods known in the art. The assays that can be used include, but are not limited to, competitive and non-competitive assay systems using techniques such as Western blots, radioimmunoassays, ELISA, "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, immunoradiometric assays, fluorescent immunoassays, clotting assays, factor VIII inhibitor assays to name but a few. Such assays are routine and well known in the art (see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York). Exemplary are described briefly below but are not intended by way of limitation.

[00254] The Bethesda assay and the Nijmegen modification of the Bethesda assay are factor VIII inhibitor assays well-known as methods to detect FVIII inhibitors (Kasper CK, et al. Proceedings: A more uniform measurement of factor VIII inhibitors. Thromb Diath Haemorrh. (1975) 34(2):612). However, the assays can be modified to assay binding of inhibitors to FVIII compositions using inhibitors such as polyclonal or monoclonal anti-FVIII antibodies, including the antibodies of Table 10, and methods such as described in Example 52. Briefly, the modified Bethesda assay involves mixing titered volumes of the test sample with an equal volume of an inhibitor at a set concentration. The mixtures are incubated for 2 hours at 37°C prior to analysis of the factor concentration by a coagulation assay such as a chromogenic assay. Similarly, a reference plasma with native factor VIII level is incubated that then assayed as the positive control. The endpoint is the titer resulting in 50% of the FVIII activity of the positive control, reported as Bethesda units. In the Nijimegen modification of the Bethesda assay, the assay samples are stabilized with imidazole buffer and the control sample is mixed with deficient plasma instead of buffer (Verbruggen B, et al. The Nijmegen modification of the Bethesda assay for factor VIII:C inhibitors: improved specificity and reliability. Thromb Haemost. (1995) 73(2):247-251).

[00255] Western blot analysis generally comprises preparing protein samples, electrophoresis of the protein samples in a polyacrylamide gel (e.g., 8%–20% SDS-PAGE depending on the molecular weight of the antigen), transferring the protein sample from the polyacrylamide gel to a membrane such as nitrocellulose, PVDF or nylon, blocking the membrane in blocking solution (e.g., PBS with 3% BSA or non-fat milk), washing the membrane in washing buffer (e.g., PBS-Tween™ 20), blocking the membrane with primary antibody (the antibody of interest) diluted in blocking buffer, washing the membrane in washing buffer, blocking the membrane with a secondary antibody (which recognizes the primary antibody, e.g., an anti-human antibody) conjugated to an enzymatic substrate (e.g., horseradish peroxidase or alkaline phosphatase) or radioactive molecule (e.g., 32 P or 125 I) diluted in blocking buffer, washing the membrane in wash buffer, and detecting the presence of the antigen. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the signal detected and to reduce the background noise. For further discussion regarding western blot protocols see, e.g.,

Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 10.8.1.

[00256] ELISA assays can detect antibodies to FVIII independent of their ability to block the procoagulant activity of FVIII, and have been utilized for the detection of anti-FVIII developing in hemophilia A patients. In a population of 131 patients with hemophilia A with inhibitors, the ELISA technique resulted in 97.7% sensitivity and 78.8% specificity, and had a high negative predictive value (98.6%) [Martin, P. G., et al. Evaluation of a novel ELISA screening test for detection of factor VIII inhibitory antibodies in haemophiliacs. Clin Lab Haematol (1999) 21:125-128]. Other investigators have found a highly significant correlation between the Bethesda titer and the absorbance values in an ELISA assay for detecting anti-FVIII Abs (Towfighi, F., et al. Comparative measurement of anti-factor VIII antibody by Bethesda assay and ELISA reveals restricted isotype profile and epitope specificity. Acta Haematol (2005) 114:84-90), with the added advantage of the ability to detect non-inhibitory anti-FVIII antibodies. Assay protocols comprise preparing the binding ligand, which may include a sample comprising either factor VIII polypeptide or the CFXTEN fusion protein, coating the well of a 96 well microtiter plate with the antibody, adding the ligand test sample and incubating, then adding a detection antibody and incubating prior to washing and adding a alkaline phosphatase- or peroxidase-conjugated secondary antibody and incubating for an additional period before the addition of TMB substrate and processing for reading by spectrophotometer at 450nm. In ELISAs the antibody or inhibitor of interest does not have to be conjugated to a detectable compound; instead, a second antibody (which recognizes the antibody or inhibitor of interest) conjugated to a detectable compound may be added to the well. Further, instead of coating the well with the antibody, the ligand may be coated to the well. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the signal detected as well as other variations of ELISAs known in the art (*see*, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 11.2.1).

[00257] Standard or modified coagulation assays are used to measure reduced binding of FVIII binding agents. In one exemplary method (further described in Example 28), the optimal concentration of a given FVIII inhibitor to utilize in the assay is first determined by a titration experiment using varying amounts of the inhibitory antibody incubated at 37°C for 2 hrs with the base vector expressing wild-type FVIII containing a His/Myc double tag. The FVIII activity is measured by the Coatest assay procedure described herein. The lowest concentration that results in optimal inhibition of FVIII activity is employed in the assay. In the assay, the FVIII inhibitor antibody at the optimal concentration is mixed with individual test samples and incubated at 37°C for 2 hrs. The resulting test samples are then collected and utilized in the Coatest activity assay, along with untreated aliquots of the CFXTEN and positive control in order to assess the residual and baseline FVIII activity for each test sample.

[00258] The invention provides methods of making CFXTEN that exhibit reduced binding to FVIII binding agents, including FVIII inhibitors, and retention of procoagulant activity. In one embodiment, the method to make a CFXTEN with reduced binding to FVIII inhibitors comprises the steps of selecting a FVIII sequence with at least 90% sequence identity to a sequence of Table 1, selecting one, two, three,

four, five, or six or more XTEN each with at least 70%, or at least 80%, or at least 90%, or at least 95-99% sequence identity to XTEN sequences of comparable length from Table 4, creating expression constructs designed to locate said XTEN at or proximal to locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9, expressing and recovering the resulting CFXTEN, and assaying the resulting fusion proteins in an assay described herein in order to confirm the reduced binding of the CFXTEN fusion protein. By the inventive method, a CFXTEN exhibits at least 5% reduced, or at least 10% reduced, or at least 15% reduced, or at least 20% reduced, or at least 25% reduced, or at least 40% reduced, or at least 50% reduced, or at least 60% reduced, or at least 70% reduced, or at least 80% reduced binding to a FVIII binding agent including, but not limited to the antibodies of Table 10 or anti-FVIII antibodies from a hemophilia A subject, and retains at least about 10%, or at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70% procoagulant activity compared to the corresponding FVIII not linked to XTEN.

[00259] Up to 8-10% of hemophilia A patients have antibodies that bind FVIII without affecting its procoagulant properties; they are not, therefore categorized as FVIII inhibitors. However, the binding of antibodies to FVIII is believed to lead to immune complexes that are cleared by the innate immune response or are more susceptible to proteolytic degradation (Kazatchkine MD. Circulating immune complexes containing anti-VIII antibodies in multi-transfused patients with haemophilia A. Clin Exp Immunol. (1980) 39(2):315 320). Accordingly, it is an object of the invention to provide CFXTEN fusion proteins comprising one or more XTEN that exhibit reduced binding of antibodies to FVIII that are not inhibitors, wherein the degradation or clearance of the CFXTEN is reduced at least 5%, or 10%, or 15%, or 20%, or 30%, or 40%, or 50%, or 60%, or 70% or less compared to a corresponding FVIII not linked to XTEN or to native FVIII bound by such antibodies. The reduced binding of antibodies to CFXTEN compared to FVIII not linked to XTEN or to native FVIII can be assayed by *in vitro* and *in vivo* methods. In vitro methods include the aforementioned ELISA and Western blot methods. The reduced degradation or clearance of CFXTEN can be assessed *in vivo* by use of animal models or in human clinical trials. In one type of trial, factor VIII or CFXTEN are administered separately, preferably by intravenous infusion, to cohorts of patients having factor VIII deficiency who have antibodies that promote degradation or clearance of therapeutic human factor VIII. The dosage of the administered test article is in a range between 5 and 50 IU/kg body weight, preferably 10-45 IU/kg, and most preferably 40 IU/kg body weight. Approximately 1 hour after each administration, the recovery of factor VIII or CFXTEN from blood samples is measured in a functional one-stage or chromogenic coagulation assay to assess activity and by ELISA, HPLC, or similar assay to qualify the amount of intact factor VIII equivalent. Samples are taken again approximately 5-10 hours after infusion, and recovery is measured. Total recovery and the rate of disappearance of factor VIII from the samples is predictive of the antibody titer, and the comparison of results from the factor VIII and CFXTEN indicates the degree of reduced clearance and/or degradation of the CFXTEN. In one embodiment, the CFXTEN fusion protein exhibits at least 5% reduced, or at least 10% reduced, or at least 15% reduced, or at least 20% reduced, or at least 25% reduced, or at least 40% reduced, or at least 50% reduced, or at least 60%

reduced, or at least 70% reduced, or at least 80% reduced binding to an anti-FVIII antibody that promotes clearance but does not otherwise inhibit the procoagulant activity of intact native FVIII. In another embodiment, the CFXTEN fusion protein exhibits at least 5% reduced, or at least 10% reduced, or at least 15% reduced, or at least 20% reduced, or at least 25% reduced, or at least 40% reduced, or at least 50% reduced, or at least 60% reduced, or at least 70% reduced, or at least 80% reduced binding to an anti-FVIII antibody that promotes the degradation of FVIII. In the foregoing embodiments of this paragraph, the reduced binding of the anti-FVIII antibody is alternatively characterized by an increased K_D value of the FVIII antibody to the fusion protein compared to the FVIII of at least two-fold, or three-fold, or four-fold, or five-fold, or 10-fold, or 33-fold, or 100-fold, or 330-fold, or at least 1000-fold compared to the binding to the corresponding FVIII not linked to XTEN. In one embodiment, the CFXTEN fusion proteins comprising one or more XTEN exhibiting reduced reactivity to an anti-FVIII antibody exhibits an increased terminal half-life when administered to a subject with anti-FVIII antibodies of at least 48 h, or at least 72 h, or at least 96 h, or at least 120 h, or at least 144 h, or at least 14 days, or at least 21 days compared to FVIII not linked to XTEN. In the foregoing embodiment, the subject can be a human hemophilia A subject or it can be a mouse hemophilia A subject with circulating anti-FVIII antibodies.

[00260] Another aspect of the present invention is the use of CFXTEN fusion protein for a specific therapy of a coagulopathy in a subject with a FVIII inhibitor. The invention provides a method of treating a subject with circulating FVIII inhibitor(s) comprising the step of administering a clotting-effective amount of a CFXTEN fusion protein to the subject wherein the fusion protein exhibits greater procoagulant activity and/or clotting-effective concentrations of longer duration compared to either a corresponding factor VIII not linked to XTEN or compared to native factor VIII administered to the subject using a comparable amount and route of administration. In one embodiment of the method, the FVIII inhibitor in the subject is an anti-FVIII antibody. In another embodiment, the FVIII inhibitor is a neutralizing anti-FVIII antibody. In one embodiment, the FVIII inhibitor is an anti-FVIII antibody that binds to the A1 domain of FVIII. In another embodiment, the FVIII inhibitor is an anti-FVIII antibody that binds to the A2 domain of FVIII. In another embodiment, the FVIII inhibitor is an anti-FVIII antibody that binds to the A3 domain of FVIII. In another embodiment, the FVIII inhibitor is an anti-FVIII antibody that binds to the C1 domain of FVIII. In another embodiment, the FVIII inhibitor is an anti-FVIII antibody that binds to the C2 domain of FVIII. In another embodiment, the FVIII inhibitor is an anti-FVIII antibody that binds to both the C2 and A2 domain of FVIII. In another embodiment, the FVIII inhibitor binds to a FVIII epitope capable of being bound by one or more antibodies of Table 10. In another embodiment, the FVIII inhibitor is a polyclonal antibody from a hemophilia A subject with FVIII inhibitor antibodies.

[00261] An object of the present invention is the creation of CFXTEN with XTEN inserted to maximize the steric interference of FVIII binding agents that would otherwise bind to FVIII and neutralize procoagulant activity or result in the clearance or degradation of FVIII. Accordingly, in one approach the invention provides CFXTEN comprising one or more XTEN wherein the XTEN are inserted

proximal to a binding site of a FVIII inhibitor or anti-FVIII antibody. In one embodiment, an XTEN is linked to the FVIII at a location selected from Table 5, Table 6, Table 7, Table 8, and Table 9 that is within about 50, or about 100, or about 150, or about 200, or about 250, or about 300 amino acids of a FVIII epitope that is bound by an antibody of Table 10. In another embodiment, the XTEN is linked to the FVIII within about 50, or about 100, or about 150, or about 200, or about 250, or about 300 amino acids of a FVIII epitope in the A2 or C2 domain that is bound by an antibody of Table 10. Accordingly, the invention provides CFXTEN fusion proteins comprising one or more XTEN wherein binding by FVIII inhibitors to the FVIII component of the fusion protein is reduced compared to the corresponding FVIII not linked to XTEN or to native FVIII and the CFXTEN retains procoagulant activity. In the foregoing embodiments hereinabove described in this paragraph, the fusion proteins can be assayed by the assays described herein below, the assays of the Examples, or other assays known in the art, and the inhibitors can be an antibody of Table 10, can be polyclonal anti-FVIII, or can be blood or plasma from a hemophilia A subject with FVIII inhibitors.

[00262] In another aspect, CFXTEN are designed to maximize the regions over which XTEN can adopt random coil conformations covering the fusion protein, thereby resulting in steric hindrance for anti-FVIII antibodies that would otherwise bind epitopes on the FVIII component of the fusion protein. It is believed that the incorporation of multiple XTEN into a CFXTEN provides a higher total hydrodynamic radius of the XTEN component compared to CFXTEN with fewer XTEN yet having approximately the same total of XTEN amino acids. Empirically, the hydrodynamic radius for a protein can be calculated based on size exclusion chromatography, and results of several fusion proteins using such methods are described in the Examples. Alternatively, the radius for XTEN polypeptides, such as those incorporated in the embodiments disclosed herein, can be approximated by mathematical formulae because the limited types of amino acids utilized have known characteristics that can be quantified. In one embodiment, the maximum radius of a single XTEN polypeptide is calculated (hereinafter "XTEN Radius") according to the formula given by Equation II:

$$\text{XTEN Radius} = (\sqrt{\text{XTEN length}}\ 0.2037) + 3.4627 \qquad \text{II}$$

[00263] In another embodiment, the sum of the maximum of the XTEN Radii for all XTEN segments in a CFXTEN is calculated (hereinafter "Sum XTEN Radii") according to the formula given by Equation III:

$$\text{Sum XTEN Radii} = \sum_{i=1}^{n} XTEN\ Radius_i \qquad \text{III}$$

wherein: n = the number of XTEN segments

and i is an iterator

[00264] In another embodiment, the ratio of the SUM XTEN Radii of a CFXTEN comprising multiple XTEN to that of an XTEN Radius for a single XTEN of an equivalent length (in total amino acid residues to that of the CFXTEN) is calculated (hereinafter "Ratio XTEN Radii") according to the formula given by Equation IV:

$$\text{Ratio XTEN Radii} = \frac{\sum_{i=1}^{n} XTEN\ Radius_i}{\left(\sqrt{\sum_{i=1}^{n} XTEN\ Length_i} * 0.2037\right) + 3.4627} \quad \text{IV}$$

wherein: n = the number of XTEN segments

and i is an iterator

[00265] In applying the Equations to the XTEN, it will be understood by one of skill in the art that the calculated values represent maximum values that could vary or be reduced depending on the host cell utilized for expression of the XTEN polypeptide. It is believed that while *E. coli* expression would result in XTEN that achieves the calculated values, expression in eukaryotic host cells in which XTEN may be glycosylated could result in a radius of the polypeptide less than the maximum calculated value. Such differences can be quantified by methods such as size exclusion chromatography, the methods of which are detailed in the Examples.

[00266] In order to design CFTEN that maximize the area over which XTEN can adopt random coil conformations, it was discovered that CFXTEN designs with Ratio XTEN Radii above 2 provide greater coverage over the fusion protein than designs with values <2. Accordingly, in one embodiment the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0, or 2.1, or 2.2, or 2.3, or 2.4, or 2.5, or 2.6, or 2.7, or 2.8, or 2.9, or 3.0, or 3.1, or 3.2, or 3.3, or 3.4, or 3.5 or greater. In some embodiments, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater comprise at least three XTEN with each XTEN having at least 42 to about 288 amino acids and wherein at least two of the XTEN are linked to the fusion protein with no less than about 100, or about 200, or about 300, or about 400, or about 500 amino acids of separation between the two XTEN. In other embodiments, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater comprise at least four XTEN with each XTEN having at least 42 to about 288 amino acids and wherein at least three of the XTEN are linked to the fusion protein with no less than about 100, or about 200, or about 300, or about 400 amino acids of separation between any two of the three XTEN.

[00267] In another embodiment, the invention provides a CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater, the CFXTEN comprises at least three XTEN with each XTEN having at least 42 to about 288 amino acids and wherein at least two of the three of the XTEN linked to the fusion protein are separated by an amino acid sequence of at least 100, or about 200, or about 300 to about 400 amino acids, and the third XTEN is linked within the B domain (or fragment thereof) or within the C domain (or the terminus thereof). In another embodiment, the invention provides a CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater, the CFXTEN comprises at least four XTEN with each XTEN having at least 42 to about 288 amino acids and wherein at least three of the four of the XTEN linked to the fusion protein are separated by an amino acid sequence of at least 300 to about 400 amino acids and the fourth XTEN is linked within the B domain (or fragment thereof) or within the C domain (or the terminus thereof).

[00268] In yet other embodiments, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater, the CFXTEN comprises at least five XTEN with four XTEN having at least 42 to about 144 amino acids wherein at least four of the XTEN are linked to the fusion protein with no less

than about 100, 200, or about 300, or about 400 amino acids of separation between any two of the four XTEN and a fifth XTEN is linked within the B domain (or fragment thereof) or within the C domain (or the terminus thereof). In one embodiment, the invention provides a CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater, the CFXTEN comprises at least five XTEN with four XTEN having at least 42 to about 144 amino acids wherein at least three of the XTEN linked to the fusion protein are separated by an amino acid sequence of at least 300 to about 400 amino acids, the fourth XTEN is linked within the B domain (or fragment thereof) and a fifth XTEN is linked within the C domain (or the terminus thereof).

[00269] In one aspect, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0, or 2.1, or 2.2, or 2.3, or 2.4, or 2.5, or 2.6, or 2.7, or 2.8, or 2.9, or 3.0, or 3.1, or 3.2, or 3.3, or 3.4, or 3.5 or greater, and the composition does not comprise certain sequences. In one embodiment of the foregoing, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater with the proviso that the fusion protein does not comprise a sequence from any one of Table 50 or Table 51. In another embodiment of the foregoing, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater with the proviso that the fusion protein does not comprise a sequence having an AG family XTEN sequence. In another embodiment of the foregoing, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater with the proviso that the fusion protein does not comprise a sequence selected from GTPGSGTASSSP (SEQ ID NO: 31), GSSTPSGATGSP (SEQ ID NO: 32), GSSPSASTGTGP (SEQ ID NO: 33), GASPGTSSTGSP (SEQ ID NO: 34). In another embodiment of the foregoing, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater with the proviso that the fusion protein does not comprise any one of the sequences selected from GTPGSGTASSSP (SEQ ID NO: 31), GSSTPSGATGSP (SEQ ID NO: 32), GSSPSASTGTGP (SEQ ID NO: 33), GASPGTSSTGSP (SEQ ID NO: 34) and GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP (SEQ ID NO: 59). In another embodiment of the foregoing, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater with the proviso that the fusion protein does not comprise a sequence selected from GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP (SEQ ID NO: 59), PGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSS (SEQ ID NO: 71), or PGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSG

TASSSPGSSTPSGATGS (SEQ ID NO: 80). In another embodiment of the foregoing, the invention provides CFXTEN in which the Ratio XTEN Radii is at least 2.0-3.5 or greater with the proviso that the fusion protein does not comprise an XTEN sequence consisting of GSEPATSGSETPGTSESATPESGPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGSEPATSGSETPGSEPATSGSETPGSEPATSGSETPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAP (SEQ ID NO: 59), PGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSS (SEQ ID NO: 71), or PGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGS (SEQ ID NO: 80).

[00270] In one aspect, the present invention provides methods to create CFXTEN with XTEN inserted to maximize the steric interference of FVIII binding agents that would otherwise bind to FVIII and neutralize procoagulant activity or result in the clearance or degradation of FVIII. Accordingly, in one embodiment, the invention provides a method comprising the steps of selecting a FVIII sequence with at least 90% sequence identity to a sequence of Table 1, selecting three or more XTEN from Table 4 in which the Ratio XTEN Radii is at least 2.0, or 2.1, or 2.2, or 2.3, or 2.4, or 2.5, or 2.6, or 2.7, or 2.8, or 2.9, or 3.0, or 3.1, or 3.2, or 3.3, or 3.4, or 3.5 or greater, creating expression constructs designed to locate said XTEN at or proximal to locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9, wherein the three or more XTEN are at least 300 to 400 amino acids, expressing and recovering the resulting CFXTEN, and assaying the resulting fusion proteins in an assay described herein in order to confirm the reduced binding of the CFXTEN fusion protein. By the inventive method, a CFXTEN exhibits at least 5% reduced, or at least 10% reduced, or at least 15% reduced, or at least 20% reduced, or at least 25% reduced, or at least 40% reduced, or at least 50% reduced, or at least 60% reduced, or at least 70% reduced, or at least 80% reduced binding to a FVIII binding agent including, but not limited to the antibodies of Table 10, and exhibits procoagulant activity.

5. CFXTEN Fusion Protein Configurations with Spacer and Cleavage Sequences

[00271] In another aspect, the invention provides CFXTEN configured with one or more spacer sequences incorporated into or adjacent to the XTEN that are designed to incorporate or enhance a functionality or property to the composition, or as an aid in the assembly or manufacture of the fusion protein compositions. Such properties include, but are not limited to, inclusion of cleavage sequence(s) to permit release of components, inclusion of amino acids compatible with nucleotide restrictions sites to permit linkage of XTEN-encoding nucleotides to FVIII-encoding nucleotides or that facilitate construction of expression vectors, and linkers designed to reduce steric hindrance in regions of CFXTEN fusion proteins.

[00272] In an embodiment, a spacer sequence can be introduced between an XTEN sequence and a FVIII component to decrease steric hindrance such that the FVIII component may assume its desired tertiary structure and/or interact appropriately with its target substrate or processing enzyme. For spacers and methods of identifying desirable spacers, see, for example, George, *et al.* (2003) Protein Engineering 15:871–879. In one embodiment, the spacer comprises one or more peptide sequences that are between 1–50 amino acid residues in length, or about 1–25 residues, or about 1-10 residues in length. Spacer sequences, exclusive of cleavage sites, can comprise any of the 20 natural L amino acids, and will preferably have XTEN-like properties in that the majority of residues will be hydrophilic amino acids that are sterically unhindered such as, but not limited to, glycine (G), alanine (A), serine (S), threonine (T), glutamate (E), proline (P) and aspartate (D). The spacer can be a single glycine residue, polyglycines or polyalanines, or is predominately a mixture of combinations of glycine, serine and alanine residues. In one embodiment, a spacer sequence, exclusive of cleavage site amino acids, has about 1 to 10 amino acids that consist of amino acids selected from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E), and proline (P) and are substantially devoid of secondary structure; e.g., less than about 10%, or less than about 5% as determined by the Chou-Fasman and/or GOR algorithms. In one embodiment, the spacer sequence is GPEGPS (SEQ ID NO: 1612). In another embodiment, the spacer sequence is GPEGPS (SEQ ID NO: 1612) linked to a cleavage sequence of Table 12. In addition, spacer sequences are designed to avoid the introduction of T-cell epitopes which can, in part, be achieved by avoiding or limiting the number of hydrophobic amino acids utilized in the spacer; the determination of epitopes is described above and in the Examples.

[00273] In a particular embodiment, the CFXTEN fusion protein comprises one or more spacer sequences linked at the junction(s) between the payload FVIII sequence and the one or more XTEN incorporated into the fusion protein, wherein the spacer sequences comprise amino acids that are compatible with nucleotides encoding restriction sites. In another embodiment, the CFXTEN fusion protein comprises one or more spacer sequences linked at the junction(s) between the payload FVIII sequence and the one more XTEN incorporated into the fusion protein wherein the spacer sequences comprise amino acids that are compatible with nucleotides encoding restriction sites and the amino acids and the one more spacer sequence amino acids are chosen from glycine (G), alanine (A), serine (S), threonine (T), glutamate (E), and proline (P). In another embodiment, the CFXTEN fusion protein comprises one or more spacer sequences linked at the junction(s) between the payload FVIII sequence and one more XTEN incorporated into the fusion protein wherein the spacer sequences comprise amino acids that are compatible with nucleotides encoding restriction sites and the one more spacer sequences are chosen from the sequences of Table 11. The exact sequence of each spacer sequence is chosen to be compatible with cloning sites in expression vectors that are used for a particular CFXTEN construct. In one embodiment, the spacer sequence has properties compatible with XTEN. In one embodiment, the spacer sequence is GAGSPGAETA (SEQ ID NO: 178). For XTEN sequences that are incorporated internal to the FVIII sequence, each XTEN would generally be flanked by two spacer sequences comprising amino acids compatible with restriction sites, while XTEN attached to the N- or C-terminus

would only require a single spacer sequence at the junction of the two components and another at the opposite end for incorporation into the vector. As would be apparent to one of ordinary skill in the art, the spacer sequences comprising amino acids compatible with restriction sites that are internal to FVIII could be omitted from the construct when an entire CFXTEN gene is synthetically generated.

Table 11: Spacer Sequences Compatible with Restriction Sites

Spacer Sequence	Restriction Enzyme
GSPG (SEQ ID NO: 174)	BsaI
ETET (SEQ ID NO: 175)	BsaI
PGSSS (SEQ ID NO: 176)	BbsI
GAP	AscI
GPA	FseI
GPSGP (SEQ ID NO: 177)	SfiI
AAA	SacII
TG	AgeI
GT	KpnI
GAGSPGAETA (SEQ ID NO: 178)	SfiI
ASS	XhoI

[00274] In another aspect, the present invention provides CFXTEN configurations with cleavage sequences incorporated into the spacer sequences. In some embodiments, spacer sequences in a CFXTEN fusion protein composition comprise one or more cleavage sequences, which are identical or different, wherein the cleavage sequence may be acted on by a protease, as shown in FIG. 12, to release FVIII, a FVIII component (e.g., the B domain) or XTEN sequence(s) from the fusion protein. In one embodiment, the incorporation of the cleavage sequence into the CFXTEN is designed to permit release of the FVIII component that becomes active or more active (with respect to its ability serve as a membrane binding site for factors IXa and X) upon its release from the XTEN. In the foregoing embodiment, the procoagulant activity of FVIII component of the CFXTEN is increased after cleavage by at least 30%, or at least 40%, or at least 50%, or at least 60%, or at least 70%, or at least 80%, or at least 90% compared to the intact CFXTEN. The cleavage sequences are located sufficiently close to the FVIII sequences, generally within 18, or within 12, or within 6, or within 2 amino acids of the FVIII sequence, such that any remaining residues attached to the FVIII after cleavage do not appreciably interfere with the activity (e.g., such as binding to a clotting protein) of the FVIII, yet provide sufficient access to the protease to be able to effect cleavage of the cleavage sequence. In some cases, the CFXTEN comprising the cleavage sequences will also have one or more spacer sequence amino acids between the FVIII and the cleavage sequence or the XTEN and the cleavage sequence to facilitate access of the protease; the spacer amino acids comprising any natural amino acid, including glycine, serine and alanine as preferred amino acids. In one embodiment, the cleavage site is a sequence that can be cleaved by a protease endogenous to the mammalian subject such that the CFXTEN can be cleaved after administration to a subject. In such case, the CFXTEN can serve as a prodrug or a circulating depot for

the FVIII. In a particular construct of the foregoing, the CFXTEN would have one or two XTEN linked to the N- and/or the C-terminus of a FVIII-BDD via a cleavage sequence that can be acted upon by an activated coagulation factor, and would have an additional XTEN located between the processing amino acids at position R740 and R1689 such that the XTEN could be released, leaving a form of FVIII similar to native activated FVIII. In one embodiment of the foregoing construct, the FVIII that is released from the fusion protein by cleavage of the cleavage sequence exhibits at least about a two-fold, or at least about a three-fold, or at least about a four-fold, or at least about a five-fold, or at least about a six-fold, or at least about a eight-fold, or at least about a ten-fold, or at least about a 20-fold increase in activity compared to the intact CFXTEN fusion protein.

[00275] Examples of cleavage sites contemplated by the invention include, but are not limited to, a polypeptide sequence cleavable by a mammalian endogenous protease selected from FXIa, FXIIa, kallikrein, FVIIIa, FVIIIa, FXa, FIIa (thrombin), Elastase-2, granzyme B, MMP-12, MMP-13, MMP-17 or MMP-20, or by non-mammalian proteases such as TEV, enterokinase, PreScission™ protease (rhinovirus 3C protease), and sortase A. Sequences known to be cleaved by the foregoing proteases and others are known in the art. Exemplary cleavage sequences contemplated by the invention and the respective cut sites within the sequences are presented in Table 12, as well as sequence variants thereof. For CFXTEN comprising incorporated cleavage sequence(s), it is generally preferred that the one or more cleavage sequences are substrates for activated clotting proteins. For example, thrombin (activated clotting factor II) acts on the sequence LTPRSLLV (SEQ ID NO: 1618) [Rawlings N.D., et al. (2008) *Nucleic Acids Res.*, **36**: D320], which is cut after the arginine at position 4 in the sequence. Active FIIa is produced by cleavage of FII by FXa in the presence of phospholipids and calcium and is down stream from factor VIII in the coagulation pathway. Once activated, its natural role in coagulation is to cleave fibrinogen, which then in turn, begins clot formation. FIIa activity is tightly controlled and only occurs when coagulation is necessary for proper hemostasis. By incorporation of the LTPRSLLV sequence (SEQ ID NO: 1618) into the CFXTEN between and linking the FVIII and the XTEN components, the XTEN is removed from the adjoining FVIII concurrent with activation of either the extrinsic or intrinsic coagulation pathways when coagulation is required physiologically, thereby selectively releasing FVIII. In another embodiment, the invention provides CFXTEN with incorporated FXIa cleavage sequences between the FVIII and XTEN component(s) that are acted upon only by initiation of the intrinsic coagulation system, wherein a procoagulant form of FVIII is released from XTEN by FXIa to participate in the coagulation cascade. While not intending to be bound by any particular theory, it is believed that the CFXTEN of the foregoing embodiment would sequester the FVIII away from the other coagulation factors except at the site of active clotting, thus allowing for larger doses (and therefore longer dosing intervals) with minimal safety concerns.

[00276] Thus, cleavage sequences, particularly those susceptible to the procoagulant activated clotting proteins listed in Table 12, would provide for sustained release of FVIII that, in certain embodiments of the CFXTEN, can provide a higher degree of activity for the FVIII component released from the intact form of the CFXTEN, as well as additional safety margin for high doses of CFXTEN administered to a

subject. In one embodiment, the invention provides CFXTEN comprising one or more cleavage sequences operably positioned to release the FVIII from the fusion protein upon cleavage, wherein the one or more cleavage sequences has at least about 86%, or at least about 92%, or 100% sequence identity to a sequence selected from Table 12.

[00277] In some embodiments, only the two or three amino acids flanking both sides of the cut site (four to six amino acids total) are incorporated into the cleavage sequence that, in turn, is incorporated into the CFXTEN of the embodiments, providing, e.g., XTEN release sites. In other embodiments, the incorporated cleavage sequence of Table 12 can have one or more deletions or insertions or one or two or three amino acid substitutions for any one or two or three amino acids in the known sequence, wherein the deletions, insertions or substitutions result in reduced or enhanced susceptibility but not an absence of susceptibility to the protease, resulting in an ability to tailor the rate of release of the FVIII from the XTEN. Exemplary substitutions within cleavage sequences that are utilized in the CFXTEN of the invention are shown in Table 12.

Table 12: Protease Cleavage Sequences

Protease Acting Upon Sequence	Exemplary Cleavage Sequence	SEQ ID NO:	Minimal Cut Site	SEQ ID NO:
FXIa	KLTR↓AET	179	KD/FL/T/R↓VA/VE/GT/GV	
FXIa	DFTR↓VVG	180	KD/FL/T/R↓VA/VE/GT/GV	
FXIIa	TMTR↓IVGG	181	NA	
Kallikrein	SPFR↓STGG	182	-/-/FL/RY↓SR/RT/-/-	
FVIIa	LQVR↓IVGG	183	NA	
FIXa	PLGR↓IVGG	184	-/-/G/R↓-/-/-/-	
FXa	IEGR↓TVGG	185	IA/E/GFP/R↓STI/VFS/-/G	
FIIa (thrombin)	LTPR↓SLLV	186	-/-/PLA/R↓SAG/-/-/-	
Elastase-2	LGPV↓SGVP	187	-/-/-/VIAT↓-/-/-/-	
Granzyme-B	VAGD↓SLEE	188	V/-/-/D↓-/-/-/-	
MMP-12	GPAG↓LGGA	189	G/PA/-/G↓L/-/G/-	190
MMP-13	GPAG↓LRGA	191	G/P/-/G↓L/-/GA/-	192
MMP-17	APLG↓LRLR	193	-/PS/-/-↓LQ/-/LT/-	
MMP-20	PALP↓LVAQ	194	NA	
TEV	ENLYFQ↓G	195	ENLYFQ↓G/S	196
Enterokinase	DDDK↓IVGG	197	DDDK↓IVGG	198
Protease 3C (PreScission™)	LEVLFQ↓GP	199	LEVLFQ↓GP	200
Sortase A	LPKT↓GSES	201	L/P/KEAD/T↓G/-/EKS/S	202

↓indicates cleavage site NA: not applicable

the listing of multiple amino acids before, between, or after a slash indicate alternative amino acids that can be substituted at the position; "-" indicates that any amino acid may be substituted for the corresponding amino acid indicated in the middle column

6. Exemplary CFXTEN Fusion Protein Sequences

[00278] Non-limiting examples of sequences of fusion proteins containing a single FVIII linked to one or more XTEN are presented in Table 21. The exemplary amino acid sequences of Table 21 (and the DNA sequences that encode them) contain his tags for purification purposes that, as would be apparent to one of skill in the art, can be deleted from the sequence without having an effect on the procoagulant activity of the CFXTEN fusion protein. In one embodiment, the CFXTEN of Table 21 further comprise amino acids on the N-terminus corresponding to that of native human FVIII (namely, the sequence MQIELSTCFFLCLLRFCFS (SEQ ID NO: 1611)) to aid in the expression and secretion of the CFXTEN fusion protein. In one embodiment, a CFXTEN composition comprises a fusion protein having at least about 80% sequence identity compared to a CFXTEN from Table 21, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or about 100% sequence identity as compared to a CFXTEN from Table 21, when optimally aligned. In another embodiment, a CFXTEN composition comprises a fusion protein from Table 21 in which the C-terminal his-his-his-his-his-his sequence (SEQ ID NO: 1700) deleted. However, the invention also contemplates substitution of any of the FVIII sequences of Table 1 for a FVIII component of the CFXTEN of Table 21, and/or substitution of any sequence of any one of Tables 3, 4, and 13-17 for an XTEN component of the CFXTEN of Table 21. Generally, the resulting CFXTEN of the foregoing examples retain at least a portion of the procoagulant activity of the corresponding FVIII not linked to the XTEN. In the foregoing fusion proteins hereinabove described in this paragraph, the CFXTEN fusion protein can further comprise one or more cleavage sequences; e.g., a sequence from Table 12, the cleavage sequence being located between the FVIII and the XTEN sequences or between adjacent FVIII domains linked by XTEN. In some embodiments comprising cleavage sequence(s), the intact CFXTEN composition has less activity but a longer half-life in its intact form compared to a corresponding FVIII not linked to the XTEN, but is designed such that upon administration to a subject, the FVIII component is gradually released from the fusion protein by cleavage at the cleavage sequence(s) by endogenous proteases, whereupon the FVIII component exhibits procoagulant activity.

[00279] The CFXTEN compositions of the embodiments can be evaluated for activity using assays or *in vivo* parameters as described herein (e.g., *in vitro* coagulation assays, assays of Table 49, or a pharmacodynamic effect in a preclinical hemophilia model or in clinical trials in humans, using methods as described in the Examples or other methods known in the art for assessing FVIII activity) to determine the suitability of the configuration or the FVIII sequence variant, and those CFXTEN compositions (including after cleavage of any incorporated XTEN-releasing cleavage sites) that retain at least about 30%, or about 40%, or about 50%, or about 55%, or about 60%, or about 70%, or about 80%, or about 90%, or about 95% or more activity compared to native FVIII sequence are considered suitable for use in the treatment of FVIII-related conditions.

V). PROPERTIES OF THE CFXTEN COMPOSITIONS OF THE INVENTION

(a) Pharmacokinetic Properties of CFXTEN

[00280] It is an object of the present invention to provide CFXTEN fusion proteins and pharmaceutical compositions comprising CFXTEN with enhanced pharmacokinetics compared to FVIII not linked to XTEN. The pharmacokinetic properties of a FVIII enhanced by linking a given XTEN to the FVIII include, but are not limited to, terminal half-life, area under the curve (AUC), C_{max}, volume of distribution, maintaining the biologically active CFXTEN above a minimum effective blood unit concentration for a longer period of time compared to the FVIII not linked to XTEN. The enhanced properties permit less frequent dosing and/or a longer-lived procoagulant effect compared to a comparable dose of FVIII not linked to XTEN. Enhancement of one or more of these properties can resulting benefits in the treatment of factor VIII-related conditions.

[00281] Exogenously administered factor VIII has been reported to have a terminal half-life in humans of approximately 12-14 hours when complexed with normal von Willebrand factor protein, whereas in the absence of von Willebrand factor, the half-life of factor VIII is reduced to 2 hours (Tuddenham EG, et al., Br J Haematol. (1982) 52(2):259-267; Bjorkman, S., et al. Clin Pharmacokinet. (2001) 40:815). As a result of the enhanced properties conferred by XTEN, the CFXTEN, when used at the dose and dose regimen determined to be appropriate for the subject and its underlying condition, can achieve a circulating concentration resulting in a desired procoagulant or clinical effect for an extended period of time compared to a comparable dose of the corresponding FVIII not linked to XTEN. As used herein, a "comparable dose" means a dose with an equivalent moles/kg or International Units/kg (IU/kg) for the composition that is administered to a subject. It will be understood in the art that a "comparable dose" of FVIII not linked to XTEN would represent a lesser weight of drug but would have essentially the same IUs or mole-equivalents of CFXTEN in the dose.

[00282] An international unit ("IU") of factor VIII is defined in the art as the coagulant activity present in 1 ml of normal human plasma. A normal, non-hemophilic individual human is expected to have about 100 IU/dL factor VIII activity. In hemophilia A, the doses required to treat are dependent on the condition. For minor bleeding, doses of native or recombinant factor VIII of 20 to 40 IU/kg are typically administered, as necessary. For moderate bleeding, doses of 30 to 60 IU/kg are administered as necessary, and for major bleeding, doses of 80 to 100 IU/kg may be required, with repeat doses of 20 to 25 IU/kg given every 8 to 12 hours until the bleeding is resolved. For prophylaxis against bleeding in patients with severe hemophilia A, the usual doses of native or recombinant FVIII preparations are 20 to 40 IU/kg body weight at intervals of about 2 to 3 days. A standard equation for estimating an appropriate dose of a composition comprising FVIII is:

Required units = body weight (kg) x desired factor VIII rise (IU/dL or % of normal) x 0.5 (IU/kg per IU/dL).

[00283] In many cases, the therapeutic levels for FVIII in subjects of different ages or degree of disease have been established and are available in published literature or are stated on the drug label for approved products containing the FVIII. For example, the Subcommittee on Factor VIII and Factor IX of the

Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis posted, on the ISTH Website 29 November, 2000, that the most widely used measure of hemophilia A is established by determining the circulating concentrations of plasma FVIII procoagulant levels, with persons with <1% (< 0.01 IU/ml) factor VIII defined as severe; 1-5% (0.01 - 0.05 IU/ml) as moderately severe; and >5-40% (0.05 - <0.40 IU/ml) as mild, where normal is 1 IU/ml of factor VIIIC (100%). The therapeutic levels can be established for new compositions, including those CFXTEN and pharmaceutical compositions comprising CFXTEN of the disclosure, using standard methods. In practicing the present invention, it will be understood that any dosage of CFXTEN that is effective may be used for treating bleeding episodes or maintaining hemostasis. The methods for establishing the therapeutic levels and dosing schedules for a given composition are known to those of skill in the art (*see*, e.g., Goodman & Gilman's The Pharmacological Basis of Therapeutics, 11th Edition, McGraw-Hill (2005)). For example, by using dose-escalation studies in subjects with the target condition to determine efficacy or a desirable pharmacologic effect, appearance of adverse events, and determination of circulating blood levels, the therapeutic blood levels for a given subject or population of subjects can be determined for a given drug or biologic. The dose escalation studies would evaluate the activity of a CFXTEN through studies in a subject or group of hemophilia A subjects. The studies would monitor blood levels of procoagulant, as well as physiological or clinical parameters as known in the art or as described herein for one or more parameters associated with the factor VIII related condition, or clinical parameters associated with a beneficial outcome, together with observations and/or measured parameters to determine the no effect dose, adverse events, minimum effective dose and the like, together with measurement of pharmacokinetic parameters that establish the determined or derived circulating blood levels. The results can then be correlated with the dose administered and the blood concentrations of the therapeutic that are coincident with the foregoing determined parameters or effect levels. By these methods, a range of doses and blood concentrations can be correlated to the minimum effective dose as well as the maximum dose and blood concentration at which a desired effect occurs or is maintained and the period for which it can be maintained, thereby establishing the therapeutic blood levels and dosing schedule for the composition. Thus, by the foregoing methods, a C_{min} blood level is established, below which the CFXTEN fusion protein would not have the desired pharmacologic effect and a C_{max} blood level, above which side effects such as thrombosis may occur (Brobrow, RS, JABFP (2005) 18(2):147-149), establishing the therapeutic window for the composition.

[00284] One of skill in the art can, by the means disclosed herein or by other methods known in the art, confirm that the administered CFXTEN remains at therapeutic blood levels to maintain hemostasis for the desired interval or requires adjustment in dose or length or sequence of XTEN. Further, the determination of the appropriate dose and dose frequency to keep the CFXTEN within the therapeutic window establishes the therapeutically effective dose regimen; the schedule for administration of multiple consecutive doses using a therapeutically effective dose of the fusion protein to a subject in need thereof resulting in consecutive C_{max} peaks and/or C_{min} troughs that remain above therapeutically-effective concentrations and result in an improvement in at least one measured parameter relevant for the

target condition. In one embodiment, the CFXTEN or a pharmaceutical compositions comprising CFXTEN administered at an appropriate dose to a subject results in blood concentrations of the CFXTEN fusion protein that remains above the minimum effective concentration to maintain hemostasis for a period at least about two-fold longer compared to the corresponding FVIII not linked to XTEN and administered at a comparable dose; alternatively at least about three-fold longer; alternatively at least about four-fold longer; alternatively at least about five-fold longer; alternatively at least about six-fold longer; alternatively at least about seven-fold longer; alternatively at least about eight-fold longer; alternatively at least about nine-fold longer, alternatively at least about ten-fold longer, or at least about twenty-fold longer or greater compared to the corresponding FVIII not linked to XTEN and administered at a comparable dose. As used herein, an "appropriate dose" means a dose of a drug or biologic that, when administered to a subject, would result in a desirable therapeutic or pharmacologic effect (e.g., hemostasis) and/or a blood concentration within the therapeutic window.

[00285] In practicing the invention, CFXTEN with longer terminal half-life are generally preferred, so as to improve patient convenience, to increase the interval between doses and to reduce the amount of drug required to achieve a sustained effect. The enhanced PK parameters allow for reduced dosing of the subject compositions, compared to FVIII not linked to XTEN, particularly for those hemophilia A subjects receiving routine prophylaxis.

[00286] As described more fully in the Examples pertaining to pharmacokinetic characteristics of fusion proteins comprising XTEN, it was observed that increasing the total length of the XTEN, singly or in combination, confers a disproportionate increase in the terminal half-life of a fusion protein comprising the XTEN. Accordingly, the invention provides CFXTEN fusion proteins and pharmaceutical compositions comprising CFXTEN wherein the CFXTEN exhibits an enhanced half-life when administered to a subject. In some embodiments, the invention provides monomeric CFXTEN fusion proteins comprising one or more XTEN wherein the number and location of the XTEN are selected to confer an increase in the terminal half-life for the CFXTEN administered to a subject compared to the corresponding FVIII not linked to the XTEN and administered at a comparable dose, wherein the increase is at least about two-fold longer, or at least about three-fold, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about seven-fold, or at least about eight-fold, or at least about nine-fold, or at least about ten-fold, or at least about 15-fold, or at least a 20-fold, or at least a 40-fold or greater increase in terminal half-life compared to the FVIII not linked to the XTEN. In other embodiments, the invention provides CXTEN compositions and pharmaceutical compositions comprising CFXTEN wherein the administration of a composition to a subject in need thereof results in a terminal half-life that is at least 12 h greater, or at least about 24 h greater, or at least about 48 h greater, or at least about 96 h greater, or at least about 144 h greater, or at least about 7 days greater, or at least about 14 days greater, or at least about 21 days greater compared to a comparable dose of FVIII not linked to XTEN. In another embodiment, administration of a coagulation-effective dose of a CFXTEN fusion protein to a subject in need thereof can result in a gain in time between consecutive doses necessary to maintain blood levels of about 0.1 IU/ml of at least 48 h, or at least 72 h, or at least about

96 h, or at least about 120 h, or at least about 7 days, or at least about 14 days, or at least about 21 days between consecutive doses compared to a FVIII not linked to XTEN and administered at a comparable dose.

[00287] In one embodiment, the present invention provides CFXTEN fusion proteins and pharmaceutical compositions comprising CFXTEN that exhibit, when administered to a subject in need thereof, an increase in AUC of at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90%, or at least about a 100%, or at least about 150%, or at least about 200%, or at least about 300%, or at least about 500%, or at least about 1000%, or at least about a 2000% compared to the corresponding FVIII not linked to the XTEN and administered to a subject at a comparable dose. The pharmacokinetic parameters of a CFXTEN can be determined by standard methods involving dosing, the taking of blood samples at timed intervals, and the assaying of the protein using ELISA, HPLC, radioassay, clotting assays, the assays of Table 49, or other methods known in the art or as described herein, followed by standard calculations of the data to derive the half-life and other PK parameters.

[00288] In one embodiment, a smaller IU amount of about two-fold less, or about three-fold less, or about four-fold less, or about five-fold less, or about six-fold less, or about eight-fold less, or about 10-fold less or greater of the fusion protein is administered in comparison to the corresponding FVIII not linked to the XTEN under a dose regimen needed to maintain hemostasis and the fusion protein achieves a comparable area under the curve as the corresponding IU amount of the FVIII not linked to the XTEN needed to maintain hemostasis. In another embodiment, the CFXTEN fusion protein or a pharmaceutical compositions comprising CFXTEN requires less frequent administration for routine prophylaxis of a hemophilia A subject, wherein the dose of fusion protein is administered about every four days, about every seven days, about every 10 days, about every 14 days, about every 21 days, or about monthly to the subject, and the fusion protein achieves a comparable area under the curve as the corresponding FVIII not linked to the XTEN and administered to the subject. In yet other embodiments, an accumulative smaller IU amount of about 5%, or about 10%, or about 20%, or about 40%, or about 50%, or about 60%, or about 70%, or about 80%, or about 90% less of the fusion protein is administered to a subject in comparison to the corresponding IU amount of the FVIII not linked to the XTEN under a dose regimen needed to maintain a blood concentration of 0.1 IU/ml, yet the fusion protein achieves at least a comparable area under the curve as the corresponding FVIII not linked to the XTEN. The accumulative smaller IU amount is measure for a period of at least about one week, or about 14 days, or about 21 days, or about one month.

[00289] In one aspect, the invention provides CFXTEN compositions designed to reduce binding by FVIII binding agents, thereby increasing the terminal half-life of CFXTEN administered to a subject, while still retaining procoagulant activity. It is believed that the CFXTEN of the present invention have comparatively higher and/or sustained activity achieved by reduced active clearance of the molecule by the addition of unstructured XTEN to the FVIII coagulation factor. The clearance mechanisms to remove FVIII from the circulation have yet to be fully elucidated. Uptake, elimination, and inactivation

of coagulation proteins can occur in the circulatory system as well as in the extravascular space. Coagulation factors are complex proteins that interact with a large number of other proteins, lipids, and receptors, and many of these interactions can contribute to the elimination of CFs from the circulation. The protein von Willebrand factor is an example of a FVIII binding agent that binds to FVIII. Factor VIII and von Willebrand factor (VWF) circulate in the blood as a tight, non-covalently linked complex in which VWF serves as a carrier that likely contributes to the protection of FVIII from active cleavage mechanisms, yet nevertheless results in a limitation on the terminal half-life of FVIII. For example: (i) VWF stabilizes the heterodimeric structure of FVIII; (ii) VWF protects FVIII from proteolytic degradation by phospholipid-binding proteases like activated protein C and activated FX (FXa); (iii) VWF interferes with binding of FVIII to negatively charged phospholipid surfaces exposed within activated platelets; (iv) VWF inhibits binding of FVIII to activated FIX (FIXa), thereby denying FVIII access to the FX-activating complex; and (v) VWF prevents the cellular uptake of FVIII (Lenting, P.J., et al., J Thrombosis and Haemostasis (2007) 5(7):1353-1360). In addition, LDL receptor-related protein (LRP1, also known as α2-macrogobulin receptor or CD91) has been identified as a candidate clearance receptor for FVIII, with LRP1 binding sites identified on both chains of the heterodimer form of FVIII (Lenting PJ, et al.,. J Biol Chem (1999) 274: 23734–23739; Saenko EL, et al., J Biol Chem (1999) 274: 37685–37692). LRPs are involved in the clearance of a diversity of ligands including proteases, inhibitors of the Kunitz type, protease serpin complexes, lipases and lipoproteins (Narita, et al., Blood (1998) 2:555-560). It has been shown that the light chain, but not the heavy chain, of factor VIII binds to surface-exposed LRP1 receptor protein (Lentig et al. (J Biol Chem (1999) 274(34):23734-23739; and U.S. Pat. No. 6,919,311), which suggests that LRP1 may play an essential role in the active clearance of proteins like FVIII. While the VWF–FVIII interaction is of high affinity (<1 nM), the complex is nevertheless in a dynamic equilibrium, such that a small but significant portion of the FVIII molecules (5–8%) circulate as a free protein (Leyte A, et al., Biochem J (1989) 257: 679–683; Noe DA. Haemostasis (1996) 26: 289–303). As such, a portion of native FVIII is unprotected by VWF, allowing active clearance mechanisms to remove the unprotected FVIII from the circulation.

[00290] In one embodiment, the invention provides CFXTEN that associate with VWF but have enhanced protection from active clearance receptors conferred by the incorporation of two more XTEN at one or more locations within the FVIII molecule (e.g., locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9 or FIGS. 8-9), wherein the XTEN interfere with the interaction of the resulting CFXTEN with those clearance receptors with the result that the pharmacokinetic properties of the CFXTEN is enhanced compared to the corresponding FVIII not linked to XTEN. In another embodiment, the invention provides CFXTEN that have reduced binding affinity with VWF of at least 5% less, or about 10%, or about 20%, or about 40%, or about 50%, or about 60%, or about 70% less, but are nevertheless configured to have enhanced protection from active clearance receptors conferred by the incorporation of XTEN at one or more locations within the FVIII molecule, wherein the XTEN interfere with the interaction of factor VIII with those receptors. In the foregoing embodiments, the CFXTEN have an increased terminal half-life of at least about 12 h, or 24 h, or 48 h, or 72 h, or 96 h, or 120 h, or

144 h, or 7 days, or 10 days, or 14 days, or 21 days compared to the FVIII not linked to XTEN. The invention provides a method to create CFXTEN with reduced clearance wherein the CFXTEN fusion proteins created with the multiple insertions are evaluated for inhibition of binding to clearance receptors, compared to FVIII not linked to XTEN, using *in vitro* binding assays or *in vivo* pharmacokinetic models described herein or other assays known in the art, and selecting those that demonstrate reduced binding yet retain procoagulant FVIII activity. In addition, the foregoing fusion proteins can be optimized to have increased Ratio XTEN Radii of at least 2.0-3.5 in order to achieve pharmacokinetic properties that are further enhanced. Table 5, Table 6, Table 7, Table 8, and Table 9 and FIGS. 8-9 provide non-limiting examples of XTEN insertion points within the factor VIII sequence. Using such insertion points, the invention contemplates CFXTEN compositions that have configurations with multiple XTEN inserted with about 100, or about 200, or about 300, or about 400, or about 500 amino acids separating at least three XTEN to further increase the protection against active clearance mechanisms and, hence, increase the terminal half-life of the CFXTEN. Not to be bound by a particular theory, the XTEN of the CFXTEN compositions with high net charge (e.g., CFXTEN comprising AE family XTEN) are expected, as described above, to have less non-specific interactions with various negatively-charged surfaces such as blood vessels, tissues, or various receptors, which would further contribute to reduced active clearance. Conversely, the XTEN of the CFXTEN compositions with a low (or no) net charge (e.g., CFXTEN comprising AG family XTEN) are expected to have a higher degree of interaction with surfaces that, while contributing to active clearance, can potentiate the activity of the associated coagulation factor, given the known contribution of cell (e.g., platelets) and vascular surfaces to the coagulation process and the intensity of activation of coagulation factors (Zhou, R., et al., Biomaterials (2005) 26(16):2965-2973; London, F., et al. Biochemistry (2000) 39(32):9850–9858). The invention, in part, takes advantage of the fact that certain ligands wherein reduced binding to a clearance receptor, either as a result of a decreased on-rate or an increased off-rate, may be effected by the obstruction of a receptor site by an inserted XTEN forming random coil, resulting in the reduced binding. The choice of the particular configuration of the CFXTEN fusion protein can be tested by methods disclosed herein to confirm those configurations that reduce the degree of binding to a clearance receptor such that a reduced rate of active clearance is achieved. In one embodiment, the CFXTEN comprises a FVIII-XTEN sequence that has one or more XTEN inserted at locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9 or FIGS. 8-9 wherein the terminal half-life of the CFXTEN is increased at least about two-fold, or at least about three-fold, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about eight-fold, or at least about ten-fold, or at least about twenty-fold compared to a FVIII not linked to an XTEN. In another embodiment, the CFXTEN comprises a FVIII-XTEN sequence that has a first and at least a second XTEN inserted at a first and second location selected from Table 5, Table 6, Table 7, Table 8, and Table 9 or FIGS. 8-9 wherein the terminal half-life of the CFXTEN is increased at least about two-fold, or at least about three-fold, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about eight-fold, or at least about ten-fold, or at least about twenty-fold compared to a FVIII not linked to an XTEN. In yet another embodiment, the CFXTEN comprises a

FVIII-XTEN sequence that incorporates multiple XTEN sequences using three of more XTEN insertion locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9 or FIGS. 8-9 separated by about 100, or about 200, or about 300, or about 400, or about 500 amino acids, wherein the terminal half-life of the CFXTEN is increased at least about two-fold, or at least about three-fold, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about eight-fold, or at least about ten-fold, or at least about twenty-fold compared to a FVIII not linked to an XTEN. In the foregoing embodiments hereinabove described in this paragraph, the XTEN incorporated into the CFXTEN configurations can be identical or they can be different, and can have at least about 80%, or 90%, or 91%, or 92%, or 93%, or 94%, or 95%, or 96%, or 97%, or 98%, or 99%, sequence identity to a sequence from any one of Tables 3, 4, and 13-17, and can optionally include one or more cleavage sequences from Table 12, facilitating release of one or more of the XTEN from the CFXTEN fusion protein.

[00291] In one embodiment, the invention provides CFXTEN that enhance the pharmacokinetics of the fusion protein by linking one or more XTEN to the FVIII component of the fusion protein wherein the fusion protein has an increase in apparent molecular weight factor of at least about two-fold, or at least about three-fold, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about seven-fold, or at least about eight-fold, or at least about ten-fold, or at least about twelve-fold, or at least about fifteen-fold, and wherein the terminal half-life of the CFXTEN when administered to a subject is increased at least about two fold, or at least about four fold, or at least about eight fold, or at least about 10-fold or more compared to the corresponding FVIII not linked to XTEN. In the foregoing embodiment, wherein at least two XTEN molecules are incorporated into the CFXTEN, the XTEN can be identical or they can be of a different sequence composition, net charge, or length. The XTEN can have at least about 80%, or 90%, or 91%, or 92%, or 93%, or 94%, or 95%, or 96%, or 97%, or 98%, or 99%, sequence identity to a sequence from any one of Tables 3, 4, and 13-17, and can optionally include one or more cleavage sequences from Table 12, facilitating release of one or more of the XTEN from the CFXTEN fusion protein.

[00292] Thus, the invention provides CFXTEN compositions in which the degree of activity, bioavailability, half-life or physicochemical characteristic of the fusion protein can be tailored by the selection and placement of the type and length of the XTEN in the CFXTEN compositions. Accordingly, the invention contemplates compositions in which a FVIII from Table 1 and XTEN or XTEN fragment from any one of Tables 3, 4, or 13-17 are produced, for example, in a configuration selected from any one of formulae I-VIII or the XTEN are inserted at locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9 or FIGS. 8-9 such that the construct has the desired property.

[00293] The invention provides methods to produce the CFXTEN compositions that can maintain the FVIII component at therapeutic levels in a subject in need thereof for at least a two-fold, or at least a three-fold, or at least a four-fold, or at least a five-fold greater period of time compared to comparable dosages of the corresponding FVIII not linked to XTEN. In one embodiment of the method, the subject is receiving routine prophylaxis to prevent bleeding episodes. In another embodiment of the method, the subject is receiving treatment for a bleeding episode. In another embodiment of the method, the subject

is receiving treatment to raise the circulating blood concentration of procoagulant FVIII above 1%, or above 1-5%, or above 5-40% relative to FVIII concentrations in normal plasma. "Procoagulant" as used herein has its general meaning in the art and generally refers to an activity that promotes clot formation, either in an *in vitro* assay or *in vivo*. The method to produce the compositions that can maintain the FVIII component at therapeutic levels includes the steps of selecting one or more XTEN appropriate for conjugation to a FVIII to provide the desired pharmacokinetic properties in view of a given dose and dose regimen, creating a gene construct that encodes the CFXTEN in one of the configurations disclosed herein, transforming an appropriate host cell with an expression vector comprising the encoding gene, expressing the fusion protein under suitable culture conditions, recovering the CFXTEN, administration of the CFXTEN to a mammal followed by assays to verify the pharmacokinetic properties and the activity of the CFXTEN fusion protein (e.g., the ability to maintain hemostasis or serve as a procoagulant) and the safety of the administered composition. Those compositions exhibiting the desired properties are selected for further use. CFXTEN created by the methods provided herein can result in increased efficacy of the administered composition by, amongst other properties, maintaining the circulating concentrations of the procoagulant FVIII component at therapeutic levels for an enhanced period of time.

[00294] The invention provides methods to assay the CFXTEN fusion proteins of differing composition or configuration in order to provide CFXTEN with the desired degree of procoagulant and therapeutic activity and pharmacokinetic properties, as well as a sufficient safety profile. Specific *in vitro* and *in vivo* assays or animal models are used to assess the activity and functional characteristics of each configured CFXTEN and/or FVIII component to be incorporated into CFXTEN, including but not limited to the assays of the Examples, those assays of Table 49, as well as the following assays or other such assays known in the art for assaying the properties and effects of FVIII. Functional assays can be conducted that allow determination of coagulation activity, such as one-stage clotting assay and two-stage clotting assay (Barrowcliffe TW, Semin Thromb Hemost. (2002) 28(3):247-256), activated partial prothrombin (aPTT) assays (Belaaouaj AA et al., J. Biol. Chem. (2000) 275:27123-8; Diaz-Collier JA. Haemost (1994) 71:339-46), chromogenic FVIII assays (Lethagen, S., et al., Scandinavian J Haematology (1986) 37:448–453), or animal model pharmacodynamic assays including bleeding time or thrombelastography (TEG or ROTEM), among others. Other assays include determining the binding affinity of a CFXTEN for the target substrate using binding or competitive binding assays, such as Biacore assays with chip-bound receptors or binding proteins or ELISA assays, as described in US Patent 5,534,617, assays described in the Examples herein, radio-receptor assays, or other assays known in the art. Other assays to determine the binding of FVIII inhibitors to CFXTEN include the Bethesda assay or the Nijmegen modification of the Bethesda assay. The foregoing assays can also be used to assess FVIII sequence variants (assayed as single components or as CFXTEN fusion proteins) and can be compared to the native FVIII to determine whether they have the same degree of procoagulant activity as the native CF, or some fraction thereof such that they are suitable for inclusion in CFXTEN; e.g., at least about 10%, or at least about 20$, or about 30%, or at least about 40%, or at least about 50%, or at least about

60%, or at least about 70%, or at least about 80%, or at least about 90% of the activity compared to the native FVIII.

[00295] Dose optimization is important for all drugs. A therapeutically effective dose or amount of the CFXTEN varies according to factors such as the disease state, age, sex, and weight of the individual, and the ability of the administered fusion protein to elicit a desired response in the individual. For example, a standardized single dose of FVIII for all patients presenting with diverse bleeding conditions or abnormal clinical parameters (e.g., neutralizing antibodies) may not always be effective. Hemophilia A patients with trauma, who have undergone surgery, or that have high titers of FVIII inhibitory antibodies generally will require higher and more frequent dosing. Generally, dosage level is adjusted in frequency, duration, and units in keeping with the severity and duration of each patient's bleeding episode. Accordingly, the CFXTEN is included in the pharmaceutically acceptable carrier, delivery vehicle, or stabilizer in an amount sufficient to deliver to a patient a therapeutically effective amount of the fusion protein to stop bleeding, as measured by standard clotting assays. A consideration of these factors is well within the purview of the ordinarily skilled clinician for the purpose of determining the therapeutically or pharmacologically effective amount of the CFXTEN and the appropriated dosing schedule, versus that amount that would result in insufficient potency such that clinical improvement or the arrest of bleeding is not achieved.

[00296] The invention provides methods to establish a dose regimen for the CFXTEN pharmaceutical compositions of the invention. The methods include administration of consecutive doses of a therapeutically effective amount of the CFXTEN pharmaceutical composition using variable periods of time between doses to determine that interval of dosing sufficient to achieve and/or maintain the desired parameter, blood level or clinical effect; such consecutive doses of a therapeutically effective amount at the effective interval establishes the therapeutically effective dose regimen for the CFXTEN for a factor VIII-related disease state or condition. A prophylactically effective amount refers to an amount of CFXTEN required for the period of time necessary to prevent a physiologic or clinical result or event; e.g., delayed onset of a bleeding episode or maintaining blood concentrations of procoagulant FVIII or equivalent above a threshold level (e.g., 1-5% to 5-40% of normal). In the methods of treatment, the dosage amount of the CFXTEN that is administered to a subject ranges from about 5 to 300 IU/kg/dose, or from about 10 to 100 IU/kg/dose, or from about 20 to about 65 IU/kg/dose, or from about 20 to about 40 IU/kg/dose for a subject. A suitable dosage may also depend on other factors that may influence the response to the drug; e.g., bleeding episodes generally requiring higher doses at more frequent intervals compared to prophylaxis.

[00297] In some embodiments, the method comprises administering a therapeutically-effective amount of a pharmaceutical composition comprising a CFXTEN fusion protein composition and at least one pharmaceutically acceptable carrier to a subject in need thereof, wherein the administration results in a greater improvement in at least one parameter or physiologic condition associated with a FVIII deficiency or coagulopathy, or results in a more favorable clinical outcome mediated by the FVIII component of the CFXTEN compared to the effect on the parameter, condition or clinical outcome

mediated by administration of a pharmaceutical composition comprising a FVIII not linked to XTEN and administered at a comparable dose. Non-limiting examples of parameters that are improved include blood concentration of procoagulant FVIII, a reduced activated partial prothrombin (aPTT) assay time, a reduced one-stage or two-stage clotting assay time, delayed onset of a bleeding episode, a reduced chromogenic FVIII assay time, a reduced bleeding time, resolution of a bleeding event, or a reduced Bethesda titer to the CFXTEN relative to native FVIII. In one embodiment of the foregoing, the improvement is achieved by administration of the CFXTEN pharmaceutical composition at a dose that achieves a circulating concentration of procoagulant FVIII (or equivalent) above a threshold level (e.g., 1-5% to 5-40% of normal FVIII levels), thereby establishing the therapeutically effective dose. In another embodiment of the foregoing, the improvement is achieved by administration of multiple consecutive doses of the CFXTEN pharmaceutical composition using a therapeutically effective dose regimen that maintains a circulating concentration of procoagulant FVIII (or equivalent) above a threshold level (e.g., 1-5% to 5-40% of normal FVIII levels) for the length of the dosing period. In another embodiment of the method, the administration of at least two consecutive doses of the CFXTEN pharmaceutical composition using a therapeutically effective dose regimen maintains a circulating concentration of procoagulant FVIII (or equivalent) above about 1%,, 2%, 3%, 4%, 5%, 10%, 15%, 20%, 30%, or 40% of normal FVIII levels for a period that is at least about three-fold longer; alternatively at least about four fold longer; alternatively at least about five fold longer; alternatively at least about six fold longer; alternatively at least about seven-fold longer; alternatively at least about eight-fold longer; alternatively at least about nine-fold longer or at least about ten-fold longer compared to a FVIII not linked to XTEN and administered using a therapeutically effective dose regimen

[00298] In one embodiment, the CFXTEN or a pharmaceutical compositions comprising CFXTEN administered at a therapeutically effective dose regimen results in a gain in time of at least about three-fold longer; alternatively at least about four-fold longer; alternatively at least about five-fold longer; alternatively at least about six-fold longer; alternatively at least about seven-fold longer; alternatively at least about eight-fold longer; alternatively at least about nine-fold longer or at least about ten-fold longer between at least two consecutive C_{max} peaks and/or C_{min} troughs for blood levels of the fusion protein compared to the corresponding biologically active protein of the fusion protein not linked to the XTEN and administered at a comparable dose regimen to a subject. In another embodiment, the CFXTEN administered at a therapeutically effective dose regimen results in a comparable improvement in one, or two, or three or more measured parameters using less frequent dosing or a lower total dosage in IUs of the fusion protein of the pharmaceutical composition compared to the corresponding biologically active protein component(s) not linked to the XTEN and administered to a subject using a therapeutically effective dose regimen for the FVIII. The measured parameters include any of the clinical, biochemical, or physiological parameters disclosed herein, or others known in the art for assessing subjects with factor VIII-related conditions.

(b) Pharmacology and Pharmaceutical Properties of CFXTEN

[00299] The present invention provides CFXTEN compositions comprising FVIII covalently linked to XTEN that have enhanced pharmaceutical and pharmacology properties compared to FVIII not linked to XTEN, as well as methods to enhance the therapeutic and/or procoagulant effect of the FVIII components of the compositions. In addition, the invention provides CFXTEN compositions with enhanced properties compared to those art-known fusion proteins of factor VIII containing albumin, immunoglobulin polypeptide partners, polypeptides of shorter length and/or polypeptide partners with repetitive sequences. In addition, CFXTEN fusion proteins provide significant advantages over chemical conjugates, such as pegylated constructs of FVIII, notably the fact that recombinant CFXTEN fusion proteins can be made in host cell expression systems, which can reduce time and cost at both the research and development and manufacturing stages of a product, as well as result in a more homogeneous, defined product with less toxicity from both the product and metabolites of the CFXTEN compared to pegylated conjugates.

[00300] As therapeutic agents, the CFXTEN possesses a number of advantages over therapeutics not comprising XTEN, including one or more of the following non-limiting properties: increased solubility, increased thermal stability, reduced immunogenicity, increased apparent molecular weight, reduced renal clearance, reduced proteolysis, reduced metabolism, enhanced therapeutic efficiency, less frequent dosage regimen with increased time between doses capable of maintaining hemostasis in a subject with hemophilia A, the ability to administer the CFXTEN composition subcutaneously or intramuscularly, a "tailored" rate of absorption when administered subcutaneously or intramuscularly, enhanced lyophilization stability, enhanced serum/plasma stability, increased terminal half-life, increased solubility in blood stream, decreased binding by neutralizing antibodies, decreased active clearance, tailored substrate binding affinity, stability to degradation, stability to freeze-thaw, stability to proteases, stability to ubiquitination, ease of administration, compatibility with other pharmaceutical excipients or carriers, persistence in the subject, increased stability in storage (e.g., increased shelf-life), and the like. The net effect of the enhanced properties is that the use of a CFXTEN composition can result in an overall enhanced therapeutic effect compared to a FVIII not linked to XTEN, result in economic benefits associated with less frequent dosing, and/or result in improved patient compliance when administered to a subject with a factor VIII-related condition.

[00301] The invention provides CFXTEN compositions and pharmaceutical compositions comprising CFXTEN wherein the administration of the composition results in an improvement in at least one of the clinical or biochemical parameters disclosed herein as being useful for assessing the subject diseases, conditions or disorders. Non-limiting examples of parameters that are improved include blood concentrations of procoagulant FVIII, a reduced activated partial prothrombin (aPTT) assay time, a reduced one-stage or two-stage clotting assay time, delayed onset of a bleeding episode, a reduced chromogenic FVIII assay time, a reduced bleeding time, resolution of a bleeding event, or a reduced Bethesda titer to the CFXTEN relative to native FVIII. The enhanced pharmacokinetic properties of the subject CFXTEN permits using an accumulatively lower IU dose of fusion protein to maintain the parameter compared to the corresponding FVIII component not linked to the XTEN. In one

embodiment, the total dose in IUs of an CFXTEN of the embodiments needed to achieve and maintain the improvement in the at least one parameter for about 2-7 days is at least about three-fold lower, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about eight-fold, or at least about 10-fold lower compared to the corresponding FVIII component not linked to the XTEN. In another embodiment, the total dose in IUs of a subject CFXTEN needed to achieve and maintain the improvement in the at least one parameter over two, three or four consecutive doses is at least about three-fold lower, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about eight-fold, or at least about 10-fold lower compared to the corresponding FVIII component not linked to the XTEN. Alternatively, the invention provides certain embodiments of CFXTEN wherein the period between consecutive administrations that results in achieving and maintaining the improvement in at least one parameter is at least about three-fold, or at least about four-fold, or at least about five-fold, or at least about six-fold, or at least about eight-fold, or at least about 10-fold longer compared to the corresponding FVIII component not linked to the XTEN and administered at a comparable IU dose. Alternatively, the invention provides certain embodiments of CFXTEN wherein administration of 25 IU/kg results in a 30% improvement in a aPTT assay (or similar coagulation assay) time in a hemophilia A subject compared to 25 IU/kg of the corresponding FVIII not linked to XTEN when assayed at about 2-7 days after administration. In yet another embodiment, the invention provides CFXTEN wherein administration of 25 IU/kg results in a 30% improvement in a bleeding time assay time in a hemophilia A subject compared to 25 IU/kg of the corresponding FVIII not linked to XTEN when assayed at about 2-7 days after administration.

[00302] In one embodiment, XTEN as a fusion partner increases the solubility of the FVIII payload. Accordingly, where enhancement of the pharmaceutical or physicochemical properties of the FVIII is desirable, such as the degree of aqueous solubility or stability, the length and/or the motif family composition of the XTEN sequences incorporated into the fusion protein may each be selected to confer a different degree of solubility and/or stability on the respective fusion proteins such that the overall pharmaceutical properties of the CFXTEN composition are enhanced. The CFXTEN fusion proteins can be constructed and assayed, using methods described herein, to confirm the physicochemical properties and the choice of the XTEN length sequence or location adjusted, as needed, to result in the desired properties. In one embodiment, the CFXTEN has an aqueous solubility that is at least about 25% greater compared to a FVIII not linked to the XTEN, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 75%, or at least about 100%, or at least about 200%, or at least about 300%, or at least about 400%, or at least about 500%, or at least about 1000% greater than the corresponding FVIII not linked to XTEN.

[00303] The invention provides methods to produce and recover expressed CFXTEN from a host cell with enhanced solubility and ease of recovery compared to FVIII not linked to XTEN. In one embodiment, the method includes the steps of transforming a eukaryotic host cell with a polynucleotide encoding a CFXTEN with one or more XTEN components of cumulative sequence length greater than about 100, or greater than about 200, or greater than about 400, or greater than about 600, or greater than

about 800, or greater than about 1000, or greater than about 2000, or greater than about 3000 amino acid residues, expressing the CFXTEN fusion protein in the host cell under suitable culture and induction conditions, and recovering the expressed fusion protein in soluble form. In one embodiment, the one or more XTEN of the CFXTEN fusion proteins each have at least about 80% sequence identity, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% sequence identity compared to one or more XTEN selected from any one of Tables 4, and 13-17, or fragments thereof, and the FVIII have at least about 80% sequence identity, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, or 100% sequence identity compared to a FVIII selected from Table 1, and the CFXTEN components are in an N- to C-terminus configuration selected from any one of the configuration embodiments disclosed herein.

VI). USES OF THE CFXTEN COMPOSITIONS

[00304] The invention provides methods and regimens for achieving a beneficial effect in a factor VIII-related condition by the administration of compositions comprising CFXTEN. As used herein, "factor VIII-related condition" is intended to include, but is not limited to factor VIII deficiencies, bleeding disorders related to factor VIII deficiency, hemophilia A, neutralization of factor VIII by anti-FVIII antibodies or other factor VIII inhibitors, and bleeding episodes resulting from trauma or surgery or vascular injury and other such conditions that can be ameliorated or corrected by administration of FVIII to a subject. The inventive methods achieve a beneficial effect while addressing disadvantages and/or limitations of other methods of treatment using factor VIII preparations that have a relatively short terminal half-life, require frequent administrations, are neutralized by inhibitors or have unfavorable pharmacoeconomics.

[00305] Hemostasis is regulated by multiple protein factors, and such proteins, as well as analogues thereof, have found utility in the treatment of factor VIII-related conditions. However, the use of commercially-available FVIII has met with less than optimal success in the management of subjects afflicted with such conditions. In particular, dose optimization and frequency of dosing is important for FVIII used in maintaining circulating FVIII concentrations above threshold levels needed for hemostasis, as well as the treatment or prevention of bleeding episodes in hemophilia A subjects. The fact that commercially-available FVIII products have a short half-life necessitates frequent dosing in order to achieve clinical benefit, which results in difficulties in the management of such patients.

[00306] As established by the Subcommittee on Factor VIII and Factor IX of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis (posted on the ISTH Website 29 November, 2000), the most widely used measure of the severity of hemophilia A is established by determining the circulating concentrations of plasma FVIII procoagulant levels, with persons with <1% (< 0.01 IU/ml) factor VIII defined as severe; 1-5% (0.01 - 0.05 IU/ml) as moderately severe; and >5-40% (0.05 - <0.40 IU/ml) as mild, where normal is 1 IU/ml of factor VIIIC (100%).

[00307] The invention provides methods of treating a subject suffering from or at risk of developing a factor VIII-related condition. More particularly, the invention provides methods for treating or preventing controlling bleeding in subject. The subject can be any animal but preferably is a human. In one embodiment, the method comprises administering a coagulation-effective amount of a CFXTEN composition to the subject in need thereof. In another embodiment, the method comprises the step of administering to the subject with a bleed a coagulation-effective amount of a pharmaceutical composition that includes a CFXTEN, wherein the administration results in an arrest or attenuation of the bleeding. As used herein, "coagulation-effective amount" is an amount of a FVIII composition that, when administered to a subject, is sufficient to effect hemostasis or other beneficial or desired therapeutic (including preventative) result. In practicing the present invention, it will be understood that a coagulation-effective amount can be administered in one or more administrations. Precise coagulation-effective amounts of the pharmaceutical composition to be administered will be guided by the judgment of the practitioner, however, the unit dose will generally depend on the severity or cause of the bleeding and the amount of pre-existing FVIII in the subject. In a particular embodiment of the method of treating a bleed, a coagulation-effective amount of a pharmaceutical compositions comprising CFXTEN is administered to a subject suffering from a bleeding episode, wherein the administration results in the resolution of the bleeding for a duration at least two-fold, or at least three-fold, or at least four-fold longer compared to a FVIII not linked to XTEN and administered to a comparable subject with a comparable bleed at a comparable dose.

[00308] In another embodiment, the administration of a coagulation-effective amount of a CFXTEN composition to a subject with a factor VIII-related condition results in a 10%, or 20%, or 30%, or 40%, or 50%, or 60%, or 70% or greater improvement of one or more biochemical, physiological or clinical parameters associated with the FVIII condition, compared to the FVIII not linked to XTEN, when measured at between 2 and 7 days after administration. In another embodiment, the administration of a coagulation-effective amount of a CFXTEN composition to the subject in need thereof results in an improvement of one or more biochemical, physiological or clinical parameters associated with the FVIII condition for a period at least two-fold longer, or at least four-fold longer, or at least five-fold longer, or at least six-fold longer compared to period achieved by a FVIII not linked to XTEN and administered at a comparable dose. Non-limiting examples of parameters that are improved for a longer duration include blood concentrations of procoagulant FVIII, a reduced activated partial prothrombin (aPTT) assay time, a reduced one-stage or two-stage clotting assay time, delayed onset of a bleeding episode, a reduced chromogenic FVIII assay time, a reduced bleeding time, among other FVIII-related parameters known in the art. In the foregoing embodiments of the paragraph, the administered CFXTEN comprises a FVIII with at least about 80%, or at least about 90%, or at least about 95%, or at least about 97%, or at least about 99% sequence identity to a factor VIII of Table 1 and one or more XTEN sequences with at least about 80%, or at least about 90%, or at least about 95%, or at least about 97%, or at least about 99% sequence identity to an XTEN of Table 4 inserted into the FVIII at one or more locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9, or as depicted in FIGS. 8-9. In certain embodiments, at

least one XTEN insertion site of the CFXTEN is selected from amino acids 32, 220, 224, 336, 339, 390, 399, 416, 603, 1656, 1711, 1725, 1905 and 1910 (numbered relative to mature native human FVIII).

[00309] In a particular embodiment of the method of treatment, a coagulation-effective amount of CFXTEN fusion protein administered to a subject suffering from hemophilia A is sufficient to increase the circulating FVIII procoagulant concentration to greater than 0.05 IU/ml and to maintain hemostasis for at least about 24 h, or at least about 48 h, or at least about 72 h, or at least about 96 h, or at least about 120 h, or at least about 144 h, or at least about 168 h, or greater. In another embodiment, the administration of a coagulation-effective amount of a pharmaceutical composition comprising CFXTEN to a subject in need thereof results in a greater reduction in a one-stage clotting assay time of at least about 5%, or about 10%, or about 20%, or about 30%, or about 40%, or about 50%, or about 60%, or about 70%, or more in a blood sample from the subject at 2-7 days after the administration compared to the assay time in a subject after administration of a comparable amount of the corresponding FVIII not linked to XTEN. In another embodiment, the administration of a therapeutically effective amount of a CFXTEN or a pharmaceutical compositions comprising CFXTEN to a subject in need thereof results in a greater reduction in the activated partial prothrombin time of at least about 5%, or about 10%, or about 20%, or about 30%, or about 40%, or about 50%, or about 60%, or about 70%, or more in a blood sample from the subject 2-7 days after administration compared to the activated partial prothrombin time in a subject after administration of a comparable amount of the corresponding FVIII not linked to XTEN. In another embodiment, the administration of a CFXTEN or a pharmaceutical compositions comprising CFXTEN to a subject in need thereof using a therapeutically effective amount results in maintenance of activated partial prothrombin times within 30% of normal in a blood sample from the subject for a period of time that is at least two-fold, or at least about three-fold, or at least about four-fold longer compared to that of a FVIII not linked to XTEN and administered to a subject using a comparable dose.

[00310] In one embodiment of the method of treatment, the CFXTEN fusion protein is formulated and administered as a pharmaceutical composition comprising the CFXTEN in admixture with a pharmaceutically acceptable excipient. Methods for making pharmaceutical formulations are well known in the art. Techniques and formulations generally may be found in Remington's Pharmaceutical Sciences, 18th Edition, Mack Publishing Co., Easton, Pa. 1990 (See, also, Wang and Hanson, Parenteral Formulations of Proteins and Peptides: Stability and Stabilizers, Journal of Parenteral Science and Technology , Technical Report No. 10, Supp. 42-2S (1988)).

[00311] In another aspect, the invention provides a regimen for treating a hemophilia A patient, said regimen comprising a composition comprising a CFXTEN fusion protein. In one embodiment of the regimen for treating a hemophilia A patient, the regimen further comprises the step of determining the amount of pharmaceutical composition comprising the CFXTEN needed to achieve hemostasis in the patient. In some embodiments of the regimen, (i) a smaller IU amount of about two-fold less, or about three-fold less, or about four-fold less, or about five-fold less, or about six-fold less, or about eight-fold less, or about 10-fold less of the pharmaceutical composition comprising CFXTEN is administered to a subject in need thereof in comparison to the corresponding coagulation factor not linked to the XTEN

under an otherwise same dose regimen, and the fusion protein achieves a comparable area under the curve (based on IU/ml) and/or a comparable therapeutic effect as the corresponding FVIII not linked to the XTEN; (ii) the pharmaceutical composition is administered less frequently (e.g., every three days, about every seven days, about every 10 days, about every 14 days, about every 21 days, or about monthly) in comparison to the corresponding FVIII not linked to the XTEN under an otherwise same dose amount, and the fusion protein achieves a comparable area under the curve and/or a comparable therapeutic effect as the corresponding coagulation factor not linked to the XTEN; or (iii) an accumulative smaller IU amount of at least about 20%, or about 30%, or about 40%, or about 50%, or about 60%, or about 70%, or about 80%, or about 90% less of the pharmaceutical composition is administered in comparison to the corresponding FVIII not linked to the XTEN under an otherwise same dose schedule and the CFXTEN fusion protein achieves a comparable therapeutic effect as the corresponding FVIII not linked to the XTEN. The accumulative smaller IU amount is measured for a period of at least about one week, or about 14 days, or about 21 days, or about one month. In the foregoing embodiments, the therapeutic effect can be determined by any of the measured parameters described herein, including but not limited to blood concentration of procoagulant FVIII, a reduced activated partial prothrombin (aPTT) assay time, a reduced one-stage or two-stage clotting assay time, delayed onset of a bleeding episode, a reduced chromogenic FVIII assay time, a reduced bleeding time, resolution of a bleeding event, or a reduced Bethesda titer to the CFXTEN relative to native FVIII, fibrinogen levels, or other assays known in the art for assessing coagulopathies of FVIII. In another embodiment, the invention provides CFXTEN for use in a regimen for a treating a hemophilia A subject comprising administering an CFXTEN composition in two or more successive doses to the subject at an effective amount, wherein the adminstration results in at least a 10%, or 20%, or 30%, or 40%, or 50%, or 60%, or 70%, or 80%, or 90% greater improvement of at least one, two, or three parameters associated with the disease compared to a FVIII not linked to XTEN and administered using a comparable dose.

[00312] In one aspect, the present invention relates to a method of preventing or treating the bleeding in a patient, optionally a haemophilia A patient, having pre-existing inhibitor(s) against FVIII. Inhibitory antibodies against FVIII commonly develop in hemophiliacs, where the overall incidence of developing an inhibitor is 15-30%, particularly in haemophiliacs who are heavily exposed to FVIII concentrates (Algiman et al. Natural antibodies to factor VIII (anti-hemophilic factor) in healthy individuals. PNAS USA (1992) 89: 3795-3799). However, inhibitory antibodies also occur in patients in auto-immune disorders, malignancies (such as lymphoproliferative disorders, lymphomas and solid tumors), during pregnancy and in the post-partum state. Inhibition can also occur when antibodies interfere with the binding of FVIII to FIX and FX. Simultaneously or alternatively, anti-FVIII antibodies can interfere with the binding of von Willebrand factor and/or phospholipids to FVIII, affecting coagulation and/or half-life of FVIII. The presence of inhibitory antibodies is often first detected with symptoms such as easy bruising and uncontrolled bleeding, and is usually referred to as acquired hemophilia. Anti-FVIII antibodies can be determined by different methods including quantitation of anti-FVIII activity in coagulation assays, ELISA for FVIII inhibitors and purification using chromatography and

immunoadsorption (Algiman et al., 1992). Accordingly, the inventive methods are used in the treatment or prevention of any condition associated with or characterized by the presence of inhibitory antibodies to FVIII. In one embodiment, the invention provides a method of treating a patient having a pre-existing inhibitor against FVIII, the method comprising the step of administering to the patient a coagulation-effective amount of a CFXTEN fusion protein that must be administered to achieve hemostasis, wherein the coagulation-effective amount of fusion protein administered is reduced in comparison to the amount of FVIII not linked to XTEN (or native FVIII) that must be administered to achieve hemostasis. In the method, the reduced amount of CFXTEN is about two-fold, or three-fold, or four-fold, or five-fold less in IU/kg compared to the corresponding FVIII not linked to XTEN. In another embodiment of the method, the amount of CFXTEN that is administered as a dose to achieve hemostasis is at least 20 to 40 IU/kg less, or 30 to 60 IU/kg less, or 40 to 80 IU/kg less, or 60 to 100 IU/kg less, or 100 to 140 IU/kg less, or 120 to 180 IU/kg less, or 140 to 200 IU/kg less compared to the corresponding FVIII not linked to XTEN or to native FVIII required to achieve hemostasis. In another embodiment, the invention provides a method of treating a bleeding episode in a hemophilia A subject having a titer of at least 10, or 20, or 30, or 40, or 50, or 75, or 100, or 150, or 200 or more Bethesda units against a FVIII not linked to XTEN, wherein the dose of CFXTEN fusion protein required to arrest the bleeding epidose is at least two-fold, or three-fold, or four-fold, or five-fold, or six-fold, or seven-fold, or eight-fold, or nine-fold, or 10-fold less in comparison to the amount of FVIII not linked to XTEN (or native FVIII) that must be administered to achieve hemostasis in a comparable subject. It will be understood by one of skill in the art that the amount of procoagulant administered to maintain hemostasis will depend on the severity of FVIII deficiency and/or the frequency or duration of bleeding.

[00313] A particular object of the present invention relates to use of CFXTEN with reduced binding by FVIII inhibitors that bind the A2 and/or C2 domains of Factor VIII as a drug. Such a drug is advantageously used for maintaining hemostasis in a patient suffering from haemophilia, wherein such patient has circulating FVIII inhibitors directed against the A2 domain and/or C2 domain of Factor VIII. In one embodiment, the invention provides a method of treatment, the method comprising the step of administering to the patient with a A2 domain-binding inhibitor a coagulation-effective amount of a CFXTEN fusion protein, wherein the CFXTEN exhibits at least 10%, or 20%, or 30%, or 40%, or 50%, or 60%, or 70%, or 80% or less binding to an inhibitor that binds the A2 domain of FVIII, compared to the FVIII not linked to XTEN or to native FVIII, and wherein the administration results in hemostasis. In another embodiment, the invention provides a method of treatment, the method comprising the step of administering to the patient with a C2 domain-binding inhibitor a coagulation-effective amount of a CFXTEN fusion protein, wherein the CFXTEN exhibits at least 10%, or 20%, or 30%, or 40%, or 50%, or 60%, or 70%, or 80% or less binding to an inhibitor that binds the C2 domain of FVIII, compared to the FVIII not linked to XTEN or to native FVIII, and wherein the administration results in hemostasis. The reduced binding of the subject CFXTEN can be assayed directly by ELISA that detects FVIII inhibitors, or measured indirectly by demonstration of reduced inhibition of FVIII activity of the CFXTEN compared to native FVIII in the presence of an inhibitor as measured by a factor VIII

chromogenic test or one-step assay as described herein, or other suitable coagulation methods known in the art. Alternatively, the subject CFXTEN can be measured for reduced (or absence of) inhibition in the presence of known inhibitors by use of a modified Bethesda assay. According to a particular aspect of the present invention, a CFXTEN useful in the methods has reduced reactivity to one or more antibodies from Table 10, as well as naturally-occurring antibodies found in hemophilia patients. For testing purposes, such and other inhibitory antibodies can be obtained from humans (i.e. from the serum of patients which have inhibitory antibodies) or can be obtained from mice, guinea pigs, horses, goats, non-human primates and other mammals by immunization with FVIII, or fragments thereof, more particularly with a fragment comprising the all or part of the A2 or C2 domain, whether in polyclonal or monoclonal form.

[00314] The invention further contemplates that the CFXTEN used in accordance with the methods provided herein can be administered in conjunction with other treatment methods and compositions (e.g., other coagulation proteins) useful for treating factor VIII-related conditions, or conditions for which coagulation factor is adjunctive therapy; e.g., bleeding episodes due to injury or surgery.

[00315] In another aspect, the invention provides methods of preparing a drug for a factor VIII-related condition, comprising combining a factor VIII sequence selected from Table 1 with one or more XTEN selected from Table 4 inserted in one or more insertion sites selected from Table 5, Table 6, Table 7, Table 8, and Table 9 to result in a drug that retains at least a portion of the activity of the native FVIII. The invention provides a method of preparing a pharmaceutical composition, comprising the step of combining the drug of the foregoing embodiment with at least one pharmaceutically acceptable carrier. In one embodiment of the method of preparing a drug for a factor VIII-related condition, the factor VIII has a sequence with at least about 80%, or at least about 90%, or at least about 95%, or at least about 97%, or at least about 99% sequence identity compared to a sequence selected from Table 1 and the one or more XTEN has a sequence with at least about 80%, or at least about 90%, or at least about 95%, or at least about 97%, or at least about 99% sequence identity compared to a sequence selected from any one of Tables 3, 4, and 13-17, or a fragment thereof, wherein the one or more XTEN are inserted in one or more locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9. In a particular embodiment of the foregoing, at least one XTEN insertion site is selected from amino acids 32, 220, 224, 336, 339, 390, 399, 416, 603, 1656, 1711, 1725, 1905 and 1910 (numbered relative to mature native human FVIII).In another embodiment of the method, the CFXTEN comprises a sequence with at least about 80%, or at least about 90%, or at least about 95%, or at least about 97%, or at least about 99% sequence identity compared to a sequence selected from any one of Table 21.

[00316] In another aspect, the invention provides a method of making the CFXTEN compositions to achieve desired pharmacokinetic, pharmacologic or pharmaceutical properties. In general, the steps in the design and production of the inventive fusion protein compositions, as illustrated in FIGS. 11-13, include: (1) the selection of a FVIII (e.g., native proteins, sequences of Table 1, analogs or derivatives with activity) to treat the particular condition; (2) selecting one or more XTEN (e.g., sequences with at least 80% identity to sequences set forth in Table 4) that will confer the desired pharmacokinetic and

physicochemical characteristics on the resulting CFXTEN (e.g., the administration of the CFXTEN composition to a subject results in the fusion protein being maintained above 0.05-0.4 IU/ml for a greater period compared to FVIII not linked to XTEN); (3) establishing a desired N- to C-terminus configuration of the CFXTEN to achieve the desired efficacy or PK parameters (e.g., selecting one or more insertion sites from Table 5, Table 6, Table 7, Table 8, and Table 9); (4) establishing the design of the expression vector encoding the configured CFXTEN; (5) transforming a suitable host with the expression vector; and (6) expressing and recovering the resultant isolated CFXTEN fusion protein. In one embodiment of the method of making CFXTEN, the XTEN for insertion are evaluated by the application of Equation IV to maximize the Ratio XTEN Radii for the fusion protein construct, with the XTEN resulting in values greater than 2.0, or 2.1, or 2.2, or 2.3, or 2.4, or 2.5, or 2.6, or 2.7, or 2.8, or 2.9. or 3.0 being preferred. For those CFXTEN for which an increase in half-life or an increased period of time spent above the minimum coagulation-effective concentration is desired, the XTEN chosen for incorporation generally have at least about 144, or about 288, or about 432, or about 576, or about 864, or about 875, or about 912, or about 923 amino acid residues where a single XTEN is to be incorporated into the CFXTEN. In another embodiment, the CFXTEN comprises a first XTEN of the foregoing lengths, and at least a second XTEN of about 36, or about 42, or about 72, or about 144, or about 288, or about 576, or about 864, or about 875, or about 912, or about 923, or about 1000 or more amino acid residues. The location of the XTEN within the fusion protein can include one, two, three, four, five or more locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9 or FIGS. 8-9. In one embodiment, the method of design includes an insertion of XTEN into the FVIII of at least one site selected from amino acids 32, 220, 224, 336, 339, 390, 399, 416, 603, 1656, 1711, 1725, 1905 and 1910 (numbered relative to mature native human FVIII).

[00317] In another aspect, the invention provides methods of making CFXTEN compositions to improve ease of manufacture, result in increased stability, increased water solubility, and/or ease of formulation, as compared to the native FVIII. In one embodiment, the invention includes a method of increasing the water solubility of a FVIII comprising the step of linking the FVIII with at least about 80%, or about 90%, or about 95% identity to a sequence from Table 1 to one or more XTEN at one, two, three, four, five or more locations selected from Table 5, Table 6, Table 7, Table 8, and Table 9 or FIGS. 8-9 wherein the XTEN is a sequence with at least about 80%, or about 90%, or about 95% sequence identity compared to a sequence from any one of Tables 3, 4, and 13-17 such that a higher concentration in soluble form of the resulting CFXTEN can be achieved, under physiologic conditions, compared to the FVIII in an un-fused state. In a particular embodiment, the CFXTEN comprises a FVIII linked to two, three, four, or five XTEN having at least about 24, or about 36, or about 48, or about 60, or about 72, or about 84, or about 96, or about 144, or about 288 amino acid residues inserted at sites selected from Table 5, Table 6, Table 7, Table 8, and Table 9 or FIGS. 8-9, in which the solubility of the fusion protein under physiologic conditions is at least three-fold greater than the corresponding FVIII not linked to XTEN, or alternatively, at least four-fold, or five-fold, or six-fold, or seven-fold, or eight-fold, or nine-fold, or at least 10-fold, or at least 20-fold, or at least 30-fold, or at least 50-fold, or at least 60-fold or

greater than FVIII not linked to XTEN. Factors that contribute to the property of XTEN to confer increased water solubility of CFs when incorporated into a fusion protein include the high solubility of the XTEN fusion partner and the low degree of self-aggregation between molecules of XTEN in solution, as well as expanding the hydrophilicity of FVIII external loops into which the XTEN is inserted. In some embodiments, the method results in a CFXTEN fusion protein wherein the water solubility is at least about 20%, or at least about 30% greater, or at least about 50% greater, or at least about 75% greater, or at least about 90% greater, or at least about 100% greater, or at least about 150% greater, or at least about 200% greater, or at least about 400% greater, or at least about 600% greater, or at least about 800% greater, or at least about 1000% greater, or at least about 2000% greater under physiologic conditions, compared to the un-fused FVIII. In one embodiment, the XTEN of the CFXTEN fusion protein is a sequence with at least about 80%, or about 90%, or about 95% sequence identity compared to a sequence from any one of Tables 3, 4, and 13-17. In another embodiment, the invention includes a method of increasing the shelf-life of a FVIII comprising the step of linking the FVIII with one or more XTEN at one or more sites selected from Table 5, Table 6, Table 7, Table 8, and Table 9, wherein the shelf-life of the resulting CFXTEN is extended compared to the FVIII in an un-fused state. As used herein, shelf-life refers to the period of time over which the procoagulant activity of a FVIII or CFXTEN that is in solution, lyophilized or in some other storage formulation remains stable without undue loss of activity or that remains within release specifications established for the pharmaceutical composition. A FVIII that degrades or aggregates generally has reduced functional activity or reduced bioavailability compared to one that remains in solution. Factors that contribute to the ability of the method to extend the shelf life of FVIII when incorporated into a fusion protein include increased water solubility, reduced self-aggregation in solution, and increased heat stability of the XTEN fusion partner. In particular, the low tendency of XTEN to aggregate facilitates methods of formulating pharmaceutical preparations containing higher drug concentrations of CFs, and the heat-stability of XTEN contributes to the property of CFXTEN fusion proteins to remain soluble and functionally active for extended periods. The method results in CFXTEN fusion proteins with prolonged or extended shelf-life that exhibit greater activity relative to a FVIII standard that has been subjected to the same storage and handling conditions. The standard may be the un-fused full-length FVIII or a commercially-available FVIII pharmaceutical composition. In one embodiment, the method includes the step of formulating the isolated CFXTEN with one or more pharmaceutically acceptable excipients that enhance the ability of the XTEN to retain its unstructured conformation and for the CFXTEN to remain soluble in the formulation for a time that is greater than that of the corresponding un-fused FVIII. In one embodiment, the method comprises linking a FVIII selected from Table 1 to one or more XTEN selected from any one of Tables 3, 4, and 13-17 inserted at one or more sites selected from Table 5, Table 6, Table 7, Table 8, and Table 9 and admixing with at least one pharmaceutically acceptable excipient to create a pharmaceutical composition that retains greater than about 100% of the procoagulant activity, or greater than about 105%, 110%, 120%, 130%, 150% or 200% of the procoagulant activity of a FVIII standard subjected to the same storage and handling conditions when compared at a time point of at least 90 days, or at least 6 months, or at least 12

months. Shelf-life may also be assessed in terms of functional activity remaining after storage, normalized to functional activity when storage began. In some embodiments, CFXTEN pharmaceutical compositions of the invention retain about 50% more procoagulant activity, or about 60%, 70%, 80%, or 90% more of the procoagulant activity of a FVIII standard when subjected to the same conditions for the same period of up to 2 weeks, or 4 weeks, or 6 weeks or longer under various temperature conditions. In one embodiment, the CFXTEN pharmaceutical composition retains at least about 50%, or about 60%, or at least about 70%, or at least about 80%, and most preferably at least about 90% or more of its original activity in solution when heated at 80°C for 10 min. In another embodiment, the CFXTEN pharmaceutical composition retains at least about 50%, preferably at least about 60%, or at least about 70%, or at least about 80%, or alternatively at least about 90% or more of its original activity in solution when heated or maintained at 37°C for about 7 days. In another embodiment, CFXTEN pharmaceutical composition retains at least about 80% or more of its functional activity after exposure to a temperature of about 30°C to about 70°C over a period of time of about one hour to about 18 hours. In the foregoing embodiments hereinabove described in this paragraph, the retained activity of the CFXTEN pharmaceutical compositions is at least about two-fold, or at least about three-fold, or at least about four-fold, or at least about five-fold, or at least about six-fold greater at a given time point than that of a corresponding pharmaceutical composition comprising FVIII not linked to the XTEN.

VII). THE NUCLEIC ACIDS SEQUENCES OF THE INVENTION

[00318] The present invention provides isolated polynucleic acids encoding CFXTEN chimeric fusion proteins and sequences complementary to polynucleic acid molecules encoding CFXTEN chimeric fusion proteins, including homologous variants thereof. In another aspect, the invention encompasses methods to produce polynucleic acids encoding CFXTEN chimeric fusion proteins and sequences complementary to polynucleic acid molecules encoding CFXTEN chimeric fusion protein, including homologous variants thereof. In general, and as illustrated in FIGS. 11-13, the methods of producing a polynucleotide sequence coding for a CFXTEN fusion protein and expressing the resulting gene product include assembling nucleotides encoding FVIII and XTEN, ligating the components in frame, incorporating the encoding gene into an expression vector appropriate for a host cell, transforming the appropriate host cell with the expression vector, and culturing the host cell under conditions causing or permitting the fusion protein to be expressed in the transformed host cell, thereby producing the biologically-active CFXTEN polypeptide, which is recovered as an isolated fusion protein by standard protein purification methods known in the art. Standard recombinant techniques in molecular biology is used to make the polynucleotides and expression vectors of the present invention.

[00319] In accordance with the invention, nucleic acid sequences that encode CFXTEN (or its complement) are used to generate recombinant DNA molecules that direct the expression of CFXTEN fusion proteins in appropriate host cells. For the purposes of the invention, nucleic acid encoding a signal peptide corresponding to that of native human FVIII (encoding MQIELSTCFFLCLLRFCFS (SEQ ID NO: 1611)) can be added to any of the encoding constructs described herein to aid in the expression

and secretion of the CFXTEN fusion protein. In one embodiment, the nucleic acid add is ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCGATTCTGCTTTAGT (SEQ ID NO: 1613), or the complement thereof.

[00320] Several cloning strategies are suitable for performing the present invention, many of which is used to generate a construct that comprises a gene coding for a fusion protein of the CFXTEN composition of the present invention, or its complement. In some embodiments, the cloning strategy is used to create a gene that encodes a monomeric CFXTEN that comprises at least a first FVIII and at least a first XTEN polypeptide, or their complement. In one embodiment of the foregoing, the gene comprises a sequence encoding a FVIII or sequence variant. In other embodiments, the cloning strategy is used to create a gene that encodes a monomeric CFXTEN that comprises nucleotides encoding at least a first molecule of FVIII or its complement and a first and at least a second XTEN or their complement that is used to transform a host cell for expression of the fusion protein of the CFXTEN composition. In the foregoing embodiments hereinabove described in this paragraph, the genes can further comprise nucleotides encoding spacer sequences that also encode cleavage sequence(s).

[00321] In designing a desired XTEN sequences, it was discovered that the non-repetitive nature of the XTEN of the inventive compositions is achieved despite use of a "building block" molecular approach in the creation of the XTEN-encoding sequences. This was achieved by the use of a library of polynucleotides encoding peptide sequence motifs, described above, that are then ligated and/or multimerized to create the genes encoding the XTEN sequences (see FIGS. 11 and 12 and Examples). Thus, while the XTEN(s) of the expressed fusion protein may consist of multiple units of as few as four different sequence motifs, because the motifs themselves consist of non-repetitive amino acid sequences, the overall XTEN sequence is rendered non-repetitive. Accordingly, in one embodiment, the XTEN-encoding polynucleotides comprise multiple polynucleotides that encode non-repetitive sequences, or motifs, operably linked in frame and in which the resulting expressed XTEN amino acid sequences are non-repetitive.

[00322] In one approach, a construct is first prepared containing the DNA sequence corresponding to CFXTEN fusion protein. DNA encoding the FVIII of the compositions is obtained synthetically, from a commercial source, or from a cDNA library prepared using standard methods from tissue or isolated cells believed to possess FVIII mRNA and to express it at a detectable level. If necessary, the coding sequence can be obtained using conventional primer extension procedures as described in Sambrook, *et al.*, *supra*, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA. One can then use polymerase chain reaction (PCR) methodology to amplify the target DNA or RNA coding sequence to obtain sufficient material for the preparation of the CFXTEN constructs containing the FVIII gene. Assays can then be conducted to confirm that the hybridizing full-length genes are the desired FVIII gene(s). By these conventional methods, DNA can be conveniently obtained from a cDNA library prepared from such sources. The FVIII encoding gene(s) can also created by standard synthetic procedures known in the art (e.g., automated nucleic acid synthesis using, for example one of the methods described in Engels et al. (Agnew. Chem. Int. Ed. Engl., 28:716-734 1989)),

using DNA sequences obtained from publicly available databases, patents, or literature references. Such procedures are well known in the art and well described in the scientific and patent literature. For example, sequences can be obtained from Chemical Abstracts Services (CAS) Registry Numbers (published by the American Chemical Society) and/or GenBank Accession Numbers (e.g., Locus ID, NP_XXXXX, and XP_XXXXX) Model Protein identifiers available through the National Center for Biotechnology Information (NCBI) webpage, available on the world wide web at ncbi.nlm.nih.gov that correspond to entries in the CAS Registry or GenBank database that contain an amino acid sequence of the protein of interest or of a fragment or variant of the protein. In one embodiment, the FVIII encoding gene encodes a protein sequence from Table 1, or a fragment or variant thereof.

[00323] A gene or polynucleotide encoding the FVIII portion of the subject CFXTEN protein, in the case of an expressed fusion protein that comprises a single FVIII, is then cloned into a construct, which is a plasmid or other vector under control of appropriate transcription and translation sequences for high level protein expression in a biological system. In a later step, a second gene or polynucleotide coding for the XTEN is genetically fused to the nucleotides encoding the N- and/or C-terminus of the FVIII gene by cloning it into the construct adjacent and in frame with the gene(s) coding for the FVIII. This second step occurs through a ligation or multimerization step. In the foregoing embodiments hereinabove described in this paragraph, it is to be understood that the gene constructs that are created can alternatively be the complement of the respective genes that encode the respective fusion proteins.

[00324] The gene encoding for the XTEN can be made in one or more steps, either fully synthetically or by synthesis combined with enzymatic processes, such as restriction enzyme-mediated cloning, PCR and overlap extension, including methods more fully described in the Examples. The methods disclosed herein can be used, for example, to ligate short sequences of polynucleotides encoding XTEN into longer XTEN genes of a desired length and sequence. In one embodiment, the method ligates two or more codon-optimized oligonucleotides encoding XTEN motif or segment sequences of about 9 to 14 amino acids, or about 12 to 20 amino acids, or about 18 to 42 amino acids, or about 42 to about 144 amino acids, or about 144 to about 288 amino acids, or 288 to about 864 amino acids or longer, or any combination of the foregoing ranges of motif or segment lengths.

[00325] Alternatively, the disclosed method is used to multimerize XTEN-encoding sequences into longer sequences of a desired length; e.g., a gene encoding 36 amino acids of XTEN can be dimerized into a gene encoding 72 amino acids, then 144, then 288, etc. Even with multimerization, XTEN polypeptides can be constructed such that the XTEN-encoding gene has low or virtually no repetitiveness through design of the codons selected for the motifs of the shortest unit being used, which can reduce recombination and increase stability of the encoding gene in the transformed host.

[00326] Genes encoding XTEN with non-repetitive sequences are assembled from oligonucleotides using standard techniques of gene synthesis. The gene design can be performed using algorithms that optimize codon usage and amino acid composition. In one method of the invention, a library of relatively short XTEN-encoding polynucleotide constructs is created and then assembled, as described above. The resulting genes are then assembled with genes encoding FVIII or regions of FVIII, as

illustrated in FIGS. 11 and 12, and the resulting genes used to transform a host cell and produce and recover the CFXTEN for evaluation of its properties, as described herein.

[00327] In another aspect, the invention provides isolated nucleic acids comprising a polynucleotide sequence encoding the CFXTEN fusion protein embodiments described herein. In one embodiment, the isolated nucleic acid comprises a polynucleotide sequence selected from (a) a sequence having at least about 80% sequence identity, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, or about 99%, to about 100% sequence identity compared to a sequence of comparable length selected from Table 21, when optimally aligned, or (b) the complement of the polynucleotide of (a). In another embodiment, the isolated nucleic acid comprises the sequence ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCGATTCTGCTTTAGT (SEQ ID NO: 1613) linked to the 5' end of the nucleic acid of (a) or the complement of the sequence linked to the 3' end of (b).

[00328] Polynucleotide libraries

[00329] In another aspect, the invention provides libraries of polynucleotides that encode XTEN sequences that are used to assemble genes that encode XTEN of a desired length and sequence.

[00330] In certain embodiments, the XTEN-encoding library constructs comprise polynucleotides that encode polypeptide segments of a fixed length. As an initial step, a library of oligonucleotides that encode motifs of 9-14 amino acid residues can be assembled. In a preferred embodiment, libraries of oligonucleotides that encode motifs of 12 amino acids are assembled.

[00331] The XTEN-encoding sequence segments can be dimerized or multimerized into longer encoding sequences, as depicted schematically in FIG. 13. Dimerization or multimerization can be performed by ligation, overlap extension, PCR assembly or similar cloning techniques known in the art. This process of can be repeated multiple times until the resulting XTEN-encoding sequences have reached the organization of sequence and desired length, providing the XTEN-encoding genes. As will be appreciated, a library of polynucleotides that encodes, e.g., 12 amino acid motifs can be dimerized and/or ligated into a library of polynucleotides that encode 36 amino acids. Libraries encoding motifs of different lengths; e.g., 9-14 amino acid motifs leading to libraries encoding 27 to 42 amino acids are contemplated by the invention. In turn, the library of polynucleotides that encode 27 to 42 amino acids, and preferably 36 amino acids (as described in the Examples) can be serially dimerized into a library containing successively longer lengths of polynucleotides that encode XTEN sequences of a desired length for incorporation into the gene encoding the CFXTEN fusion protein, as disclosed herein.

[00332] A more efficient way to optimize the DNA sequence encoding XTEN is based on combinatorial libraries. The gene encoding XTEN can be designed and synthesized in segment such that multiple codon versions are obtained for each segment. These segments can be randomly assembled into a library of genes such that each library member encodes the same amino acid sequences but library members comprise a large number of codon versions. Such libraries can be screened for genes that result in high-level expression and/or a low abundance of truncation products. The process of combinatorial

gene assembly is illustrated in FIG. 18. The genes in FIG. 18 are assembled from 6 base fragments and each fragment is available in 4 different codon versions. This allows for a theoretical diversity of 4096.

[00333] In some embodiments, libraries are assembled of polynucleotides that encode amino acids that are limited to specific sequence XTEN families; e.g., the AD, AE, AF, AG, AM, or AQ sequences of Table 4. In other embodiments, libraries comprise sequences that encode two or more of the motif family sequences from Table 3. The names and sequences of representative, non-limiting polynucleotide sequences of libraries that encode 36mers are presented in Tables 13-17, and the methods used to create them are described more fully in the respective Examples. In other embodiments, libraries that encode XTEN are constructed from segments of polynucleotide codons linked in a randomized sequence that encode amino acids wherein at least about 80%, or at least about 90%, or at least about 91%, or at least about 92%, or at least about 93%, or at least about 94%, or at least about 95%, or at least about 97%, or at least about 98%, or at least about 99% of the codons are selected from the group consisting of condons for glycine (G), alanine (A), serine (S), threonine (T), glutamate (E) and proline (P) amino acids. The libraries can be used, in turn, for serial dimerization or ligation to achieve polynucleotide sequence libraries that encode XTEN sequences, for example, of 42, 48, 72, 144, 288, 576, 864, 875, 912, 923, 1318 amino acids, or up to a total length of about 3000 amino acids, as well as intermediate lengths, in which the encoded XTEN can have one or more of the properties disclosed herein, when expressed as a component of a CFXTEN fusion protein. In some cases, the polynucleotide library sequences may also include additional bases used as "sequencing islands," described more fully below.

[00334] FIG. 14 is a schematic flowchart of representative, non-limiting steps in the assembly of a XTEN polynucleotide construct and a CFXTEN polynucleotide construct in the embodiments of the invention. Individual oligonucleotides **501** are annealed into sequence motifs **502** such as a 12 amino acid motif ("12-mer"), which is ligated to additional sequence motifs from a library to create a pool that encompasses the desired length of the XTEN **504**, as well as ligated to a smaller concentration of an oligo containing BbsI, and KpnI restriction sites **503.** The resulting pool of ligation products is gel-purified and the band with the desired length of XTEN is cut, resulting in an isolated XTEN gene with a stopper sequence **505**. The XTEN gene is cloned into a stuffer vector. In this case, the vector encodes an optional CBD sequence **506** and a GFP gene **508.** Digestion is than performed with BbsI/HindIII to remove **507** and **508** and place the stop codon. The resulting product is then cloned into a BsaI/HindIII digested vector containing a gene encoding the FVIII, resulting in the gene **500** encoding an FVIII-XTEN fusion protein.

[00335] One may clone the library of XTEN-encoding genes into one or more expression vectors known in the art. To facilitate the identification of well-expressing library members, one can construct the library as fusion to a reporter protein. Non-limiting examples of suitable reporter genes are green fluorescent protein, luciferace, alkaline phosphatase, and beta-galactosidase. By screening, one can identify short XTEN sequences that can be expressed in high concentration in the host organism of choice. Subsequently, one can generate a library of random XTEN dimers and repeat the screen for high

level of expression. Subsequently, one can screen the resulting constructs for a number of properties such as level of expression, protease stability, or binding to antiserum.

[00336] One aspect of the invention is to provide polynucleotide sequences encoding the components of the fusion protein wherein the creation of the sequence has undergone codon optimization. Of particular interest is codon optimization with the goal of improving expression of the polypeptide compositions and to improve the genetic stability of the encoding gene in the production hosts. For example, codon optimization is of particular importance for XTEN sequences that are rich in glycine or that have very repetitive amino acid sequences. Codon optimization is performed using computer programs (Gustafsson, C., et al. (2004) *Trends Biotechnol*, **22**: 346-53), some of which minimize ribosomal pausing (Coda Genomics Inc.). In one embodiment, one can perform codon optimization by constructing codon libraries where all members of the library encode the same amino acid sequence but where codon usage is varied. Such libraries can be screened for highly expressing and genetically stable members that are particularly suitable for the large-scale production of XTEN-containing products. When designing XTEN sequences one can consider a number of properties. One can minimize the repetitiveness in the encoding DNA sequences. In addition, one can avoid or minimize the use of codons that are rarely used by the production host (e.g. the AGG and AGA arginine codons and one leucine codon in *E. coli*). In the case of *E. coli*, two glycine codons, GGA and GGG, are rarely used in highly expressed proteins. Thus codon optimization of the gene encoding XTEN sequences can be very desirable. DNA sequences that have a high level of glycine tend to have a high GC content that can lead to instability or low expression levels. Thus, when possible, it is preferred to choose codons such that the GC-content of XTEN-encoding sequence is suitable for the production organism that will be used to manufacture the XTEN.

[00337] In one embodiment, polynucleotide libraries are constructed using the disclosed methods wherein all members of the library encode the same amino acid sequence but where codon usage for the respective amino acids in the sequence is varied or optimized for the intended host cell. Such libraries can be screened for highly expressing and genetically stable members that are particularly suitable for the large-scale production of XTEN-containing products. In one embodiment, the libraries are optimized for expression in a eukaryotic host cell.

[00338] Optionally, one can sequence clones in the library to eliminate isolates that contain undesirable sequences. The initial library of short XTEN sequences allows some variation in amino acid sequence. For instance one can randomize some codons such that a number of hydrophilic amino acids can occur in a particular position. During the process of iterative multimerization one can screen the resulting library members for other characteristics like solubility or protease resistance in addition to a screen for high-level expression.

[00339] Once the gene that encodes the XTEN of desired length and properties is selected, it is genetically fused at the desired location to the nucleotides encoding the FVIII gene(s) by cloning it into the construct adjacent and in frame with the gene coding for FVIII, or alternatively between nucleotides encoding adjacent domains of the FVIII, or alternatively within a sequence encoding a given FVIII domain, or alternatively in frame with nucleotides encoding a spacer/cleavage sequence linked to a

terminal XTEN. The invention provides various permutations of the foregoing, depending on the CFXTEN to be encoded. For example, a gene encoding a CFXTEN fusion protein comprising a FVIII and two XTEN, such as embodied by formula VI, as depicted above, the gene would have polynucleotides encoding FVIII, encoding two XTEN, which can be identical or different in composition and sequence length. In one non-limiting embodiment of the foregoing, the FVIII polynucleotides would encode factor VIII and the polynucleotides encoding the C-terminus XTEN would encode an XTEN of 288 amino acids and the polynucleotides encoding an internal XTEN adjacent to the C-terminus of the A2 domain would encode an XTEN of 144 amino acids. The step of cloning the FVIII genes into the XTEN construct can occur through a ligation or multimerization step, as shown in FIG. 14. The constructs encoding CFXTEN fusion proteins can be designed in different configurations of the components XTEN, CF, and spacer sequences, such as the configurations of formulae I-VIII. In one embodiment, the construct comprises polynucleotide sequences complementary to, or those that encode a monomeric polypeptide of components in the following order (5' to 3') FVIII, an XTEN internal to the B domain, and a C-terminal XTEN. In another embodiment, the construct comprises polynucleotide sequences complementary to, or those that encode a monomeric polypeptide of components in the following order (5' to 3') FVIIII, spacer sequence linked to the C-terminus, and XTEN. The spacer polynucleotides can optionally comprise sequences encoding cleavage sequences. As will be apparent to those of skill in the art, multiple permutations of FVIII domains and inserted XTEN are possible.

[00340] Homology, sequence similarity or sequence identity of nucleotide or amino acid sequences may also be determined conventionally by using known software or computer programs such as the BestFit or Gap pairwise comparison programs (GCG Wisconsin Package, Genetics Computer Group, 575 Science Drive, Madison, Wis. 53711). BestFit uses the local homology algorithm of Smith and Waterman (Advances in Applied Mathematics. 1981. 2: 482-489), to find the best segment of identity or similarity between two sequences. Gap performs global alignments: all of one sequence with all of another similar sequence using the method of Needleman and Wunsch, (Journal of Molecular Biology. 1970. 48:443-453). When using a sequence alignment program such as BestFit, to determine the degree of sequence homology, similarity or identity, the default setting may be used, or an appropriate scoring matrix may be selected to optimize identity, similarity or homology scores.

[00341] Nucleic acid sequences that are "complementary" are those that are capable of base-pairing according to the standard Watson-Crick complementarity rules. As used herein, the term "complementary sequences" means nucleic acid sequences that are substantially complementary, as may be assessed by the same nucleotide comparison set forth above, or as defined as being capable of hybridizing to the polynucleotides that encode the CFXTEN sequences under stringent conditions, such as those described herein.

[00342] The resulting polynucleotides encoding the CFXTEN chimeric fusion proteins can then be individually cloned into an expression vector. The nucleic acid sequence is inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or

more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan. Such techniques are well known in the art and well described in the scientific and patent literature.

[00343] Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage that may conveniently be subjected to recombinant DNA procedures, and the choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e., a vector, which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated. Representative plasmids are illustrated in FIG. 17, with encoding regions for different configurations of FVIII and XTEN components portrayed.

[00344] The invention provides for the use of plasmid vectors containing replication and control sequences that are compatible with and recognized by the host cell, and are operably linked to the CFXTEN gene for controlled expression of the CFXTEN fusion proteins. The vector ordinarily carries a replication site, as well as sequences that encode proteins that are capable of providing phenotypic selection in transformed cells. Such vector sequences are well known for a variety of bacteria, yeast, and viruses. Useful expression vectors that can be used include, for example, segments of chromosomal, non-chromosomal and synthetic DNA sequences. "Expression vector" refers to a DNA construct containing a DNA sequence that is operably linked to a suitable control sequence capable of effecting the expression of the DNA encoding the fusion protein in a suitable host. The requirements are that the vectors are replicable and viable in the host cell of choice. Low- or high-copy number vectors may be used as desired.

[00345] Other suitable vectors include, but are not limited to, derivatives of SV40 and pcDNA and known bacterial plasmids such as col EI, pCRl, pBR322, pMal-C2, pET, pGEX as described by Smith, et al., Gene 57:31-40 (1988), pMB9 and derivatives thereof, plasmids such as RP4, phage DNAs such as the numerous derivatives of phage I such as NM98 9, as well as other phage DNA such as M13 and filamentous single stranded phage DNA; yeast plasmids such as the 2 micron plasmid or derivatives of the 2m plasmid, as well as centomeric and integrative yeast shuttle vectors; vectors useful in eukaryotic cells such as vectors useful in insect or mammalian cells; vectors derived from combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or the expression control sequences; and the like. Yeast expression systems that can also be used in the present invention include, but are not limited to, the non-fusion pYES2 vector (Invitrogen), the fusion pYESHisA, B, C (Invitrogen), pRS vectors and the like.

[00346] The control sequences of the vector include a promoter to effect transcription, an optional operator sequence to control such transcription, a sequence encoding suitable mRNA ribosome binding sites, and sequences that control termination of transcription and translation. The promoter may be any

DNA sequence, which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell.

[00347] Examples of suitable promoters for directing the transcription of the DNA encoding the FVIII polypeptide variant in mammalian cells are the SV40 promoter (Subramani et al., *Mol. Cell. Biol.* 1 (1981), 854-864), the MT-1 (metallothionein gene) promoter (Palmiter et al., *Science* 222 (1983), 809-814), the CMV promoter (Boshart et al., *Cell* 41:521-530, 1985) or the adenovirus 2 major late promoter (Kaufman and Sharp, *Mol. Cell. Biol,* 2:1304-1319, 1982). The vector may also carry sequences such as UCOE (ubiquitous chromatin opening elements).

[00348] Examples of suitable promoters for use in filamentous fungus host cells are, for instance, the ADH3 promoter or the tpiA promoter. Examples of other useful promoters are those derived from the gene encoding *A. oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *A. niger* neutral α-amylase, *A. niger* acid stable α-amylase, *A. niger* or *A. awamoriglucoamylase* (gluA), *Rhizomucor miehei* lipase, *A. oryzae* alkaline protease, *A. oryzae* triose phosphate isomerase or *A. nidulans* acetamidase. Preferred are the TAKA-amylase and gluA promoters.

[00349] Promoters suitable for use in expression vectors with prokaryotic hosts include the β-lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)], all is operably linked to the DNA encoding CFXTEN polypeptides. Promoters for use in bacterial systems can also contain a Shine-Dalgarno (S.D.) sequence, operably linked to the DNA encoding CFXTEN polypeptides.

[00350] The invention contemplates use of other expression systems including, for example, a baculovirus expression system with both non-fusion transfer vectors, such as, but not limited to pVL941 Summers, et al., Virology 84:390-402 (1978)), pVL1393 (Invitrogen), pVL1392 (Summers, et al., Virology 84:390- 402 (1978) and Invitrogen) and pBlueBacIII (Invitrogen), and fusion transfer vectors such as, but not limited to, pAc7 00 (Summers, et al., Virology 84:390-402 (1978)), pAc701 and pAc70-2 (same as pAc700, with different reading frames), pAc360 Invitrogen) and pBlueBacHisA, B, C (; Invitrogen) can be used.

[00351] Examples of suitable promoters for directing the transcription of the DNA encoding the FVIII polypeptide variant in mammalian cells are the CMV promoter (Boshart et al., *Cell* 41:521-530, 1985), the SV40 promoter (Subramani et al., *Mol. Cell. Biol.* 1 (1981), 854-864), the MT-1 (metallothionein gene) promoter (Palmiter et al., *Science* 222 (1983), 809-814), the adenovirus 2 major late promoter (Kaufman and Sharp, *Mol. Cell. Biol,* 2:1304-1319, 1982). The vector may also carry sequences such as UCOE (ubiquitous chromatin opening elements).

[00352] The DNA sequences encoding the CFXTEN may also, if necessary, be operably connected to a suitable terminator, such as the hGH terminator (Palmiter et al., *Science* 222, 1983, pp. 809-814) or the TPI1 terminators (Alber and Kawasaki, *J. Mol. Appl. Gen.* 1, 1982, pp. 419-434) or ADH3 (McKnight et al., *The EMBO J.* 4, 1985, pp. 2093-2099). Expression vectors may also contain a set of RNA splice sites

located downstream from the promoter and upstream from the insertion site for the CFXTEN sequence itself, including splice sites obtained from adenovirus. Also contained in the expression vectors is a polyadenylation signal located downstream of the insertion site. Particularly preferred polyadenylation signals include the early or late polyadenylation signal from SV40 (Kaufman and Sharp, ibid.), the polyadenylation signal from the adenovirus 5 Elb region, the hGH terminator (DeNoto et al. Nucl. Acids Res. 9:3719-3730, 1981). The expression vectors may also include a noncoding viral leader sequence, such as the adenovirus 2 tripartite leader, located between the promoter and the RNA splice sites; and enhancer sequences, such as the SV40 enhancer.

[00353] To direct the CFXTEN of the present invention into the secretory pathway of the host cells, a secretory signal sequence (a.k.a., a leader sequence, a prepro sequence, or a pre sequence) may be included in the recombinant vector. The secretory signal sequence is operably linked to the DNA sequences encoding the CFXTEN, usually positioned 5′ to the DNA sequence encoding the CFXTEN fusion protein. The secretory signal sequence may be that, normally associated with the native FVIII protein or may be from a gene encoding another secreted protein. Non-limiting examples include OmpA, PhoA, and DsbA for *E. coli* expression, ppL-alpha, DEX4, invertase signal peptide, acid phosphatase signal peptide, CPY, or INU1 for yeast expression, and IL2L, SV40, IgG kappa and IgG lambda for mammalian expression. Signal sequences are typically proteolytically removed from the protein during the translocation and secretion process, generating a defined N terminus. Methods are disclosed in Arnau, *et al.*, Protein Expression and Purification 48: 1-13 (2006).

[00354] The procedures used to ligate the DNA sequences coding for the CFXTEN, the promoter and optionally the terminator and/or secretory signal sequence, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook, J. *et al.*, "Molecular Cloning: A Laboratory Manual," 3rd edition, Cold Spring Harbor Laboratory Press, 2001). In this manner, a chimeric DNA molecule coding for a monomeric CFXTEN fusion protein is generated within the construct. Optionally, this chimeric DNA molecule may be transferred or cloned into another construct that is a more appropriate expression vector. At this point, a host cell capable of expressing the chimeric DNA molecule can be transformed with the chimeric DNA molecule.

[00355] Non-limiting examples of mammalian cell lines for use in the present invention are the COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), BHK-21 (ATCC CCL 10)) and BHK-293 (ATCC CRL 1573; Graham et al., *J. Gen. Virol.* 36:59-72, 1977), BHK-570 cells (ATCC CRL 10314), CHO-K1 (ATCC CCL 61), CHO-S (Invitrogen 11619-012), and 293-F (Invitrogen R790-7), and the parental and derivative cell lines known in the art useful for expression of FVIII. A tk-ts13 BHK cell line is also available from the ATCC under accession number CRL 1632. In addition, a number of other cell lines may be used within the present invention, including Rat Hep I (Rat hepatoma; ATCC CRL 1600), Rat Hep II (Rat hepatoma; ATCC CRL 1548), TCMK (ATCC CCL 139), Human lung (ATCC HB 8065), NCTC 1469 (ATCC CCL 9.1), CHO (ATCC CCL 61) and DUKX cells (Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA* 77:4216-4220, 1980).

[00356] Examples of suitable yeasts cells include cells of *Saccharomyces* spp. or *Schizosaccharomyces* spp., in particular strains of *Saccharomyces cerevisiae* or *Saccharomyces kluyveri* . Methods for transforming yeast cells with heterologous DNA and producing heterologous polypeptides there from are described, e.g. in U.S. Pat. No. 4,599,311, U.S. Pat. No. 4,931,373, U.S. Pat. No. 4,870,008, 5,037,743, and U.S. Pat. No. 4,845,075. Transformed cells are selected by a phenotype determined by a selectable marker, commonly drug resistance or the ability to grow in the absence of a particular nutrient, e.g. leucine. A preferred vector for use in yeast is the POT1 vector disclosed in U.S. Pat. No. 4,931,373. The DNA sequences encoding the CFXTEN may be preceded by a signal sequence and optionally a leader sequence, e.g. as described above. Further examples of suitable yeast cells are strains of *Kluyveromyces*, such as *K. lactis, Hansenula* , e.g. *H. polymorpha* , or *Pichia* , e.g. *P. pastoris* (cf. Gleeson et al., *J. Gen. Microbiol.* 132, 1986, pp. 3459-3465; U.S. Pat. No. 4,882,279). Examples of other fungal cells are cells of filamentous fungi, e.g. *Aspergillus* spp., *Neurospora* spp., *Fusarium* spp. or *Trichoderma* spp., in particular strains of *A. oryzae, A. nidulans* or *A. niger* . The use of *Aspergillus* spp. for the expression of proteins is described in, e.g., EP 272 277, EP 238 023, EP 184 438. The transformation of *F. oxysporum* may, for instance, be carried out as described by Malardier et al., 1989 , *Gene* 78: 147-156. The transformation of *Trichoderma* spp. may be performed for instance as described in EP 244 234.

[00357] Other suitable cells that can be used in the present invention include, but are not limited to, prokaryotic host cells strains such as *Escherichia coli, (e.g., strain DH5-α), Bacillus subtilis, Salmonella typhimurium,* or strains of the genera of *Pseudomonas, Streptomyces and Staphylococcus.* Non-limiting examples of suitable prokaryotes include those from the genera: *Actinoplanes; Archaeoglobus; Bdellovibrio; Borrelia; Chloroflexus; Enterococcus; Escherichia; Lactobacillus; Listeria; Oceanobacillus; Paracoccus; Pseudomonas; Staphylococcus; Streptococcus; Streptomyces; Thermoplasma; and Vibrio.*

[00358] Methods of transfecting mammalian cells and expressing DNA sequences introduced in the cells are described in e.g., Kaufman and Sharp, *J. Mol. Biol.* 159 (1982), 601-621; Southern and Berg, *J. Mol. Appl. Genet.* 1 (1982), 327-341; Loyter et al., *Proc. Natl. Acad. Sci. USA* 79 (1982), 422-426; Wigler et al., *Cell* 14 (1978), 725; Corsaro and Pearson, *Somatic Cell Genetics* 7 (1981), 603, Graham and van der Eb, *Virology* 52 (1973), 456; and Neumann et al., *EMBO J.* 1 (1982), 841-845.

[00359] Cloned DNA sequences are introduced into cultured mammalian cells by, for example, calcium phosphate-mediated transfection (Wigler et al., Cell 14:725-732, 1978; Corsaro and Pearson, Somatic Cell Genetics 7:603-616, 1981; Graham and Van der Eb, Virology 52d:456-467, 1973), transfection with many commercially available reagents such as FuGENEG Roche Diagnostics, Mannheim, Germany) or lipofectamine (Invitrogen) or by electroporation (Neumann et al., EMBO J. 1:841-845, 1982). To identify and select cells that express the exogenous DNA, a gene that confers a selectable phenotype (a selectable marker) is generally introduced into cells along with the gene or cDNA of interest. Preferred selectable markers include genes that confer resistance to drugs such as neomycin, hygromycin, puromycin, zeocin, and methotrexate. The selectable marker may be an amplifiable selectable marker. A preferred amplifiable selectable marker is a dihydrofolate reductase (DHFR) sequence. Further examples of selectable markers are

well known to one of skill in the art and include reporters such as enhanced green fluorescent protein (EGFP), beta-galactosidase (β-gal) or chloramphenicol acetyltransferase (CAT). Selectable markers are reviewed by Thilly (Mammalian Cell Technology, Butterworth Publishers, Stoneham, Mass.). The person skilled in the art will easily be able to choose suitable selectable markers. Any known selectable marker may be employed so long as it is capable of being expressed simultaneously with the nucleic acid encoding a gene product.

[00360] Selectable markers may be introduced into the cell on a separate plasmid at the same time as the gene of interest, or they may be introduced on the same plasmid. If, on the same plasmid, the selectable marker and the gene of interest may be under the control of different promoters or the same promoter, the latter arrangement producing a dicistronic message. Constructs of this type are known in the art (for example, Levinson and Simonsen, U.S. Pat. No. 4,713,339). It may also be advantageous to add additional DNA, known as "carrier DNA," to the mixture that is introduced into the cells.

[00361] After the cells have taken up the DNA, they are grown in an appropriate growth medium, typically 1-2 days, to begin expressing the gene of interest. As used herein the term "appropriate growth medium" means a medium containing nutrients and other components required for the growth of cells and the expression of the CFXTEN of interest. Media generally include a carbon source, a nitrogen source, essential amino acids, essential sugars, vitamins, salts, phospholipids, protein and growth factors. For production of gamma-carboxylated proteins, the medium will contain vitamin K, preferably at a concentration of about 0.1 μg/ml to about 5 μg/ml. Drug selection is then applied to select for the growth of cells that are expressing the selectable marker in a stable fashion. For cells that have been transfected with an amplifiable selectable marker the drug concentration may be increased to select for an increased copy number of the cloned sequences, thereby increasing expression levels. Clones of stably transfected cells are then screened for expression of the FVIII polypeptide variant of interest.

[00362] The transformed or transfected host cell is then cultured in a suitable nutrient medium under conditions permitting expression of the CFXTEN polypeptide after which the resulting peptide may be recovered from the culture as an isolated fusion protein. The medium used to culture the cells may be any conventional medium suitable for growing the host cells, such as minimal or complex media containing appropriate supplements. Suitable media are available from commercial suppliers or may be prepared according to published recipes (e.g. in catalogues of the American Type Culture Collection). The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

[00363] Gene expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or in situ hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried

out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

[00364] Gene expression, alternatively, may be measured by immunological of fluorescent methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids or the detection of selectable markers, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence FVIII polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to FVIII and encoding a specific antibody epitope. Examples of selectable markers are well known to one of skill in the art and include reporters such as enhanced green fluorescent protein (EGFP), beta-galactosidase (β-gal) or chloramphenicol acetyltransferase (CAT).

[00365] Expressed CFXTEN polypeptide product(s) may be purified via methods known in the art or by methods disclosed herein. Procedures such as gel filtration, affinity purification (e.g., using an anti-FVIII antibody column), salt fractionation, ion exchange chromatography, size exclusion chromatography, hydroxyapatite adsorption chromatography, hydrophobic interaction chromatography and gel electrophoresis may be used; each tailored to recover and purify the fusion protein produced by the respective host cells. Additional purification may be achieved by conventional chemical purification means, such as high performance liquid chromatography. Some expressed CFXTEN may require refolding during isolation and purification. Methods of purification are described in Robert K. Scopes, Protein Purification: Principles and Practice, Charles R. Castor (ed.), Springer-Verlag 1994, and Sambrook, *et al., supra.* Multi-step purification separations are also described in Baron, *et al.*, Crit. Rev. Biotechnol. 10:179-90 (1990) and Below, *et al.*, J. Chromatogr. A. 679:67-83 (1994). For therapeutic purposes it is preferred that the CFXTEN fusion proteins of the invention are substantially pure. Thus, in a preferred embodiment of the invention the CFXTEN of the invention is purified to at least about 90 to 95% homogeneity, preferably to at least about 98% homogeneity. Purity may be assessed by, e.g., gel electrophoresis, HPLC, and amino-terminal amino acid sequencing..

VIII). PHARMACEUTICAL COMPOSITIONS

[00366] The present invention provides pharmaceutical compositions comprising CFXTEN. In one embodiment, the pharmaceutical composition comprises a CFXTEN fusion protein disclosed herein admixed with at least one pharmaceutically acceptable carrier. CFXTEN polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the polypeptide is combined in admixture with a pharmaceutically acceptable carrier vehicle, such as aqueous solutions, buffers, solvents and/or pharmaceutically acceptable suspensions, emulsions, stabilizers or excipients. Examples of non-aqueous solvents include propyl ethylene glycol, polyethylene glycol and vegetable oils. Formulations of the pharmaceutical compositions are prepared for storage by mixing the active CFXTEN ingredient having the desired

degree of purity with optional physiologically acceptable carriers, excipients (e.g., sodium chloride, a calcium salt, sucrose, or polysorbate) or stabilizers (e.g., sucrose, trehalose, raffinose, arginine, a calcium salt, glycine or histidine), as described in Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980), in the form of lyophilized formulations or aqueous solutions.

[00367] The pharmaceutical composition may be supplied as a lyophilized powder to be reconstituted prior to administration. In another embodiment, the pharmaceutical composition may be supplied in a liquid form in a vial, the contents of which can be administered directly to a patient. Alternatively, the composition is supplied as a liquid in a pre-filled syringe for administration of the composition. In another embodiment, the composition is supplied as a liquid in a pre-filled vial that can be incorporated into a pump.

[00368] The pharmaceutical compositions can be administered by any suitable means or route, including subcutaneously, subcutaneously by infusion pump, intramuscularly, and intravenously. It will be appreciated that the preferred route will vary with the disease and age of the recipient, and the severity of the condition being treated.

[00369] In one embodiment, the CFXTEN pharmaceutical composition in liquid form or after reconstitution (when supplied as a lyophilized powder) comprises coagulation factor VIII with an activity of at least 50 IU/ml, or at least 100 IU/ml, or at least 200 IU/ml, or at least 300 IU/ml, or at least 400 IU/ml, or an activity of at least 500 IU/ml, or an activity of at least 600 IU/ml, which composition is capable of increasing factor VIII activity to at least 1.5% of the normal plasma level in the blood for at least about 12 hours, or at least about 24 hours, or at least about 48 hours, or at least about 72 hours, or at least about 96 hours, or at least about 120 hours after administration of the factor VIII pharmaceutical composition to a subject in need of routine prophylaxis. In another embodiment, the CFXTEN pharmaceutical composition in liquid form or after reconstitution (when supplied as a lyophilized powder) comprises coagulation factor VIII with an activity of at least 50 IU/ml, or at least 100 IU/ml, or at least 200 IU/ml, or at least 300 IU/ml, or at least 400 IU/ml, or at least 500 IU/ml, or an activity of at least 600 IU/ml, which composition is capable of increasing factor VIII activity to at least 2.5% of the normal plasma level in the blood for at least about 12 hours, or at least about 24 hours, or at least about 48 hours, or at least about 72 hours, or at least about 96 hours, or at least about 120 hours after administration to a subject in need of routine prophylaxis. It is specifically contemplated that the pharmaceutical compositions of the foregoing can be formulated to include one or more excipients, buffers or other ingredients known in the art to be compatible with administration by the intravenous route or the subcutaneous route or the intramuscular route. Thus, in the embodiments hereinabove described in this paragraph, the pharmaceutical composition is administered subcutaneously, intramuscularly or intravenously.

[00370] The compositions of the invention may be formulated using a variety of excipients. Suitable excipients include microcrystalline cellulose (e.g. Avicel™ PH102, Avicel™ PH101), polymethacrylate, poly(ethyl acrylate, methyl methacrylate, trimethylammonioethyl methacrylate chloride) (such as Eudragit

RS-30D), hydroxypropyl methylcellulose (Methocel K100M, Premium CR Methocel K100M, Methocel E5, Opadry®), magnesium stearate, talc, triethyl citrate, aqueous ethylcellulose dispersion (Surelease®), and protamine sulfate. The slow release agent may also comprise a carrier, which can comprise, for example, solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents. Pharmaceutically acceptable salts can also be used in these slow release agents, for example, mineral salts such as hydrochlorides, hydrobromides, phosphates, or sulfates, as well as the salts of organic acids such as acetates, proprionates, malonates, or benzoates. The composition may also contain liquids, such as water, saline, glycerol, and ethanol, as well as substances such as wetting agents, emulsifying agents, or pH buffering agents. Liposomes may also be used as a carrier.

[00371] In another embodiment, the compositions of the present invention are encapsulated in liposomes, which have demonstrated utility in delivering beneficial active agents in a controlled manner over prolonged periods of time. Liposomes are closed bilayer membranes containing an entrapped aqueous volume. Liposomes may also be unilamellar vesicles possessing a single membrane bilayer or multilamellar vesicles with multiple membrane bilayers, each separated from the next by an aqueous layer. The structure of the resulting membrane bilayer is such that the hydrophobic (non-polar) tails of the lipid are oriented toward the center of the bilayer while the hydrophilic (polar) heads orient towards the aqueous phase. In one embodiment, the liposome may be coated with a flexible water soluble polymer that avoids uptake by the organs of the mononuclear phagocyte system, primarily the liver and spleen. Suitable hydrophilic polymers for surrounding the liposomes include, without limitation, PEG, polyvinylpyrrolidone, polyvinylmethylether, polymethyloxazoline, polyethyloxazoline, polyhydroxypropyloxazoline, polyhydroxypropylmethacrylamide, polymethacrylamide, polydimethylacrylamide, polyhydroxypropylmethacrylate, polyhydroxethylacrylate, hydroxymethylcellulose hydroxyethylcellulose, polyethyleneglycol, polyaspartamide and hydrophilic peptide sequences as described in U.S. Pat. Nos. 6,316,024; 6,126,966; 6,056,973; 6,043,094. Additional liposomal technologies are described in U.S. Pat. Nos. 6,759,057; 6,406,713; 6,352,716; 6,316,024; 6,294,191; 6,126,966; 6,056,973; 6,043,094; 5,965,156; 5,916,588; 5,874,104; 5,215,680; and 4,684,479. These describe liposomes and lipid-coated microbubbles, and methods for their manufacture. Thus, one skilled in the art, considering both the disclosure of this invention and the disclosures of these other patents could produce a liposome for the extended release of the polypeptides of the present invention.

[00372] For liquid formulations, a desired property is that the formulation be supplied in a form that can pass through a 25, 28, 30, 31, 32 gauge needle for intravenous, intramuscular, intraarticular, or subcutaneous administration.

[00373] Syringe pumps may also be used as slow release agents. Such devices are described in U.S. Pat. Nos. 4,976,696; 4,933,185; 5,017,378; 6,309,370; 6,254,573; 4,435,173; 4,398,908; 6,572,585; 5,298,022; 5,176,502; 5,492,534; 5,318,540; and 4,988,337. One skilled in the art, considering both the disclosure of this invention and the

disclosures of these other patents could produce a syringe pump for the extended release of the compositions of the present invention.

IX). PHARMACEUTICAL KITS

[00374] In another aspect, the invention provides a kit to facilitate the use of the CFXTEN polypeptides. The kit comprises the pharmaceutical composition provided herein, a container and a label or package insert on or associated with the container. Suitable containers include, for example, bottles, vials, syringes, etc., formed from a variety of materials such as glass or plastic. The container holds a pharmaceutical composition as a formulation that is effective for treating the FVIII-related condition and may have a sterile access port (for example the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). The package insert can list the approved indications for the drug, instructions for the reconstitution and/or administration of the drug for the use for the approved indication, appropriate dosage and safety information, and information identifying the lot and expiration of the drug. In another embodiment of the foregoing, the kit can comprise a second container that can carry a suitable diluent for the pharmaceutical composition, the use of which will provide the user with the appropriate concentration to be delivered to the subject.

EXAMPLES

[00375] Example 1: Construction of XTEN_AD36 motif segments

[00376] The following example describes the construction of a collection of codon-optimized genes encoding motif sequences of 36 amino acids. As a first step, a stuffer vector pCW0359 was constructed based on a pET vector and that includes a T7 promoter. pCW0359 encodes a cellulose binding domain (CBD) and a TEV protease recognition site followed by a stuffer sequence that is flanked by BsaI, BbsI, and KpnI sites. The BsaI and BbsI sites were inserted such that they generate compatible overhangs after digestion. The stuffer sequence is followed by a truncated version of the GFP gene and a His tag. The stuffer sequence contains stop codons and thus E. coli cells carrying the stuffer plasmid pCW0359 form non-fluorescent colonies. The stuffer vector pCW0359 was digested with BsaI and KpnI to remove the stuffer segment and the resulting vector fragment was isolated by agarose gel purification. The sequences were designated XTEN_AD36, reflecting the AD family of motifs. Its segments have the amino acid sequence $[X]_3$ where X is a 12mer peptide with the sequences: GESPGGSSGSES (SEQ ID NO: 19), GSEGSSGPGESS (SEQ ID NO: 20), GSSESGSSEGGP (SEQ ID NO: 21), or GSGGEPSESGSS (SEQ ID NO: 22). The insert was obtained by annealing the following pairs of phosphorylated synthetic oligonucleotide pairs:

AD1for: AGGTGAATCTCCDGGTGGYTCYAGCGGTTCYGARTC (SEQ ID NO: 1619)
AD1rev: ACCTGAYTCRGAACCGCTRGARCCACCHGGAGATTC (SEQ ID NO: 1620)
AD2for: AGGTAGCGAAGGTTCTTCYGGTCCDGGYGARTCYTC (SEQ ID NO: 1621)
AD2rev: ACCTGARGAYTCRCCHGGACCRGAAGAACCTTCGCT (SEQ ID NO: 1622)
AD3for: AGGTTCYTCYGAAAGCGGTTCTTCYGARGGYGGTCC (SEQ ID NO: 1623)

AD3rev: ACCTGGACCRCCYTCRGAAGAACCGCTTTCRGARGA (SEQ ID NO: 1624)

AD4for: AGGTTCYGGTGGYGAACCDTCYGARTCTGGTAGCTC (SEQ ID NO: 1625)

[00377] We also annealed the phosphorylated oligonucleotide 3KpnIstopperFor: AGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 1626) and the non-phosphorylated oligonucleotide pr_3KpnIstopperRev: CCTCGAGTGAAGACGA (SEQ ID NO: 1627). The annealed oligonucleotide pairs were ligated, which resulted in a mixture of products with varying length that represents the varying number of 12mer repeats ligated to one BbsI/KpnI segment. The products corresponding to the length of 36 amino acids were isolated from the mixture by preparative agarose gel electrophoresis and ligated into the BsaI/KpnI digested stuffer vector pCW0359. Most of the clones in the resulting library designated LCW0401 showed green fluorescence after induction, which shows that the sequence of XTEN_AD36 had been ligated in frame with the GFP gene and that most sequences of XTEN_AD36 had good expression levels.

[00378] We screened 96 isolates from library LCW0401 for high level of fluorescence by stamping them onto agar plate containing IPTG. The same isolates were evaluated by PCR and 48 isolates were identified that contained segments with 36 amino acids as well as strong fluorescence. These isolates were sequenced and 39 clones were identified that contained correct XTEN_AD36 segments. The file names of the nucleotide and amino acid constructs and the sequences for these segments are listed in Table 13.

Table 13: DNA and Amino Acid Sequences for AD 36-mer motifs (SEQ ID NOS 203-278, respectively, in order of appearance)

File name	Amino acid sequence	Nucleotide sequence
LCW0401_001_ GFP-N_A01.ab1	GSGGEPSESGSSGESPGG SSGSESGESPGGSSGSES	GGTTCTGGTGGCGAACCGTCCGAGTCTGGTAGCTCA GGTGAATCTCCGGGTGGCTCTAGCGGTTCCGAGTCA GGTGAATCTCCTGGTGGTTCCAGCGGTTCCGAGTCA
LCW0401_002_ GFP-N_B01.ab1	GSEGSSGPGESSGESPGG SSGSESGSSESGSSEGGP	GGTAGCGAAGGTTCTTCTGGTCCTGGCGAGTCTTCA GGTGAATCTCCTGGTGGTTCCAGCGGTTCTGAATCA GGTTCCTCCGAAAGCGGTTCTTCCGAGGGCGGTCCA
LCW0401_003_ GFP-N_C01.ab1	GSSESGSSEGGPGSSESG SSEGGPGESPGGSSGSES	GGTTCCTCTGAAAGCGGTTCTTCCGAAGGTGGTCCA GGTTCCTCTGAAAGCGGTTCTTCTGAGGGTGGTCCA GGTGAATCTCCGGGTGGCTCCAGCGGTTCCGAGTCA
LCW0401_004_ GFP-N_D01.ab1	GSGGEPSESGSSGSSESG SSEGGPGSGGEPSESGSS	GGTTCCGGTGGCGAACCGTCTGAATCTGGTAGCTCA GGTTCTTCTGAAAGCGGTTCTTCCGAGGGTGGTCCA GGTTCTGGTGGTGAACCTTCCGAGTCTGGTAGCTCA
LCW0401_007_ GFP-N_F01.ab1	GSSESGSSEGGPGSEGSS GPGESSGSEGSSGPGESS	GGTTCTTCCGAAAGCGGTTCTTCTGAGGGTGGTCCA GGTAGCGAAGGTTCTTCCGGTCCAGGTGAGTCTTCA GGTAGCGAAGGTTCTTCTGGTCCTGGTGAATCTTCA
LCW0401_008_ GFP-N_G01.ab1	GSSESGSSEGGPGESPGG SSGSESGSEGSSGPGESS	GGTTCCTCTGAAAGCGGTTCTTCCGAGGGTGGTCCA GGTGAATCTCCAGGTGGTTCCAGCGGTTCTGAGTCA GGTAGCGAAGGTTCTTCTGGTCCAGGTGAATCCTCA
LCW0401_012_ GFP-N_H01.ab1	GSGGEPSESGSSGSGGEP SESGSSGSEGSSGPGESS	GGTTCTGGTGGTGAACCGTCTGAGTCTGGTAGCTCA GGTTCCGGTGGCGAACCATCCGAATCTGGTAGCTCA GGTAGCGAAGGTTCTTCCGGTCCAGGTGAGTCTTCA
LCW0401_015_ GFP-N_A02.ab1	GSSESGSSEGGPGSEGSS GPGESSGESPGGSSGSES	GGTTCTTCCGAAAGCGGTTCTTCCGAAGGCGGTCCA GGTAGCGAAGGTTCTTCTGGTCCAGGCGAATCTTCA GGTGAATCTCCTGGTGGCTCCAGCGGTTCTGAGTCA
LCW0401_016_ GFP-N_B02.ab1	GSSESGSSEGGPGSSESG SSEGGPGSSESGSSEGGP	GGTTCCTCCGAAAGCGGTTCTTCTGAGGGCGGTCCA GGTTCCTCCGAAAGCGGTTCTTCCGAGGGCGGTCCA GGTTCTTCTGAAAGCGGTTCTTCCGAGGGCGGTCCA

File name	Amino acid sequence	Nucleotide sequence
LCW0401_020_ GFP-N_E02.ab1	GSGGEPSESGSSGSEGSS GPGESSGSSESGSSEGGP	GGTTCCGGTGGCGAACCGTCCGAATCTGGTAGCTCA GGTAGCGAAGGTTCTTCTGGTCCAGGCGAATCTTCA GGTTCCTCTGAAAGCGGTTCTTCTGAGGGCGGTCCA
LCW0401_022_ GFP-N_F02.ab1	GSGGEPSESGSSGSSESG SSEGGPGSGGEPSESGSS	GGTTCTGGTGGTGAACCGTCCGAATCTGGTAGCTCA GGTTCTTCCGAAAGCGGTTCTTCTGAAGGTGGTCCA GGTTCCGGTGGCGAACCTTCTGAATCTGGTAGCTCA
LCW0401_024_ GFP-N_G02.ab1	GSGGEPSESGSSGSSESG SSEGGPGESPGGSSGSES	GGTTCTGGTGGCGAACCGTCCGAATCTGGTAGCTCA GGTTCCTCCGAAAGCGGTTCTTCTGAAGGTGGTCCA GGTGAATCTCCAGGTGGTTCTAGCGGTTCTGAATCA
LCW0401_026_ GFP-N_H02.ab1	GSGGEPSESGSSGESPGG SSGSESGSEGSSGPGESS	GGTTCTGGTGGCGAACCGTCTGAGTCTGGTAGCTCA GGTGAATCTCCTGGTGGCTCCAGCGGTTCTGAATCA GGTAGCGAAGGTTCTTCTGGTCCTGGTGAATCTTCA
LCW0401_027_ GFP-N_A03.ab1	GSGGEPSESGSSGESPGG SSGSESGSGGEPSESGSS	GGTTCCGGTGGCGAACCTTCCGAATCTGGTAGCTCA GGTGAATCTCCGGGTGGTTCTAGCGGTTCTGAGTCA GGTTCTGGTGGTGAACCTTCCGAGTCTGGTAGCTCA
LCW0401_028_ GFP-N_B03.ab1	GSSESGSSEGGPGSSESG SSEGGPGSSESGSSEGGP	GGTTCCTCTGAAAGCGGTTCTTCTGAGGGCGGTCCA GGTTCTTCCGAAAGCGGTTCTTCCGAGGGCGGTCCA GGTTCTTCCGAAAGCGGTTCTTCTGAAGGCGGTCCA
LCW0401_030_ GFP-N_C03.ab1	GESPGGSSGSESGSEGSS GPGESSGSEGSSGPGESS	GGTGAATCTCCGGGTGGCTCCAGCGGTTCTGAGTCA GGTAGCGAAGGTTCTTCCGGTCCGGGTGAGTCCTCA GGTAGCGAAGGTTCTTCCGGTCCTGGTGAGTCTTCA
LCW0401_031_ GFP-N_D03.ab1	GSGGEPSESGSSGSGGEP SESGSSGSSESGSSEGGP	GGTTCTGGTGGCGAACCTTCCGAATCTGGTAGCTCA GGTTCCGGTGGTGAACCTTCTGAATCTGGTAGCTCA GGTTCTTCTGAAAGCGGTTCTTCCGAGGGCGGTCCA
LCW0401_033_ GFP-N_E03.ab1	GSGGEPSESGSSGSGGEP SESGSSGSGGEPSESGSS	GGTTCCGGTGGTGAACCTTCTGAATCTGGTAGCTCA GGTTCCGGTGGCGAACCATCCGAGTCTGGTAGCTCA GGTTCCGGTGGTGAACCATCCGAGTCTGGTAGCTCA
LCW0401_037_ GFP-N_F03.ab1	GSGGEPSESGSSGSSESG SSEGGPGSEGSSGPGESS	GGTTCCGGTGGCGAACCTTCTGAATCTGGTAGCTCA GGTTCCTCCGAAAGCGGTTCTTCTGAGGGCGGTCCA GGTAGCGAAGGTTCTTCTGGTCCGGGCGAGTCTTCA
LCW0401_038_ GFP-N_G03.ab1	GSGGEPSESGSSGSEGSS GPGESSGSGGEPSESGSS	GGTTCCGGTGGTGAACCGTCCGAGTCTGGTAGCTCA GGTAGCGAAGGTTCTTCTGGTCCGGGTGAGTCTTCA GGTTCTGGTGGCGAACCGTCCGAATCTGGTAGCTCA
LCW0401_039_ GFP-N_H03.ab1	GSGGEPSESGSSGESPGG SSGSESGSGGEPSESGSS	GGTTCTGGTGGCGAACCGTCCGAATCTGGTAGCTCA GGTGAATCTCCTGGTGGTTCCAGCGGTTCCGAGTCA GGTTCTGGTGGCGAACCTTCCGAATCTGGTAGCTCA
LCW0401_040_ GFP-N_A04.ab1	GSSESGSSEGGPGSGGEP SESGSSGSSESGSSEGGP	GGTTCTTCCGAAAGCGGTTCTTCCGAGGGCGGTCCA GGTTCCGGTGGTGAACCATCTGAATCTGGTAGCTCA GGTTCTTCTGAAAGCGGTTCTTCTGAAGGTGGTCCA
LCW0401_042_ GFP-N_C04.ab1	GSEGSSGPGESSGESPGG SSGSESGSEGSSGPGESS	GGTAGCGAAGGTTCTTCCGGTCCTGGTGAGTCTTCA GGTGAATCTCCAGGTGGCTCTAGCGGTTCCGAGTCA GGTAGCGAAGGTTCTTCTGGTCCTGGCGAGTCCTCA
LCW0401_046_ GFP-N_D04.ab1	GSSESGSSEGGPGSSESG SSEGGPGSSESGSSEGGP	GGTTCCTCTGAAAGCGGTTCTTCCGAAGGCGGTCCA GGTTCTTCCGAAAGCGGTTCTTCTGAGGGCGGTCCA GGTTCCTCCGAAAGCGGTTCTTCTGAGGGTGGTCCA
LCW0401_047_ GFP-N_E04.ab1	GSGGEPSESGSSGESPGG SSGSESGESPGGSSGSES	GGTTCTGGTGGCGAACCTTCCGAGTCTGGTAGCTCA GGTGAATCTCCGGGTGGTTCTAGCGGTTCCGAGTCA GGTGAATCTCCGGGTGGTTCCAGCGGTTCTGAGTCA
LCW0401_051_ GFP-N_F04.ab1	GSGGEPSESGSSGSEGSS GPGESSGESPGGSSGSES	GGTTCTGGTGGCGAACCATCTGAGTCTGGTAGCTCA GGTAGCGAAGGTTCTTCCGGTCCAGGCGAGTCTTCA GGTGAATCTCCTGGTGGCTCCAGCGGTTCTGAGTCA
LCW0401_053_ GFP-N_H04.ab1	GESPGGSSGSESGESPGG SSGSESGESPGGSSGSES	GGTGAATCTCCTGGTGGTTCCAGCGGTTCCGAGTCA GGTGAATCTCCAGGTGGCTCTAGCGGTTCCGAGTCA GGTGAATCTCCTGGTGGTTCTAGCGGTTCTGAATCA
LCW0401_054_ GFP-N_A05.ab1	GSEGSSGPGESSGSEGSS GPGESSGSGGEPSESGSS	GGTAGCGAAGGTTCTTCCGGTCCAGGTGAATCTTCA GGTAGCGAAGGTTCTTCTGGTCCTGGTGAATCCTCA GGTTCCGGTGGCGAACCATCTGAATCTGGTAGCTCA
LCW0401_059_ GFP-N_D05.ab1	GSGGEPSESGSSGSEGSS GPGESSGESPGGSSGSES	GGTTCTGGTGGCGAACCATCCGAATCTGGTAGCTCA GGTAGCGAAGGTTCTTCTGGTCCTGGCGAATCTTCA GGTGAATCTCCAGGTGGCTCTAGCGGTTCCGAATCA

File name	Amino acid sequence	Nucleotide sequence
LCW0401_060_ GFP-N_E05.ab1	GSGGEPSESGSSGSSESG SSEGGPGSGGEPSESGSS	GGTTCCGGTGGTGAACCGTCCGAATCTGGTAGCTCA GGTTCCTCTGAAAGCGGTTCTTCCGAGGGTGGTCCA GGTTCCGGTGGTGAACCTTCTGAGTCTGGTAGCTCA
LCW0401_061_ GFP-N_F05.ab1	GSSESGSSEGGPGSGGEP SESGSSGSEGSSGPGESS	GGTTCCTCTGAAAGCGGTTCTTCTGAGGGCGGTCCA GGTTCTGGTGGCGAACCATCTGAATCTGGTAGCTCA GGTAGCGAAGGTTCTTCCGGTCCGGGTGAATCTTCA
LCW0401_063_ GFP-N_H05.ab1	GSGGEPSESGSSGSEGSS GPGESSGSEGSSGPGESS	GGTTCTGGTGGTGAACCGTCCGAATCTGGTAGCTCA GGTAGCGAAGGTTCTTCTGGTCCTGGCGAGTCTTCA GGTAGCGAAGGTTCTTCTGGTCCTGGTGAATCTTCA
LCW0401_066_ GFP-N_B06.ab1	GSGGEPSESGSSGSSESG SSEGGPGSGGEPSESGSS	GGTTCTGGTGGCGAACCATCCGAGTCTGGTAGCTCA GGTTCTTCCGAAAGCGGTTCTTCCGAAGGCGGTCCA GGTTCTGGTGGTGAACCGTCCGAATCTGGTAGCTCA
LCW0401_067_ GFP-N_C06.ab1	GSGGEPSESGSSGESPGG SSGSESGESPGGSSGSES	GGTTCCGGTGGCGAACCTTCCGAATCTGGTAGCTCA GGTGAATCTCCGGGTGGTTCTAGCGGTTCCGAATCA GGTGAATCTCCAGGTGGTTCTAGCGGTTCCGAATCA
LCW0401_069_ GFP-N_D06.ab1	GSGGEPSESGSSGSGGEP SESGSSGESPGGSSGSES	GGTTCCGGTGGTGAACCATCTGAGTCTGGTAGCTCA GGTTCCGGTGGCGAACCGTCCGAGTCTGGTAGCTCA GGTGAATCTCCGGGTGGTTCCAGCGGTTCCGAATCA
LCW0401_070_ GFP-N_E06.ab1	GSEGSSGPGESSGSSESG SSEGGPGSEGSSGPGESS	GGTAGCGAAGGTTCTTCTGGTCCGGGCGAATCCTCA GGTTCCTCCGAAAGCGGTTCTTCCGAAGGTGGTCCA GGTAGCGAAGGTTCTTCCGGTCCTGGTGAATCTTCA
LCW0401_078_ GFP-N_F06.ab1	GSSESGSSEGGPGESPGG SSGSESGESPGGSSGSES	GGTTCCTCTGAAAGCGGTTCTTCTGAAGGCGGTCCA GGTGAATCTCCGGGTGGCTCCAGCGGTTCTGAATCA GGTGAATCTCCTGGTGGCTCCAGCGGTTCCGAGTCA
LCW0401_079_ GFP-N_G06.ab1	GSEGSSGPGESSGSEGSS GPGESSGSGGEPSESGSS	GGTAGCGAAGGTTCTTCTGGTCCAGGCGAGTCTTCA GGTAGCGAAGGTTCTTCCGGTCCTGGCGAGTCTTCA GGTTCCGGTGGCGAACCGTCCGAATCTGGTAGCTCA

[00379] Example 2: Construction of XTEN_AE36 segments

[00380] A codon library encoding XTEN sequences of 36 amino acid length was constructed. The XTEN sequence was designated XTEN_AE36. Its segments have the amino acid sequence $[X]_3$ where X is a 12mer peptide with the sequence: GSPAGSPTSTEE (SEQ ID NO: 23), GSEPATSGSETP (SEQ ID NO: 24), GTSESATPESGP (SEQ ID NO: 25), or GTSTEPSEGSAP (SEQ ID NO: 26). The insert was obtained by annealing the following pairs of phosphorylated synthetic oligonucleotide pairs:

AE1for: AGGTAGCCCDGCWGGYTCTCCDACYTCYACYGARGA (SEQ ID NO: 1628)

AE1rev: ACCTTCYTCRGTRGARGTHGGAGARCCWGCHGGGCT (SEQ ID NO: 1629)

AE2for: AGGTAGCGAACCKGCWACYTCYGGYTCTGARACYCC (SEQ ID NO: 1630)

AE2rev: ACCTGGRGTYTCAGARCCRGARGTWGCMGGTTCGCT (SEQ ID NO: 1631)

AE3for: AGGTACYTCTGAAAGCGCWACYCCKGARTCYGGYCC (SEQ ID NO: 1632)

AE3rev: ACCTGGRCCRGAYTCMGGRGTWGCGCTTTCAGARGT (SEQ ID NO: 1633)

AE4for: AGGTACYTCTACYGAACCKTCYGARGGYAGCGCWCC (SEQ ID NO: 1634)

AE4rev: ACCTGGWGCGCTRCCYTCRGAMGGTTCRGTAGARGT (SEQ ID NO: 1635)

[00381] We also annealed the phosphorylated oligonucleotide 3KpnIstopperFor: AGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 1626) and the non-phosphorylated oligonucleotide pr_3KpnIstopperRev: CCTCGAGTGAAGACGA (SEQ ID NO: 1627). The annealed oligonucleotide pairs were ligated, which resulted in a mixture of products with varying length that

represents the varying number of 12mer repeats ligated to one BbsI/KpnI segment. The products corresponding to the length of 36 amino acids were isolated from the mixture by preparative agarose gel electrophoresis and ligated into the BsaI/KpnI digested stuffer vector pCW0359. Most of the clones in the resulting library designated LCW0402 showed green fluorescence after induction which shows that the sequence of XTEN_AE36 had been ligated in frame with the GFP gene and most sequences of XTEN_AE36 show good expression.

[00382] We screened 96 isolates from library LCW0402 for high level of fluorescence by stamping them onto agar plate containing IPTG. The same isolates were evaluated by PCR and 48 isolates were identified that contained segments with 36 amino acids as well as strong fluorescence. These isolates were sequenced and 37 clones were identified that contained correct XTEN_AE36 segments. The file names of the nucleotide and amino acid constructs and the sequences for these segments are listed in Table 14.

Table 14: DNA and Amino Acid Sequences for AE 36-mer motifs (SEQ ID NOS 279-352, respectively, in order of appearance)

File name	Amino acid sequence	Nucleotide sequence
LCW0402_002_ GFP-N_A07.ab1	GSPAGSPTSTEEGTSE SATPESGPGTSTEPSE GSAP	GGTAGCCCGGCAGGCTCTCCGACCTCTACTGAGGAA GGTACTTCTGAAAGCGCAACCCCGGAGTCCGGCCCA GGTACCTCTACCGAACCGTCTGAGGGCAGCGCACCA
LCW0402_003_ GFP-N_B07.ab1	GTSTEPSEGSAPGTST EPSEGSAPGTSTEPSE GSAP	GGTACTTCTACCGAACCGTCCGAAGGCAGCGCTCCA GGTACCTCTACTGAACCTTCCGAGGGCAGCGCTCCA GGTACCTCTACCGAACCTTCTGAAGGTAGCGCACCA
LCW0402_004_ GFP-N_C07.ab1	GTSTEPSEGSAPGTSE SATPESGPGTSESATP ESGP	GGTACCTCTACCGAACCGTCTGAAGGTAGCGCACCA GGTACCTCTGAAAGCGCAACTCCTGAGTCCGGTCCA GGTACTTCTGAAAGCGCAACCCCGGAGTCTGGCCCA
LCW0402_005_ GFP-N_D07.ab1	GTSTEPSEGSAPGTSE SATPESGPGTSESATP ESGP	GGTACTTCTACTGAACCGTCTGAAGGTAGCGCACCA GGTACTTCTGAAAGCGCAACCCCGGAATCCGGCCCA GGTACCTCTGAAAGCGCAACCCCGGAGTCCGGCCCA
LCW0402_006_ GFP-N_E07.ab1	GSEPATSGSETPGTSE SATPESGPGSPAGSPT STEE	GGTAGCGAACCGGCAACCTCCGGCTCTGAAACCCCA GGTACCTCTGAAAGCGCTACTCCTGAATCCGGCCCA GGTAGCCCGGCAGGTTCTCCGACTTCCACTGAGGAA
LCW0402_008_ GFP-N_F07.ab1	GTSESATPESGPGSEP ATSGSETPGTSTEPSE GSAP	GGTACTTCTGAAAGCGCAACCCCTGAATCCGGTCCA GGTAGCGAACCGGCTACTTCTGGCTCTGAGACTCCA GGTACTTCTACCGAACCGTCCGAAGGTAGCGCACCA
LCW0402_009_ GFP-N_G07.ab1	GSPAGSPTSTEEGSPA GSPTSTEEGSEPATSG SETP	GGTAGCCCGGCTGGCTCTCCAACCTCCACTGAGGAA GGTAGCCCGGCTGGCTCTCCAACCTCCACTGAAGAA GGTAGCGAACCGGCTACCTCCGGCTCTGAAACTCCA
LCW0402_011_ GFP-N_A08.ab1	GSPAGSPTSTEEGTSE SATPESGPGTSTEPSE GSAP	GGTAGCCCGGCTGGCTCTCCTACCTCTACTGAGGAA GGTACTTCTGAAAGCGCTACTCCTGAGTCTGGTCCA GGTACCTCTACTGAACCGTCCGAAGGTAGCGCTCCA
LCW0402_012_ GFP-N_B08.ab1	GSPAGSPTSTEEGSPA GSPTSTEEGTSTEPSE GSAP	GGTAGCCCTGCTGGCTCTCCGACTTCTACTGAGGAA GGTAGCCCGGCTGGTTCTCCGACTTCTACTGAGGAA GGTACTTCTACCGAACCTTCCGAAGGTAGCGCTCCA
LCW0402_013_ GFP-N_C08.ab1	GTSESATPESGPGTST EPSEGSAPGTSTEPSE GSAP	GGTACTTCTGAAAGCGCTACTCCGGAGTCCGGTCCA GGTACCTCTACCGAACCGTCCGAAGGCAGCGCTCCA GGTACTTCTACTGAACCTTCTGAGGGTAGCGCTCCA
LCW0402_014_ GFP-N_D08.ab1	GTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSE GSAP	GGTACCTCTACCGAACCTTCCGAAGGTAGCGCTCCA GGTAGCCCGGCAGGTTCTCCTACTTCCACTGAGGAA GGTACTTCTACCGAACCTTCTGAGGGTAGCGCACCA
LCW0402_015_ GFP-N_E08.ab1	GSEPATSGSETPGSPA GSPTSTEEGTSESATP ESGP	GGTAGCGAACCGGCTACTTCCGGCTCTGAGACTCCA GGTAGCCCTGCTGGCTCTCCGACCTCTACCGAAGAA GGTACCTCTGAAAGCGCTACCCCTGAGTCTGGCCCA
LCW0402_016_ GFP-N_F08.ab1	GTSTEPSEGSAPGTSE SATPESGPGTSESATP	GGTACTTCTACCGAACCTTCCGAGGGCAGCGCACCA GGTACTTCTGAAAGCGCTACCCCTGAGTCCGGCCCA

File name	Amino acid sequence	Nucleotide sequence
	ESGP	GGTACTTCTGAAAGCGCTACTCCTGAATCCGGTCCA
LCW0402_020_ GFP-N_G08.ab1	GTSTEPSEGSAPGSEP ATSGSETPGSPAGSPT STEE	GGTACTTCTACTGAACCGTCTGAAGGCAGCGCACCA GGTAGCGAACCGGCTACTTCCGGTTCTGAAACCCCA GGTAGCCCAGCAGGTTCTCCAACTTCTACTGAAGAA
LCW0402_023_ GFP-N_A09.ab1	GSPAGSPTSTEEGTSE SATPESGPGSEPATSG SETP	GGTAGCCCTGCTGGCTCTCCAACCTCCACCGAAGAA GGTACCTCTGAAAGCGCAACCCCTGAATCCGGCCCA GGTAGCGAACCGGCAACCTCCGGTTCTGAAACCCCA
LCW0402_024_ GFP-N_B09.ab1	GTSESATPESGPGSPA GSPTSTEEGSPAGSPT STEE	GGTACTTCTGAAAGCGCTACTCCTGAGTCCGGCCCA GGTAGCCCGGCTGGCTCTCCGACTTCCACCGAGGAA GGTAGCCCGGCTGGCTCTCCAACTTCTACTGAAGAA
LCW0402_025_ GFP-N_C09.ab1	GTSTEPSEGSAPGTSE SATPESGPGTSTEPSE GSAP	GGTACCTCTACTGAACCTTCTGAGGGCAGCGCTCCA GGTACTTCTGAAAGCGCTACCCCGGAGTCCGGTCCA GGTACTTCTACTGAACCGTCCGAAGGTAGCGCACCA
LCW0402_026_ GFP-N_D09.ab1	GSPAGSPTSTEEGTST EPSEGSAPGSEPATSG SETP	GGTAGCCCGGCAGGCTCTCCGACTTCCACCGAGGAA GGTACCTCTACTGAACCTTCTGAGGGTAGCGCTCCA GGTAGCGAACCGGCAACCTCTGGCTCTGAAACCCCA
LCW0402_027_ GFP-N_E09.ab1	GSPAGSPTSTEEGTST EPSEGSAPGTSTEPSE GSAP	GGTAGCCCAGCAGGCTCTCCGACTTCCACTGAGGAA GGTACTTCTACTGAACCTTCCGAAGGCAGCGCACCA GGTACCTCTACTGAACCTTCTGAGGGCAGCGCTCCA
LCW0402_032_ GFP-N_H09.ab1	GSEPATSGSETPGTSE SATPESGPGSPAGSPT STEE	GGTAGCGAACCTGCTACCTCCGGTTCTGAAACCCCA GGTACCTCTGAAAGCGCAACTCCGGAGTCTGGTCCA GGTAGCCCTGCAGGTTCTCCTACCTCCACTGAGGAA
LCW0402_034_ GFP-N_A10.ab1	GTSESATPESGPGTST EPSEGSAPGTSTEPSE GSAP	GGTACCTCTGAAAGCGCTACTCCGGAGTCTGGCCCA GGTACCTCTACTGAACCGTCTGAGGGTAGCGCTCCA GGTACTTCTACTGAACCGTCCGAAGGTAGCGCACCA
LCW0402_036_ GFP-N_C10.ab1	GSPAGSPTSTEEGTST EPSEGSAPGTSTEPSE GSAP	GGTAGCCCGGCTGGTTCTCCGACTTCCACCGAGGAA GGTACCTCTACTGAACCTTCTGAGGGTAGCGCTCCA GGTACCTCTACTGAACCTTCCGAAGGCAGCGCTCCA
LCW0402_039_ GFP-N_E10.ab1	GTSTEPSEGSAPGTST EPSEGSAPGTSTEPSE GSAP	GGTACTTCTACCGAACCGTCCGAGGGCAGCGCTCCA GGTACTTCTACTGAACCTTCTGAAGGCAGCGCTCCA GGTACTTCTACTGAACCTTCCGAAGGTAGCGCACCA
LCW0402_040_ GFP-N_F10.ab1	GSEPATSGSETPGTSE SATPESGPGTSTEPSE GSAP	GGTAGCGAACCTGCAACCTCTGGCTCTGAAACCCCA GGTACCTCTGAAAGCGCTACTCCTGAATCTGGCCCA GGTACTTCTACTGAACCGTCCGAGGGCAGCGCACCA
LCW0402_041_ GFP-N_G10.ab1	GTSTEPSEGSAPGSPA GSPTSTEEGTSTEPSE GSAP	GGTACTTCTACCGAACCGTCCGAGGGTAGCGCACCA GGTAGCCCAGCAGGTTCTCCTACCTCCACCGAGGAA GGTACTTCTACCGAACCGTCCGAGGGTAGCGCACCA
LCW0402_050_ GFP-N_A11.ab1	GSEPATSGSETPGTSE SATPESGPGSEPATSG SETP	GGTAGCGAACCGGCAACCTCCGGCTCTGAAACTCCA GGTACTTCTGAAAGCGCTACTCCGGAATCCGGCCCA GGTAGCGAACCGGCTACTTCCGGCTCTGAAACCCCA
LCW0402_051_ GFP-N_B11.ab1	GSEPATSGSETPGTSE SATPESGPGSEPATSG SETP	GGTAGCGAACCGGCAACTTCCGGCTCTGAAACCCCA GGTACTTCTGAAAGCGCTACTCCTGAGTCTGGCCCA GGTAGCGAACCTGCTACCTCTGGCTCTGAAACCCCA
LCW0402_059_ GFP-N_E11.ab1	GSEPATSGSETPGSEP ATSGSETPGTSTEPSE GSAP	GGTAGCGAACCGGCAACCTCTGGCTCTGAAACTCCA GGTAGCGAACCTGCAACCTCCGGCTCTGAAACCCCA GGTACTTCTACTGAACCTTCTGAGGGCAGCGCACCA
LCW0402_060_ GFP-N_F11.ab1	GTSESATPESGPGSEP ATSGSETPGSEPATSG SETP	GGTACTTCTGAAAGCGCTACCCCGGAATCTGGCCCA GGTAGCGAACCGGCTACTTCTGGTTCTGAAACCCCA GGTAGCGAACCGGCTACCTCCGGTTCTGAAACTCCA
LCW0402_061_ GFP-N_G11.ab1	GTSTEPSEGSAPGTST EPSEGSAPGTSESATP ESGP	GGTACCTCTACTGAACCTTCCGAAGGCAGCGCTCCA GGTACCTCTACCGAACCGTCCGAGGGCAGCGCACCA GGTACTTCTGAAAGCGCAACCCCTGAATCCGGTCCA
LCW0402_065_ GFP-N_A12.ab1	GSEPATSGSETPGTSE SATPESGPGTSESATP ESGP	GGTAGCGAACCGGCAACCTCTGGCTCTGAAACCCCA GGTACCTCTGAAAGCGCTACTCCGGAATCTGGTCCA GGTACTTCTGAAAGCGCTACTCCGGAATCCGGTCCA
LCW0402_066_ GFP-N_B12.ab1	GSEPATSGSETPGSEP ATSGSETPGTSTEPSE GSAP	GGTAGCGAACCTGCTACCTCCGGCTCTGAAACTCCA GGTAGCGAACCGGCTACTTCCGGTTCTGAAACTCCA GGTACCTCTACCGAACCTTCCGAAGGCAGCGCACCA
LCW0402_067_ GFP-N_C12.ab1	GSEPATSGSETPGTST EPSEGSAPGSEPATSG SETP	GGTAGCGAACCTGCTACTTCTGGTTCTGAAACTCCA GGTACTTCTACCGAACCGTCCGAGGGTAGCGCTCCA GGTAGCGAACCTGCTACTTCTGGTTCTGAAACTCCA

File name	Amino acid sequence	Nucleotide sequence
LCW0402_069_ GFP-N_D12.ab1	GTSTEPSEGSAPGTST EPSEGSAPGSEPATSG SETP	GGTACCTCTACCGAACCGTCCGAGGGTAGCGCACCA GGTACCTCTACTGAACCGTCTGAGGGTAGCGCTCCA GGTAGCGAACCGGCAACCTCCGGTTCTGAAACTCCA
LCW0402_073_ GFP-N_F12.ab1	GTSTEPSEGSAPGSEP ATSGSETPGSPAGSPT STEE	GGTACTTCTACTGAACCTTCCGAAGGTAGCGCTCCA GGTAGCGAACCTGCTACTTCTGGTTCTGAAACCCCA GGTAGCCCGGCTGGCTCTCCGACCTCCACCGAGGAA
LCW0402_074_ GFP-N_G12.ab1	GSEPATSGSETPGSPA GSPTSTEEGTSESATP ESGP	GGTAGCGAACCGGCTACTTCCGGCTCTGAGACTCCA GGTAGCCCAGCTGGTTCTCCAACCTCTACTGAGGAA GGTACTTCTGAAAGCGCTACCCCTGAATCTGGTCCA
LCW0402_075_ GFP-N_H12.ab1	GTSESATPESGPGSEP ATSGSETPGTSESATP ESGP	GGTACCTCTGAAAGCGCAACTCCTGAGTCTGGCCCA GGTAGCGAACCTGCTACCTCCGGCTCTGAGACTCCA GGTACCTCTGAAAGCGCAACCCCGGAATCTGGTCCA

[00383] Example 3: Construction of XTEN_AF36 segments

[00384] A codon library encoding sequences of 36 amino acid length was constructed. The sequences were designated XTEN_AF36. Its segments have the amino acid sequence $[X]_3$ where X is a 12mer peptide with the sequence: GSTSESPSGTAP (SEQ ID NO: 27), GTSTPESGSASP (SEQ ID NO: 28), GTSPSGESSTAP (SEQ ID NO: 29), or GSTSSTAESPGP (SEQ ID NO: 30). The insert was obtained by annealing the following pairs of phosphorylated synthetic oligonucleotide pairs:

AF1for: AGGTTCTACYAGCGAATCYCCKTCTGGYACYGCWCC (SEQ ID NO: 1636)

AF1rev: ACCTGGWGCRGTRCCAGAMGGRGATTCGCTRGTAGA (SEQ ID NO: 1637)

AF2for: AGGTACYTCTACYCCKGAAAGCGGYTCYGCWTCTCC (SEQ ID NO: 1638)

AF2rev: ACCTGGAGAWGCRGARCCGCTTTCMGGRGTAGARGT (SEQ ID NO: 1639)

AF3for: AGGTACYTCYCCKAGCGGYGAATCTTCTACYGCWCC (SEQ ID NO: 1640)

AF3rev: ACCTGGWGCRGTAGAAGATTCRCCGCTMGGRGARGT (SEQ ID NO: 1641)

AF4for: AGGTTCYACYAGCTCTACYGCWGAATCTCCKGGYCC (SEQ ID NO: 1642)

AF4rev: ACCTGGRCCMGGAGATTCWGCRGTAGAGCTRGTRGA (SEQ ID NO: 1643)

[00385] We also annealed the phosphorylated oligonucleotide 3KpnIstopperFor: AGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 1626) and the non-phosphorylated oligonucleotide pr_3KpnIstopperRev: CCTCGAGTGAAGACGA (SEQ ID NO: 1627). The annealed oligonucleotide pairs were ligated, which resulted in a mixture of products with varying length that represents the varying number of 12mer repeats ligated to one BbsI/KpnI segment The products corresponding to the length of 36 amino acids were isolated from the mixture by preparative agarose gel electrophoresis and ligated into the BsaI/KpnI digested stuffer vector pCW0359. Most of the clones in the resulting library designated LCW0403 showed green fluorescence after induction which shows that the sequence of XTEN_AF36 had been ligated in frame with the GFP gene and most sequences of XTEN_AF36 show good expression.

[00386] We screened 96 isolates from library LCW0403 for high level of fluorescence by stamping them onto agar plate containing IPTG. The same isolates were evaluated by PCR and 48 isolates were identified that contained segments with 36 amino acids as well as strong fluorescence. These isolates were sequenced and 44 clones were identified that contained correct XTEN_AF36 segments. The file

names of the nucleotide and amino acid constructs and the sequences for these segments are listed in Table 15.

Table 15: DNA and Amino Acid Sequences for AF 36-mer motifs (SEQ ID NOS 353-440, respectively, in order of appearance)

File name	Amino acid sequence	Nucleotide sequence
LCW0403_004_ GFP-N_A01.ab1	GTSTPESGSASPGTSP SGESSTAPGTSPSGES STAP	GGTACTTCTACTCCGGAAAGCGGTTCCGCATCTCCA GGTACTTCTCCTAGCGGTGAATCTTCTACTGCTCCAG GTACCTCTCCTAGCGGCGAATCTTCTACTGCTCCA
LCW0403_005_ GFP-N_B01.ab1	GTSPSGESSTAPGSTS STAESPGPGTSPSGES STAP	GGTACTTCTCCGAGCGGTGAATCTTCTACCGCACCA GGTTCTACTAGCTCTACCGCTGAATCTCCGGGCCCAG GTACTTCTCCGAGCGGTGAATCTTCTACTGCTCCA
LCW0403_006_ GFP-N_C01.ab1	GSTSSTAESPGPGTSP SGESSTAPGTSTPESG SASP	GGTTCCACCAGCTCTACTGCTGAATCTCCTGGTCCAG GTACCTCTCCTAGCGGTGAATCTTCTACTGCTCCAGG TACTTCTACTCCTGAAAGCGGCTCTGCTTCTCCA
LCW0403_007_ GFP-N_D01.ab1	GSTSSTAESPGPGSTS STAESPGPGTSPSGES STAP	GGTTCTACCAGCTCTACTGCAGAATCTCCTGGCCCAG GTTCCACCAGCTCTACCGCAGAATCTCCGGGTCCAG GTACTTCCCCTAGCGGTGAATCTTCTACCGCACCA
LCW0403_008_ GFP-N_E01.ab1	GSTSSTAESPGPGTSP SGESSTAPGTSTPESG SASP	GGTTCTACTAGCTCTACTGCTGAATCTCCTGGCCCAG GTACTTCTCCTAGCGGTGAATCTTCTACCGCTCCAGG TACCTCTACTCCGGAAAGCGGTTCTGCATCTCCA
LCW0403_010_ GFP-N_F01.ab1	GSTSSTAESPGPGTST PESGSASPGSTSESPS GTAP	GGTTCTACCAGCTCTACCGCAGAATCTCCTGGTCCAG GTACCTCTACTCCGGAAAGCGGCTCTGCATCTCCAG GTTCTACTAGCGAATCTCCTTCTGGCACTGCACCA
LCW0403_011_ GFP-N_G01.ab1	GSTSSTAESPGPGTST PESGSASPGTSTPESG SASP	GGTTCTACTAGCTCTACTGCAGAATCTCCTGGCCCAG GTACCTCTACTCCGGAAAGCGGCTCTGCATCTCCAG GTACTTCTACCCCTGAAAGCGGTTCTGCATCTCCA
LCW0403_012_ GFP-N_H01.ab1	GSTSESPSGTAPGTSP SGESSTAPGSTSESPS GTAP	GGTTCTACCAGCGAATCTCCTTCTGGCACCGCTCCAG GTACCTCTCCTAGCGGCGAATCTTCTACCGCTCCAGG TTCTACTAGCGAATCTCCTTCTGGCACTGCACCA
LCW0403_013_ GFP-N_A02.ab1	GSTSSTAESPGPGSTS STAESPGPGTSPSGES STAP	GGTTCCACCAGCTCTACTGCAGAATCTCCGGGCCCA GGTTCTACTAGCTCTACTGCAGAATCTCCGGGTCCAG GTACTTCTCCTAGCGGCGAATCTTCTACCGCTCCA
LCW0403_014_ GFP-N_B02.ab1	GSTSSTAESPGPGTST PESGSASPGSTSESPS GTAP	GGTTCCACTAGCTCTACTGCAGAATCTCCTGGCCCAG GTACCTCTACCCCTGAAAGCGGCTCTGCATCTCCAG GTTCTACCAGCGAATCCCCGTCTGGCACCGCACCA
LCW0403_015_ GFP-N_C02.ab1	GSTSSTAESPGPGSTS STAESPGPGTSPSGES STAP	GGTTCTACTAGCTCTACTGCTGAATCTCCGGGTCCAG GTTCTACCAGCTCTACTGCTGAATCTCCTGGTCCAGG TACCTCCCCGAGCGGTGAATCTTCTACTGCACCA
LCW0403_017_ GFP-N_D02.ab1	GSTSSTAESPGPGSTS ESPSGTAPGSTSSTAE SPGP	GGTTCTACCAGCTCTACCGCTGAATCTCCTGGCCCAG GTTCTACCAGCGAATCCCCGTCTGGCACCGCACCAG GTTCTACTAGCTCTACCGCTGAATCTCCGGGTCCA
LCW0403_018_ GFP-N_E02.ab1	GSTSSTAESPGPGSTS STAESPGPGSTSSTAE SPGP	GGTTCTACCAGCTCTACCGCAGAATCTCCTGGCCCA GGTTCCACTAGCTCTACCGCTGAATCTCCTGGTCCAG GTTCTACTAGCTCTACCGCTGAATCTCCTGGTCCA
LCW0403_019_ GFP-N_F02.ab1	GSTSESPSGTAPGSTS STAESPGPGSTSSTAE SPGP	GGTTCTACTAGCGAATCCCCTTCTGGTACTGCTCCAG GTTCCACTAGCTCTACCGCTGAATCTCCTGGCCCAGG TTCCACTAGCTCTACTGCAGAATCTCCTGGTCCA
LCW0403_023_ GFP-N_H02.ab1	GSTSESPSGTAPGSTS ESPSGTAPGSTSESPS GTAP	GGTTCTACTAGCGAATCTCCTTCTGGTACCGCTCCAG GTTCTACCAGCGAATCCCCGTCTGGTACTGCTCCAGG TTCTACCAGCGAATCTCCTTCTGGTACTGCACCA
LCW0403_024_ GFP-N_A03.ab1	GSTSSTAESPGPGSTS STAESPGPGSTSSTAE SPGP	GGTTCCACCAGCTCTACTGCTGAATCTCCTGGCCCAG GTTCTACCAGCTCTACTGCTGAATCTCCGGGCCCAGG TTCCACCAGCTCTACCGCTGAATCTCCGGGTCCA
LCW0403_025_ GFP-N_B03.ab1	GSTSSTAESPGPGSTS STAESPGPGTSPSGES STAP	GGTTCCACTAGCTCTACCGCAGAATCTCCTGGTCCAG GTTCTACTAGCTCTACTGCTGAATCTCCGGGTCCAGG TACCTCCCCTAGCGGCGAATCTTCTACCGCTCCA
LCW0403_028_ GFP-N_D03.ab1	GSSPSASTGTGPGSST PSGATGSPGSSTPSGA TGSP	GGTTCTAGCCCTTCTGCTTCCACCGGTACCGGCCCAG GTAGCTCTACTCCGTCTGGTGCAACTGGCTCTCCAGG TAGCTCTACTCCGTCTGGTGCAACCGGCTCCCCA

File name	Amino acid sequence	Nucleotide sequence
LCW0403_029_ GFP-N_E03.ab1	GTSPSGESSTAPGTST PESGSASPGSTSSTAE SPGP	GGTACTTCCCCTAGCGGTGAATCTTCTACTGCTCCAG GTACCTCTACTCCGGAAAGCGGCTCCGCATCTCCAG GTTCTACTAGCTCTACTGCTGAATCTCCTGGTCCA
LCW0403_030_ GFP-N_F03.ab1	GSTSSTAESPGPGSTS STAESPGPGTSTPESG SASP	GGTTCTACTAGCTCTACCGCTGAATCTCCGGGTCCAG GTTCTACCAGCTCTACTGCAGAATCTCCTGGCCCAGG TACTTCTACTCCGGAAAGCGGTTCCGCTTCTCCA
LCW0403_031_ GFP-N_G03.ab1	GTSPSGESSTAPGSTS STAESPGPGTSTPESG SASP	GGTACTTCTCCTAGCGGTGAATCTTCTACCGCTCCAG GTTCTACCAGCTCTACTGCTGAATCTCCTGGCCCAGG TACTTCTACCCCGGAAAGCGGCTCCGCTTCTCCA
LCW0403_033_ GFP-N_H03.ab1	GSTSESPSGTAPGSTS STAESPGPGSTSSTAE SPGP	GGTTCTACTAGCGAATCCCCTTCTGGTACTGCACCAG GTTCTACCAGCTCTACTGCTGAATCTCCGGGCCCAGG TTCCACCAGCTCTACCGCAGAATCTCCTGGTCCA
LCW0403_035_ GFP-N_A04.ab1	GSTSSTAESPGPGSTS ESPSGTAPGSTSSTAE SPGP	GGTTCCACCAGCTCTACCGCTGAATCTCCGGGCCCA GGTTCTACCAGCGAATCCCCTTCTGGCACTGCACCA GGTTCTACTAGCTCTACCGCAGAATCTCCGGGCCCA
LCW0403_036_ GFP-N_B04.ab1	GSTSSTAESPGPGTSP SGESSTAPGTSTPESG SASP	GGTTCTACCAGCTCTACTGCTGAATCTCCGGGTCCAG GTACTTCCCCGAGCGGTGAATCTTCTACTGCACCAG GTACTTCTACTCCGGAAAGCGGTTCCGCTTCTCCA
LCW0403_039_ GFP-N_C04.ab1	GSTSESPSGTAPGSTS ESPSGTAPGTSPSGES STAP	GGTTCTACCAGCGAATCTCCTTCTGGCACCGCTCCAG GTTCTACTAGCGAATCCCCGTCTGGTACCGCACCAG GTACTTCTCCTAGCGGCGAATCTTCTACCGCACCA
LCW0403_041_ GFP-N_D04.ab1	GSTSESPSGTAPGSTS ESPSGTAPGTSTPESG SASP	GGTTCTACCAGCGAATCCCCTTCTGGTACTGCTCCAG GTTCTACCAGCGAATCCCCTTCTGGCACCGCACCAG GTACTTCTACCCCTGAAAGCGGCTCCGCTTCTCCA
LCW0403_044_ GFP-N_E04.ab1	GTSTPESGSASPGSTS STAESPGPGSTSSTAE SPGP	GGTACCTCTACTCCTGAAAGCGGTTCTGCATCTCCAG GTTCCACTAGCTCTACCGCAGAATCTCCGGGCCCAG GTTCTACTAGCTCTACTGCTGAATCTCCTGGCCCA
LCW0403_046_ GFP-N_F04.ab1	GSTSESPSGTAPGSTS ESPSGTAPGTSPSGES STAP	GGTTCTACCAGCGAATCCCCTTCTGGCACTGCACCA GGTTCTACTAGCGAATCCCCTTCTGGTACCGCACCAG GTACTTCTCCGAGCGGCGAATCTTCTACTGCTCCA
LCW0403_047_ GFP-N_G04.ab1	GSTSSTAESPGPGSTS STAESPGPGSTSESPS GTAP	GGTTCTACTAGCTCTACCGCTGAATCTCCTGGCCCAG GTTCCACTAGCTCTACCGCAGAATCTCCGGGCCCAG GTTCTACTAGCGAATCCCCTTCTGGTACCGCTCCA
LCW0403_049_ GFP-N_H04.ab1	GSTSSTAESPGPGSTS STAESPGPGTSTPESG SASP	GGTTCCACCAGCTCTACTGCAGAATCTCCTGGCCCA GGTTCTACTAGCTCTACCGCAGAATCTCCTGGTCCAG GTACCTCTACTCCTGAAAGCGGTTCCGCATCTCCA
LCW0403_051_ GFP-N_A05.ab1	GSTSSTAESPGPGSTS STAESPGPGSTSESPS GTAP	GGTTCTACTAGCTCTACTGCTGAATCTCCGGGCCCAG GTTCTACTAGCTCTACCGCTGAATCTCCGGGTCCAGG TTCTACTAGCGAATCTCCTTCTGGTACCGCTCCA
LCW0403_053_ GFP-N_B05.ab1	GTSPSGESSTAPGSTS ESPSGTAPGSTSSTAE SPGP	GGTACCTCCCCGAGCGGTGAATCTTCTACTGCACCA GGTTCTACTAGCGAATCCCCTTCTGGTACTGCTCCAG GTTCCACCAGCTCTACTGCAGAATCTCCGGGTCCA
LCW0403_054_ GFP-N_C05.ab1	GSTSESPSGTAPGTSP SGESSTAPGSTSSTAE SPGP	GGTTCTACTAGCGAATCCCCGTCTGGTACTGCTCCAG GTACTTCCCCTAGCGGTGAATCTTCTACTGCTCCAGG TTCTACCAGCTCTACCGCAGAATCTCCGGGTCCA
LCW0403_057_ GFP-N_D05.ab1	GSTSSTAESPGPGSTS ESPSGTAPGTSPSGES STAP	GGTTCTACCAGCTCTACCGCTGAATCTCCTGGCCCAG GTTCTACTAGCGAATCTCCGTCTGGCACCGCACCAG GTACTTCCCCTAGCGGTGAATCTTCTACTGCACCA
LCW0403_058_ GFP-N_E05.ab1	GSTSESPSGTAPGSTS ESPSGTAPGTSTPESG SASP	GGTTCTACTAGCGAATCTCCTTCTGGCACTGCACCAG GTTCTACCAGCGAATCTCCGTCTGGCACTGCACCAG GTACCTCTACCCCTGAAAGCGGTTCCGCTTCTCCA
LCW0403_060_ GFP-N_F05.ab1	GTSTPESGSASPGSTS ESPSGTAPGSTSSTAE SPGP	GGTACCTCTACTCCGGAAAGCGGTTCCGCATCTCCA GGTTCTACCAGCGAATCCCCGTCTGGCACCGCACCA GGTTCTACTAGCTCTACTGCTGAATCTCCGGGCCCA
LCW0403_063_ GFP-N_G05.ab1	GSTSSTAESPGPGTSP SGESSTAPGTSPSGES STAP	GGTTCTACTAGCTCTACTGCAGAATCTCCGGGCCCA GGTACCTCTCCTAGCGGTGAATCTTCTACCGCTCCAG GTACTTCTCCGAGCGGTGAATCTTCTACCGCTCCA
LCW0403_064_ GFP-N_H05.ab1	GTSPSGESSTAPGTSP SGESSTAPGTSPSGES STAP	GGTACCTCCCCTAGCGGCGAATCTTCTACTGCTCCAG GTACCTCTCCTAGCGGCGAATCTTCTACCGCTCCAGG TACCTCCCCTAGCGGTGAATCTTCTACCGCACCA
LCW0403_065_	GSTSSTAESPGPGTST	GGTTCCACTAGCTCTACTGCTGAATCTCCTGGCCCAG

File name	Amino acid sequence	Nucleotide sequence
GFP-N_A06.ab1	PESGSASPGSTSESPS GTAP	GTACTTCTACTCCGGAAAGCGGTTCCGCTTCTCCAGG TTCTACTAGCGAATCTCCGTCTGGCACCGCACCA
LCW0403_066_ GFP-N_B06.ab1	GSTSESPSGTAPGTSP SGESSTAPGTSPSGES STAP	GGTTCTACTAGCGAATCTCCGTCTGGCACTGCTCCAG GTACTTCTCCTAGCGGTGAATCTTCTACCGCTCCAGG TACTTCCCCTAGCGGCGAATCTTCTACCGCTCCA
LCW0403_067_ GFP-N_C06.ab1	GSTSESPSGTAPGTST PESGSASPGSTSSTAE SPGP	GGTTCTACTAGCGAATCTCCTTCTGGTACCGCTCCAG GTACTTCTACCCCTGAAAGCGGCTCCGCTTCTCCAGG TTCCACTAGCTCTACCGCTGAATCTCCGGGTCCA
LCW0403_068_ GFP-N_D06.ab1	GSTSSTAESPGPGSTS STAESPGPGSTSESPS GTAP	GGTTCCACTAGCTCTACTGCTGAATCTCCTGGCCCAG GTTCTACCAGCTCTACCGCTGAATCTCCTGGCCCAGG TTCTACCAGCGAATCTCCGTCTGGCACCGCACCA
LCW0403_069_ GFP-N_E06.ab1	GSTSESPSGTAPGTST PESGSASPGTSTPESG SASP	GGTTCTACTAGCGAATCCCCGTCTGGTACCGCACCA GGTACTTCTACCCCGGAAAGCGGCTCTGCTTCTCCAG GTACTTCTACCCCGGAAAGCGGCTCCGCATCTCCA
LCW0403_070_ GFP-N_F06.ab1	GSTSESPSGTAPGTST PESGSASPGTSTPESG SASP	GGTTCTACTAGCGAATCCCCGTCTGGTACTGCTCCAG GTACTTCTACTCCTGAAAGCGGTTCCGCTTCTCCAGG TACCTCTACTCCGGAAAGCGGTTCTGCATCTCCA

[00387] Example 4: Construction of XTEN_AG36 segments

[00388] A codon library encoding sequences of 36 amino acid length was constructed. The sequences were designated XTEN_AG36. Its segments have the amino acid sequence $[X]_3$ where X is a 12mer peptide with the sequence: GTPGSGTASSSP (SEQ ID NO: 31), GSSTPSGATGSP (SEQ ID NO: 32), GSSPSASTGTGP (SEQ ID NO: 33), or GASPGTSSTGSP (SEQ ID NO: 34). The insert was obtained by annealing the following pairs of phosphorylated synthetic oligonucleotide pairs:

AG1for: AGGTACYCCKGGYAGCGGTACYGCWTCTTCYTCTCC (SEQ ID NO: 1644)

AG1rev: ACCTGGAGARGAAGAWGCRGTACCGCTRCCMGGRGT (SEQ ID NO: 1645)

AG2for: AGGTAGCTCTACYCCKTCTGGTGCWACYGGYTCYCC (SEQ ID NO: 1646)

AG2rev: ACCTGGRGARCCRGTWGCACCAGAMGGRGTAGAGCT (SEQ ID NO: 1647)

AG3for: AGGTTCTAGCCCKTCTGCWTCYACYGGTACYGGYCC (SEQ ID NO: 1648)

AG3rev: ACCTGGRCCRGTACCRGTRGAWGCAGAMGGGCTAGA (SEQ ID NO: 1649)

AG4for: AGGTGCWTCYCCKGGYACYAGCTCTACYGGTTCTCC (SEQ ID NO: 1650)

AG4rev: ACCTGGAGAACCRGTAGAGCTRGTRCCMGGRGAWGC (SEQ ID NO: 1651)

[00389] We also annealed the phosphorylated oligonucleotide 3KpnIstopperFor: AGGTTCGTCTTCACTCGAGGGTAC (SEQ ID NO: 1626) and the non-phosphorylated oligonucleotide pr_3KpnIstopperRev: CCTCGAGTGAAGACGA (SEQ ID NO: 1627). The annealed oligonucleotide pairs were ligated, which resulted in a mixture of products with varying length that represents the varying number of 12mer repeats ligated to one BbsI/KpnI segment. The products corresponding to the length of 36 amino acids were isolated from the mixture by preparative agarose gel electrophoresis and ligated into the BsaI/KpnI digested stuffer vector pCW0359. Most of the clones in the resulting library designated LCW0404 showed green fluorescence after induction which shows that the sequence of XTEN_AG36 had been ligated in frame with the GFP gene and most sequences of XTEN_AG36 show good expression.

[00390] We screened 96 isolates from library LCW0404 for high level of fluorescence by stamping them onto agar plate containing IPTG. The same isolates were evaluated by PCR and 48 isolates were

identified that contained segments with 36 amino acids as well as strong fluorescence. These isolates were sequenced and 44 clones were identified that contained correct XTEN_AG36 segments. The file names of the nucleotide and amino acid constructs and the sequences for these segments are listed in Table 16.

Table 16: DNA and Amino Acid Sequences for AG 36-mer motifs (SEQ ID NOS 441-528, respectively, in order of appearance)

File name	Amino acid sequence	Nucleotide sequence
LCW0404_001_ GFP-N_A07.ab1	GASPGTSSTGSPGTPGS GTASSSPGSSTPSGATG SP	GGTGCATCCCCGGGCACTAGCTCTACCGGTTCTCCA GGTACTCCTGGTAGCGGTACTGCTTCTTCTTCTCCAG GTAGCTCTACTCCTTCTGGTGCTACTGGTTCTCCA
LCW0404_003_ GFP-N_B07.ab1	GSSTPSGATGSPGSSPS ASTGTGPGSSTPSGATG SP	GGTAGCTCTACCCCTTCTGGTGCTACCGGCTCTCCAG GTTCTAGCCCGTCTGCTTCTACCGGTACCGGTCCAGG TAGCTCTACCCCTTCTGGTGCTACTGGTTCTCCA
LCW0404_006_ GFP-N_C07.ab1	GASPGTSSTGSPGSSPS ASTGTGPGSSTPSGATG SP	GGTGCATCTCCGGGTACTAGCTCTACCGGTTCTCCAG GTTCTAGCCCTTCTGCTTCCACTGGTACCGGCCCAGG TAGCTCTACCCCGTCTGGTGCTACTGGTTCCCCA
LCW0404_007_ GFP-N_D07.ab1	GTPGSGTASSSPGSSTPS GATGSPGASPGTSSTGS P	GGTACTCCGGGCAGCGGTACTGCTTCTTCCTCTCCAG GTAGCTCTACCCCTTCTGGTGCAACTGGTTCCCCAGG TGCATCCCCTGGTACTAGCTCTACCGGTTCTCCA
LCW0404_009_ GFP-N_E07.ab1	GTPGSGTASSSPGASPG TSSTGSPGSRPSASTGT GP	GGTACCCCTGGCAGCGGTACTGCTTCTTCTTCTCCAG GTGCTTCCCCTGGTACCAGCTCTACCGGTTCTCCAGG TTCTAGACCTTCTGCATCCACCGGTACTGGTCCA
LCW0404_011_ GFP-N_F07.ab1	GASPGTSSTGSPGSSTPS GATGSPGASPGTSSTGS P	GGTGCATCTCCTGGTACCAGCTCTACCGGTTCTCCAG GTAGCTCTACTCCTTCTGGTGCTACTGGCTCTCCAGG TGCTTCCCCGGGTACCAGCTCTACCGGTTCTCCA
LCW0404_012_ GFP-N_G07.ab1	GTPGSGTASSSPGSSTPS GATGSPGSSTPSGATGS P	GGTACCCCGGGCAGCGGTACCGCATCTTCCTCTCCA GGTAGCTCTACCCCGTCTGGTGCTACCGGTTCCCCAG GTAGCTCTACCCCGTCTGGTGCAACCGGCTCCCCA
LCW0404_014_ GFP-N_H07.ab1	GASPGTSSTGSPGASPG TSSTGSPGASPGTSSTGS P	GGTGCATCTCCGGGCACTAGCTCTACTGGTTCTCCAG GTGCATCCCCTGGCACTAGCTCTACTGGTTCTCCAGG TGCTTCTCCTGGTACCAGCTCTACTGGTTCTCCA
LCW0404_015_ GFP-N_A08.ab1	GSSTPSGATGSPGSSPS ASTGTGPGASPGTSSTG SP	GGTAGCTCTACTCCGTCTGGTGCAACCGGCTCCCCA GGTTCTAGCCCGTCTGCTTCCACTGGTACTGGCCCAG GTGCTTCCCCGGGCACCAGCTCTACTGGTTCTCCA
LCW0404_016_ GFP-N_B08.ab1	GSSTPSGATGSPGSSTPS GATGSPGTPGSGTASSS P	GGTAGCTCTACTCCTTCTGGTGCTACCGGTTCCCCAG GTAGCTCTACTCCTTCTGGTGCTACTGGTTCCCCAGG TACTCCGGGCAGCGGTACTGCTTCTTCCTCTCCA
LCW0404_017_ GFP-N_C08.ab1	GSSTPSGATGSPGSSTPS GATGSPGASPGTSSTGS P	GGTAGCTCTACTCCGTCTGGTGCAACCGGTTCCCCAG GTAGCTCTACTCCTTCTGGTGCTACTGGCTCCCCAGG TGCATCCCCTGGCACCAGCTCTACCGGTTCTCCA
LCW0404_018_ GFP-N_D08.ab1	GTPGSGTASSSPGSSPS ASTGTGPGSSTPSGATG SP	GGTACTCCTGGTAGCGGTACCGCATCTTCCTCTCCAG GTTCTAGCCCTTCTGCATCTACCGGTACCGGTCCAGG TAGCTCTACTCCTTCTGGTGCTACTGGCTCTCCA
LCW0404_023_ GFP-N_F08.ab1	GASPGTSSTGSPGSSPS ASTGTGPGTPGSGTASS SP	GGTGCTTCCCCGGGCACTAGCTCTACCGGTTCTCCAG GTTCTAGCCCTTCTGCATCTACTGGTACTGGCCCAGG TACTCCGGGCAGCGGTACTGCTTCTTCCTCTCCA
LCW0404_025_ GFP-N_G08.ab1	GSSTPSGATGSPGSSTPS GATGSPGASPGTSSTGS P	GGTAGCTCTACTCCGTCTGGTGCTACCGGCTCTCCAG GTAGCTCTACCCCTTCTGGTGCAACCGGCTCCCCAGG TGCTTCTCCGGGTACCAGCTCTACTGGTTCTCCA
LCW0404_029_ GFP-N_A09.ab1	GTPGSGTASSSPGSSTPS GATGSPGSSPSASTGTG P	GGTACCCCTGGCAGCGGTACCGCTTCTTCCTCTCCAG GTAGCTCTACCCCGTCTGGTGCTACTGGCTCTCCAGG TTCTAGCCCGTCTGCATCTACCGGTACCGGCCCA
LCW0404_030_ GFP-N_B09.ab1	GSSTPSGATGSPGTPGS GTASSSPGTPGSGTASS SP	GGTAGCTCTACTCCTTCTGGTGCAACCGGCTCCCCAG GTACCCCGGGCAGCGGTACCGCATCTTCCTCTCCAG GTACTCCGGGTAGCGGTACTGCTTCTTCTTCTCCA
LCW0404_031_ GFP-N_C09.ab1	GTPGSGTASSSPGSSTPS GATGSPGASPGTSSTGS	GGTACCCCGGGTAGCGGTACTGCTTCTTCCTCTCCAG GTAGCTCTACCCCTTCTGGTGCAACCGGCTCTCCAGG

File name	Amino acid sequence	Nucleotide sequence
	P	TGCTTCTCCGGGCACCAGCTCTACCGGTTCTCCA
LCW0404_034_GFP-N_D09.ab1	GSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSP	GGTAGCTCTACCCCGTCTGGTGCTACCGGCTCTCCAGGTAGCTCTACCCCGTCTGGTGCAACCGGCTCCCCAGGTGCATCCCCGGGTACTAGCTCTACCGGTTCTCCA
LCW0404_035_GFP-N_E09.ab1	GASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSP	GGTGCTTCTCCGGGCACCAGCTCTACTGGTTCTCCAGGTACCCCGGGCAGCGGTACCGCATCTTCTTCTCCAGGTAGCTCTACTCCTTCTGGTGCAACTGGTTCTCCA
LCW0404_036_GFP-N_F09.ab1	GSSPSASTGTGPGSSTPSGATGSPGTPGSGTASSSP	GGTTCTAGCCCGTCTGCTTCCACCGGTACTGGCCCAGGTAGCTCTACCCCGTCTGGTGCAACTGGTTCCCCAGGTACCCCTGGTAGCGGTACCGCTTCTTCTTCTCCA
LCW0404_037_GFP-N_G09.ab1	GASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSP	GGTGCTTCTCCGGGCACCAGCTCTACTGGTTCTCCAGGTTCTAGCCCTTCTGCATCCACCGGTACCGGTCCAGGTAGCTCTACCCCTTCTGGTGCAACCGGCTCTCCA
LCW0404_040_GFP-N_H09.ab1	GASPGTSSTGSPGSSTPSGATGSPGSSTPSGATGSP	GGTGCATCCCCGGGCACCAGCTCTACCGGTTCTCCAGGTAGCTCTACCCCGTCTGGTGCTACCGGCTCTCCAGTAGCTCTACCCCGTCTGGTGCTACTGGCTCTCCA
LCW0404_041_GFP-N_A10.ab1	GTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSP	GGTACCCCTGGTAGCGGTACTGCTTCTTCCTCTCCAGGTAGCTCTACTCCGTCTGGTGCTACCGGTTCTCCAGGTACCCCGGGTAGCGGTACCGCATCTTCTTCTCCA
LCW0404_043_GFP-N_C10.ab1	GSSPSASTGTGPGSSTPSGATGSPGSSTPSGATGSP	GGTTCTAGCCCTTCTGCTTCCACCGGTACTGGCCCAGGTAGCTCTACCCCTTCTGGTGCTACCGGCTCCCCAGGTAGCTCTACTCCTTCTGGTGCAACTGGCTCTCCA
LCW0404_045_GFP-N_D10.ab1	GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGP	GGTGCTTCTCCTGGCACCAGCTCTACTGGTTCTCCAGGTTCTAGCCCTTCTGCTTCTACCGGTACTGGTCCAGGTTCTAGCCCTTCTGCATCCACTGGTACTGGTCCA
LCW0404_047_GFP-N_F10.ab1	GTPGSGTASSSPGASPGTSSTGSPGASPGTSSTGSP	GGTACTCCTGGCAGCGGTACCGCTTCTTCTTCTCCAGGTGCTTCTCCTGGTACTAGCTCTACTGGTTCTCCAGGTGCTTCTCCGGGCACTAGCTCTACTGGTTCTCCA
LCW0404_048_GFP-N_G10.ab1	GSSTPSGATGSPGASPGTSSTGSPGSSTPSGATGSP	GGTAGCTCTACCCCGTCTGGTGCTACCGGTTCCCCAGGTGCTTCTCCTGGTACTAGCTCTACCGGTTCTCCAGGTAGCTCTACCCCGTCTGGTGCTACTGGCTCTCCA
LCW0404_049_GFP-N_H10.ab1	GSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSP	GGTAGCTCTACCCCGTCTGGTGCTACTGGTTCTCCAGGTACTCCGGGCAGCGGTACTGCTTCTTCCTCTCCAGGTAGCTCTACCCCTTCTGGTGCTACTGGCTCTCCA
LCW0404_050_GFP-N_A11.ab1	GASPGTSSTGSPGSSPSASTGTGPGSSTPSGATGSP	GGTGCATCTCCTGGTACCAGCTCTACTGGTTCTCCAGGTTCTAGCCCTTCTGCTTCTACCGGTACCGGTCCAGGTAGCTCTACTCCTTCTGGTGCTACCGGTTCTCCA
LCW0404_051_GFP-N_B11.ab1	GSSTPSGATGSPGSSTPSGATGSPGSSTPSGATGSP	GGTAGCTCTACCCCGTCTGGTGCTACTGGCTCTCCAGGTAGCTCTACTCCTTCTGGTGCTACTGGTTCCCCAGGTAGCTCTACCCCGTCTGGTGCAACTGGCTCTCCA
LCW0404_052_GFP-N_C11.ab1	GASPGTSSTGSPGTPGSGTASSSPGASPGTSSTGSP	GGTGCATCCCCGGGTACCAGCTCTACCGGTTCTCCAGGTACTCCTGGCAGCGGTACTGCATCTTCCTCTCCAGGTGCTTCTCCGGGCACCAGCTCTACTGGTTCTCCA
LCW0404_053_GFP-N_D11.ab1	GSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSP	GGTAGCTCTACTCCTTCTGGTGCAACTGGTTCTCCAGGTTCTAGCCCGTCTGCATCCACTGGTACCGGTCCAGGTGCTTCCCCTGGCACCAGCTCTACCGGTTCTCCA
LCW0404_057_GFP-N_E11.ab1	GASPGTSSTGSPGSSTPSGATGSPGSSPSASTGTGP	GGTGCATCTCCTGGTACTAGCTCTACTGGTTCTCCAGGTAGCTCTACTCCGTCTGGTGCAACCGGCTCTCCAGGTTCTAGCCCTTCTGCATCTACCGGTACTGGTCCA
LCW0404_060_GFP-N_F11.ab1	GTPGSGTASSSPGSSTPSGATGSPGASPGTSSTGSP	GGTACTCCTGGCAGCGGTACCGCATCTTCCTCTCCAGGTAGCTCTACTCCGTCTGGTGCAACTGGTTCCCCAGGTGCTTCTCCGGGTACCAGCTCTACCGGTTCTCCA
LCW0404_062_GFP-N_G11.ab1	GSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSP	GGTAGCTCTACCCCGTCTGGTGCAACCGGCTCCCCAGGTACTCCTGGTAGCGGTACCGCTTCTTCTTCTCCAGGTAGCTCTACTCCGTCTGGTGCTACCGGCTCCCCA
LCW0404_066_GFP-N_H11.ab1	GSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSP	GGTTCTAGCCCTTCTGCATCCACCGGTACCGGCCCAGGTTCTAGCCCGTCTGCTTCTACCGGTACTGGTCCAGGTGCTTCTCCGGGTACTAGCTCTACTGGTTCTCCA
LCW0404_067_GFP-N_A12.ab1	GTPGSGTASSSPGSSTPSGATGSPGSNPSASTGTGP	GGTACCCCGGGTAGCGGTACCGCTTCTTCTTCTCCAGGTAGCTCTACTCCGTCTGGTGCTACCGGCTCTCCAGGTTCTAACCCTTCTGCATCCACCGGTACCGGCCCA

File name	Amino acid sequence	Nucleotide sequence
LCW0404_068_ GFP-N_B12.ab1	GSSPSASTGTGPGSSTPS GATGSPGASPGTSSTGS P	GGTTCTAGCCCTTCTGCATCTACTGGTACTGGCCCAG GTAGCTCTACTCCTTCTGGTGCTACCGGCTCTCCAGG TGCTTCTCCGGGTACTAGCTCTACCGGTTCTCCA
LCW0404_069_ GFP-N_C12.ab1	GSSTPSGATGSPGASPG TSSTGSPGTPGSGTASSS P	GGTAGCTCTACCCCTTCTGGTGCAACCGGCTCTCCAG GTGCATCCCCGGGTACCAGCTCTACCGGTTCTCCAG GTACTCCGGGTAGCGGTACCGCTTCTTCCTCTCCA
LCW0404_070_ GFP-N_D12.ab1	GSSTPSGATGSPGSSTPS GATGSPGSSTPSGATGS P	GGTAGCTCTACTCCGTCTGGTGCAACCGGTTCCCCAG GTAGCTCTACCCCTTCTGGTGCAACCGGCTCCCCAGG TAGCTCTACCCCTTCTGGTGCAACTGGCTCTCCA
LCW0404_073_ GFP-N_E12.ab1	GASPGTSSTGSPGTPGS GTASSSPGSSTPSGATG SP	GGTGCTTCTCCTGGCACTAGCTCTACCGGTTCTCCAG GTACCCCTGGTAGCGGTACCGCATCTTCCTCTCCAGG TAGCTCTACTCCTTCTGGTGCTACTGGTTCCCCA
LCW0404_075_ GFP-N_F12.ab1	GSSTPSGATGSPGSSPS ASTGTGPGSSPSASTGT GP	GGTAGCTCTACCCCGTCTGGTGCTACTGGCTCCCCAG GTTCTAGCCCTTCTGCATCCACCGGTACCGGTCCAGG TTCTAGCCCGTCTGCATCTACTGGTACTGGTCCA
LCW0404_080_ GFP-N_G12.ab1	GASPGTSSTGSPGSSPS ASTGTGPGSSPSASTGT GP	GGTGCTTCCCCGGGCACCAGCTCTACTGGTTCTCCAG GTTCTAGCCCGTCTGCTTCTACTGGTACTGGTCCAGG TTCTAGCCCTTCTGCTTCCACTGGTACTGGTCCA
LCW0404_081_ GFP-N_H12.ab1	GASPGTSSTGSPGSSPS ASTGTGPGTPGSGTASS SP	GGTGCTTCCCCGGGTACCAGCTCTACCGGTTCTCCAG GTTCTAGCCCTTCTGCTTCTACCGGTACCGGTCCAGG TACCCCTGGCAGCGGTACCGCATCTTCCTCTCCA

[00391] Example 5: Construction of XTEN_AE864

[00392] XTEN_AE864 was constructed from serial dimerization of XTEN_AE36 to AE72, 144, 288, 576 and 864. A collection of XTEN_AE72 segments was constructed from 37 different segments of XTEN_AE36. Cultures of E. coli harboring all 37 different 36-amino acid segments were mixed and plasmid was isolated. This plasmid pool was digested with BsaI/NcoI to generate the small fragment as the insert. The same plasmid pool was digested with BbsI/NcoI to generate the large fragment as the vector. The insert and vector fragments were ligated resulting in a doubling of the length and the ligation mixture was transformed into BL21Gold(DE3) cells to obtain colonies of XTEN_AE72.

[00393] This library of XTEN_AE72 segments was designated LCW0406. All clones from LCW0406 were combined and dimerized again using the same process as described above yielding library LCW0410 of XTEN_AE144. All clones from LCW0410 were combined and dimerized again using the same process as described above yielding library LCW0414 of XTEN_AE288. Two isolates LCW0414.001 and LCW0414.002 were randomly picked from the library and sequenced to verify the identities. All clones from LCW0414 were combined and dimerized again using the same process as described above yielding library LCW0418 of XTEN_AE576. We screened 96 isolates from library LCW0418 for high level of GFP fluorescence. 8 isolates with right sizes of inserts by PCR and strong fluorescence were sequenced and 2 isolates (LCW0418.018 and LCW0418.052) were chosen for future use based on sequencing and expression data.

[00394] The specific clone pCW0432 of XTEN_AE864 was constructed by combining LCW0418.018 of XTEN_AE576 and LCW0414.002 of XTEN_AE288 using the same dimerization process as described above.

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 180

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 180

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHIER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

CLAIMS

WHAT IS CLAIMED IS:

1. A recombinant factor VIII (FVIII) fusion protein comprising a FVIII polypeptide and a first extended recombinant polypeptide (XTEN),
wherein the FVIII polypeptide comprises an A3 domain, an A2 domain, an A1 domain, a C2 domain, a C1 domain, and a B domain or a portion of the B domain,
wherein the first XTEN is inserted into the FVIII polypeptide at a site immediately downstream of amino acid 745 of the sequence set forth in SEQ ID NO:2; and
wherein the recombinant FVIII fusion protein exhibits a prolonged terminal half-life when administered to a subject as compared to a corresponding FVIII polypeptide lacking the first XTEN.

2. The recombinant FVIII fusion protein of claim 1, wherein an FVIII processing site corresponding to amino acid residue R1648 of SEQ ID NO: 2 is eliminated in the FVIII polypeptide.

3. The recombinant FVIII fusion protein of claim 1, wherein the FVIII polypeptide is a single chain factor VIII polypeptide.

4. The recombinant FVIII fusion protein of any one of claims 1 to 3, wherein the first XTEN is at least 36 amino acids and comprises any combination of XTEN motif sequences, wherein the XTEN motif sequences comprise an amino acid sequence having at least 95% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs: 19-48, and any combination thereof.

5. The recombinant FVIII fusion protein of any one of claims 1 to 3,
wherein the first XTEN comprises an amino acid sequence at least 95% identical to an amino acid sequence selected from the group consisting of SEQ ID NOs: 50, 52, 67, 78, and any combination thereof.

6. The recombinant FVIII fusion protein of any one of claims 1 to 3, wherein the first XTEN is at least 36 amino acids and comprises any combination of XTEN motif sequences,

Date Reçue/Date Received 2022-02-04

wherein the XTEN motif sequences comprise an amino acid sequence having at least 95% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOs: 23, 24, 25, 26, and any combination thereof.

7. The recombinant FVIII fusion protein of claim 6, wherein the XTEN motif sequences comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 23, 24, 25, 26, and any combination thereof.

8. The recombinant FVIII fusion protein of claim 7, wherein the sequence motifs comprise the amino acid sequence set forth in SEQ ID NO: 23.

9. The recombinant FVIII fusion protein of claim 7, wherein the XTEN sequence motifs comprise the amino acid sequence set forth in SEQ ID NO: 24.

10. The recombinant FVIII fusion protein of claim 7, wherein the XTEN sequence motifs comprise the amino acid sequence set forth in SEQ ID NO: 25.

11. The recombinant FVIII fusion protein of claim 7, wherein the XTEN sequence motifs comprise the amino acid sequence set forth in SEQ ID NO: 26.

12. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 49, 50, 51, 52, 57, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 78, 79, 279, 281, 283, 285, 287, 289, 291, 293, 295, 297, 299, 301, 303, 305, 307, 309, 311, 313, 315, 317, 319, 321, 323, 325, 327, 329, 331, 333, 335, 337, 339, 341, 343, 345, 347, 349, 351, and any combination thereof.

13. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 50, 52, 67, and 78.

14. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises an amino acid sequence at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 50.

15. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises the amino acid sequence set forth in SEQ ID NO: 50.

Date Reçue/Date Received 2022-02-04

16. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises an amino acid sequence at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 52.

17. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises the amino acid sequence set forth in SEQ ID NO: 52.

18. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises an amino acid sequence at least 95% identical to SEQ ID NO: 67.

19. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises the amino acid sequence set forth in SEQ ID NO: 67.

20. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises an amino acid sequence at least 95% identical to SEQ ID NO:78.

21. The recombinant FVIII fusion protein of claim 1, wherein the first XTEN comprises the amino acid sequence set forth in SEQ ID NO:78.

22. The recombinant FVIII fusion protein of any one of claims 1 to 21, further comprising a Fc fragment of immunoglobulin or an FcRn binding domain.

23. The recombinant FVIII fusion protein of any one of claims 1 to 22, wherein the FVIII polypeptide comprises a deletion of amino acids corresponding to amino acids 747-1638 of SEQ ID NO: 2.

24. The recombinant FVIII fusion protein of any one of claims 1 to 23, wherein the FVIII polypeptide comprises amino acids 1-745 of SEQ ID NO: 2.

25. The recombinant FVIII fusion protein of any one of claims 1 to 24, wherein the FVIII polypeptide comprises amino acids 1649-2332 of SEQ ID NO: 2.

26. A recombinant factor VIII fusion protein comprising a factor VIII polypeptide fused to an extended recombinant polypeptide (XTEN) and a Fc fragment of immunoglobulin;

Date Reçue/Date Received 2022-02-04

wherein the factor VIII polypeptide comprises amino acid residues 1-745 and 1649-2332 of mature human factor VIII (SEQ ID NO: 2);

wherein an FVIII processing site corresponding to amino acid residue R1648 of SEQ ID NO: 2 is eliminated in the FVIII polypeptide;

wherein the XTEN is inserted within the factor VIII polypeptide immediately downstream of an amino acid corresponding to residue 745 of SEQ ID NO: 2; and

wherein the XTEN comprises the amino acid sequence set forth in SEQ ID NO: 78.

27. An isolated nucleic acid encoding the recombinant FVIII fusion protein of any one of claims 1 to 26.

28. A host cell comprising an expression vector comprising the nucleic acid of claim 27.

29. A method of making the recombinant FVIII fusion protein of any one of claims 1 to 26, wherein the method comprises culturing the host cell of claim 28 in media under conditions suitable for making the fusion protein and recovering the fusion protein.

30. A pharmaceutical composition comprising (i) the recombinant factor VIII fusion protein of any one of claims 1 to 26 and (ii) a pharmaceutically acceptable carrier.

31. Use of the recombinant FVIII fusion protein of any one of claims 1 to 26, for treating a bleeding disease or condition in a subject in need thereof.

32. Use of the pharmaceutical composition of claim 30 for treating a bleeding disease or condition in a subject in need thereof.

33. The use of claim 31 or 32, wherein the bleeding disease or condition is a bleeding episode.

34. The use of claim 31 or 32, wherein the bleeding disease or condition is a coagulopathy.

35. The use of any one of claims 31-34, wherein the bleeding disease or condition is a result of trauma or surgery.

Date Reçue/Date Received 2022-02-04

36. The use of any one of claims 31-35 wherein the pharmaceutical composition or the recombinant FVIII fusion protein is suitable for prophylactic treatment.

37. The use of claim 31 or 32, wherein the bleeding disease or condition is hemophilia A.

38. The use of claim 37 wherein the hemophilia A is mild hemophilia A.

39. The use of claim 37 wherein the hemophilia A is moderate hemophilia A.

40. The use of claim 37 wherein the hemophilia A is severe hemophilia A.

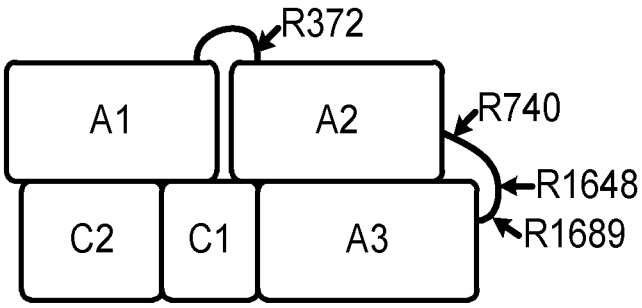
41. The use of any one of claims 31 to 40, wherein the subject is a male.

42. The use of any one of claims 31 to 40, wherein the subject is a female.

Date Reçue/Date Received 2022-02-04

FIG. 1A

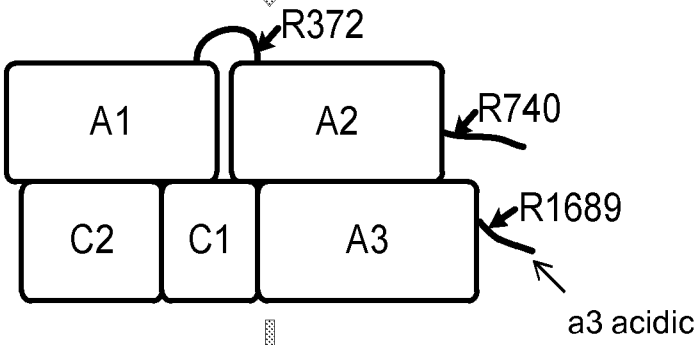
scFVIII



Intracellular processing

FIG. 1B

FVIII



Clotting event

FIG. 1C

FVIIIa

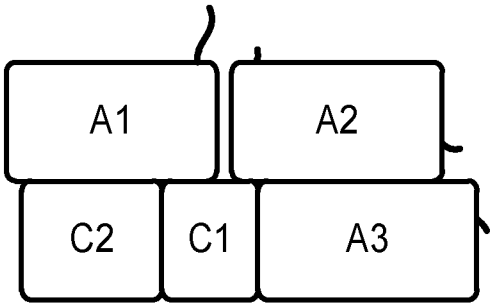


FIG. 1

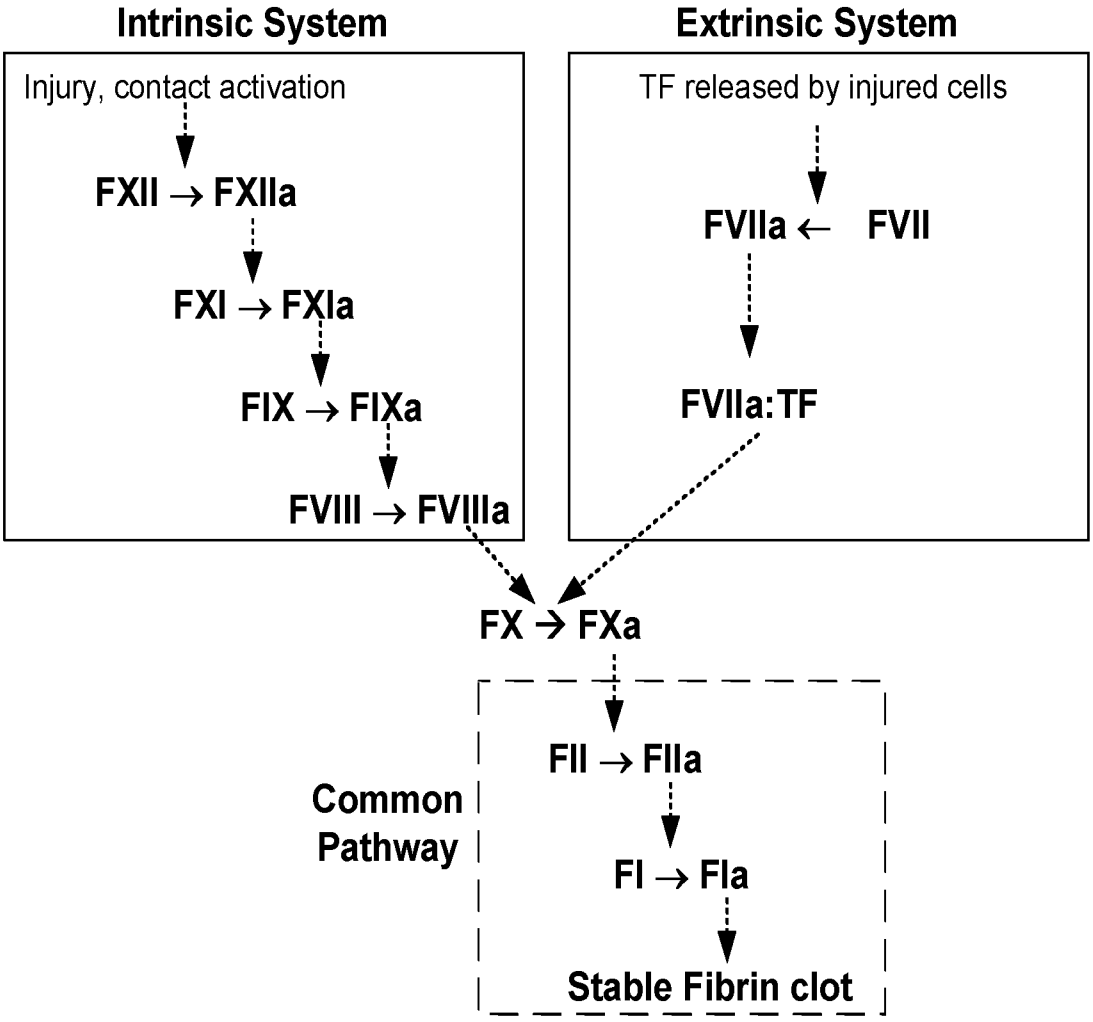

Intrinsic System
Injury, contact activation
FXII → FXIIa
FXI → FXIa
FIX → FIXa
FVIII → FVIIIa
Extrinsic System
TF released by injured cells
FVIIa ← FVII
FVIIa:TF
FX → FXa
Common Pathway
FII → FIIa
FI → FIa
Stable Fibrin clot

FIG. 2

ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFTDHLFN
IAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQ
REKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCR
EGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNR
SLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLL
MDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRF
DDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIG
RKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGI
TDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNME
RDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAG
VQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKH
KMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYE
DSYEDISAYLLSKNNAIEPRSFSQNSRHPSTRQKQFNATTIPENDIEKTDPWFAHRTPMP
KIQNVSSSDLLMLLRQSPTPHGLSLSDLQEAKYETFSDDPSPGAIDSNNSLSEMTHFRPQ
LHHSGDMVFTPESGLQLRLNEKLGTTAATELKKLDFKVSSTSNNLISTIPSDNLAAGTDN
TSSLGPPSMPVHYDSQLDTTLFGKKSSPLTESGGPLSLSEENNDSKLLESGLMNSQESSW
GKNVSSTESGRLFKGKRAHGPALLTKDNALFKVSISLLKTNKTSNNSATNRKTHIDGPSL
LIENSPSVWQNILESDTEFKKVTPLIHDRMLMDKNATALRLNHMSNKTTSSKNMEMVQQK
KEGPIPPDAQNPDMSFFKMLFLPESARWIQRTHGKNSLNSGQGPSPKQLVSLGPEKSVEG
QNFLSEKNKVVVGKGEFTKDVGLKEMVFPSSRNLFLTNLDNLHENNTHNQEKKIQEEIEK
KETLIQENVVLPQIHTVTGTKNFMKNLFLLSTRQNVEGSYDGAYAPVLQDFRSLNDSTNR
TKKHTAHFSKKGEEENLEGLGNQTKQIVEKYACTTRISPNTSQQNFVTQRSKRALKQFRL
PLEETELEKRIIVDDTSTQWSKNMKHLTPSTLTQIDYNEKEKGAITQSPLSDCLTRSHSI
PQANRSPLPIAKVSSFPSIRPIYLTRVLFQDNSSHLPAASYRKKDSGVQESSHFLQGAKK
NNLSLAILTLEMTGDQREVGSLGTSATNSVTYKKVENTVLPKPDLPKTSGKVELLPKVHI
YQKDLFPTETSNGSPGHLDLVEGSLLQGTEGAIKWNEANRPGKVPFLRVATESSAKTPSK
LLDPLAWDNHYGTQIPKEEWKSQEKSPEKTAFKKKDTILSLNACESNHAIAAINEGQNKP
EIEVTWAKQGRTERLCSQNPPVLKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIY
DEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTD
GSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGA
EPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHT
NTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHA
INGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYP
GVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITAS
GQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQ
FIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRS
TLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWR
PQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV
KVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

FIG. 3

ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFV
EFTVHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVS
YWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTY
SYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSW
HSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGM
GTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHI
SSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDN
SPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGP
QRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTLLIIFKNQ
ASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDG
PTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRN
VILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDS
LQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGE
TVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISA
YLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQSDQEEIDYDDTISVEMKK
EDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGS
VPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQ
ASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEF
DCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTI
FDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVM
AQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETV
EMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQIT
ASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGAR
QKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNP
PIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASS
YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQ
GVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLD
PPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

FIG. 4

Fig. 5A

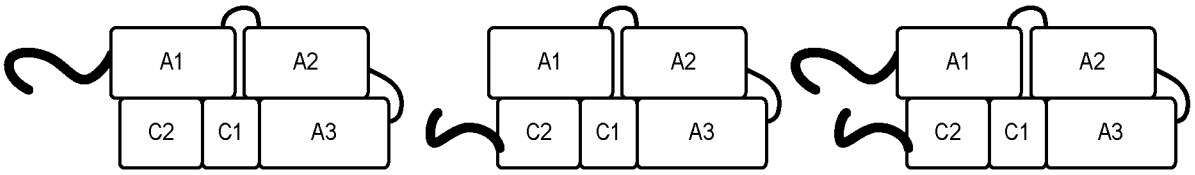


Fig. 5B

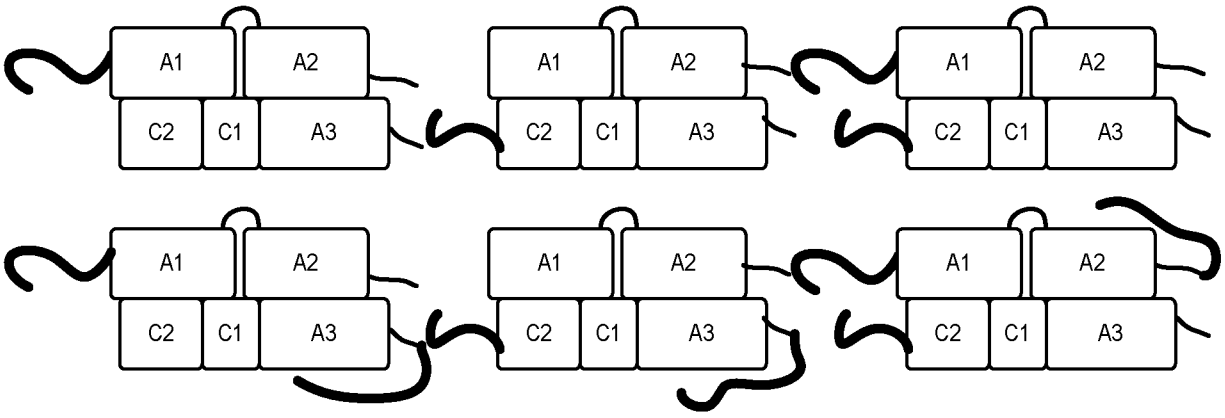


Fig. 5C

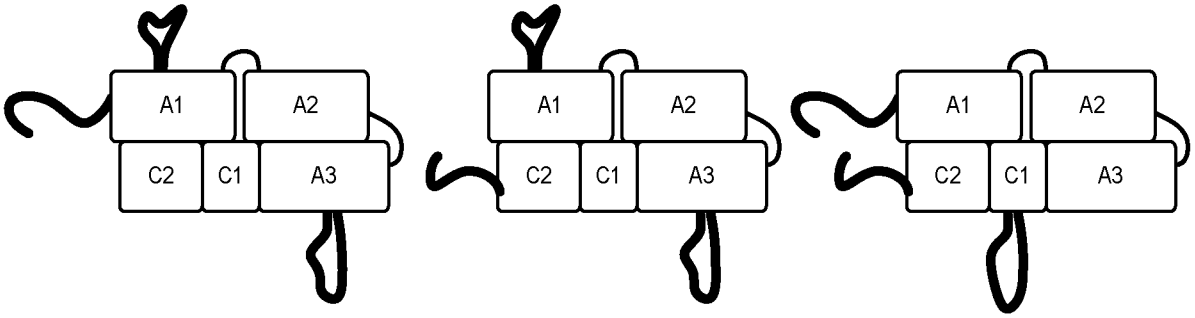


Fig. 5D

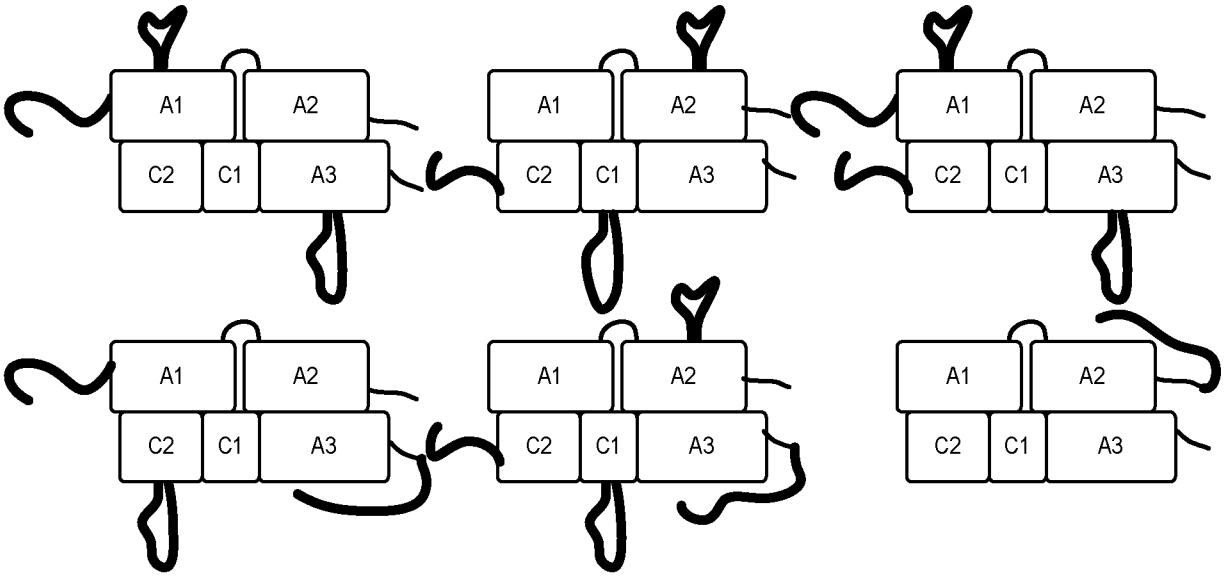


FIG. 5

Fig. 6A

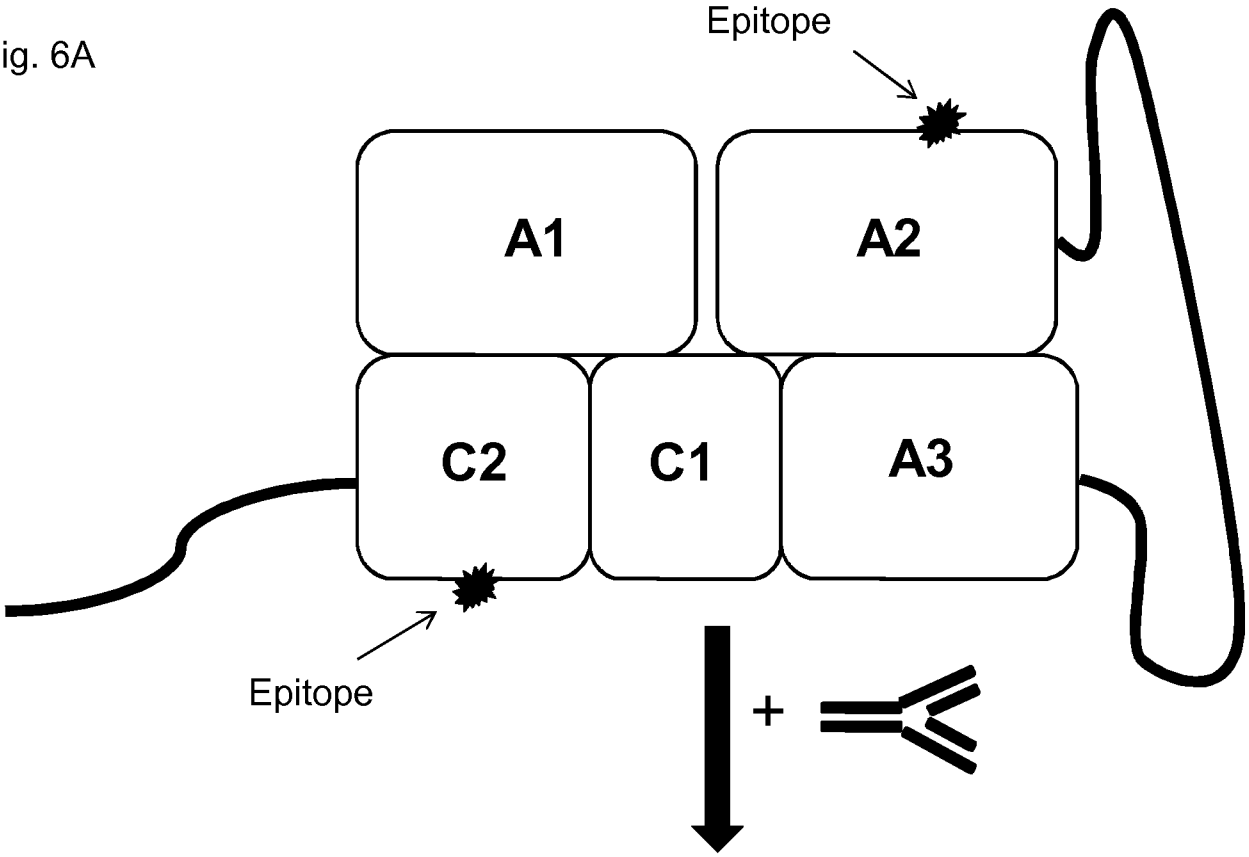


Fig. 6B

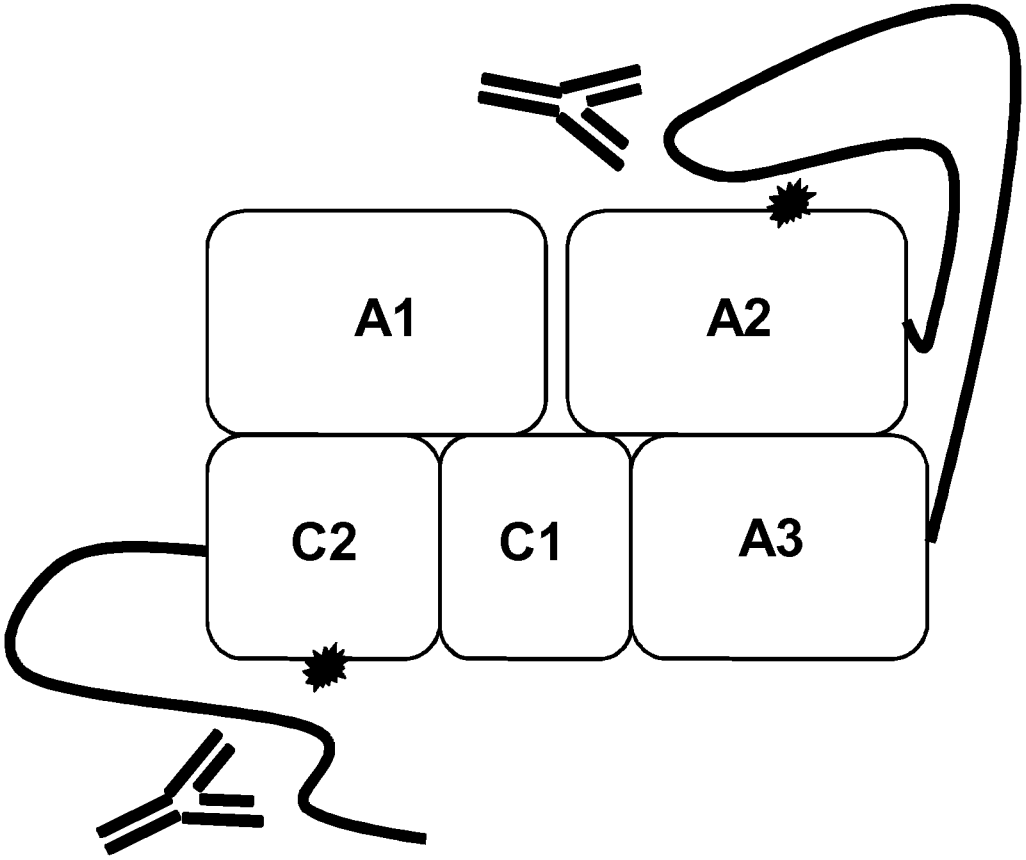


FIG. 6

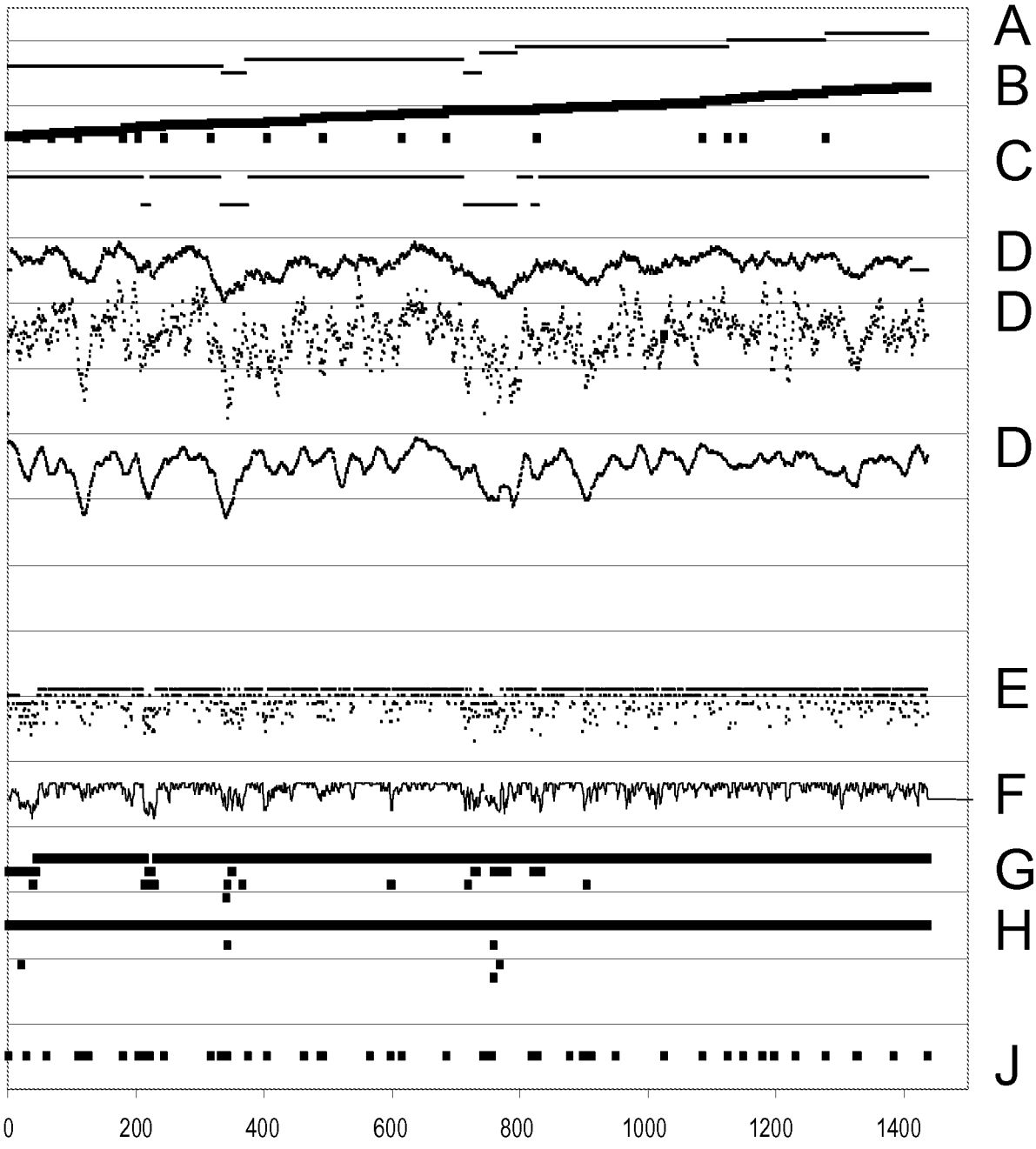


FIG. 7

ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTL
FVEFTVHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHA
VGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASD
PLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFA
VFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHR
KSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLL
MDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDL
TDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVL
APDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILG
PLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKD
FPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGP
LLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAG
VQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLS
VFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNR
GMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLK
RHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKK
TRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFT
QPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEE
DQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLE
KDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTEN
MERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYL
LSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAG
IWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYG
QWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFS
SLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPI
IARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASS
YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVT
TQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVV
NSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

FIG. 8

ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTL
FVEFTVHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHA
VGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASD
PLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFA
VFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHR
KSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLL
MDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDL
TDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVL
APDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILG
PLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKD
FPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGP
LLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAG
VQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLS
VFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNR
GMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLK
RHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKK
TRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFT
QPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEE
DQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLE
KDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTEN
MERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYL
LSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAG
IWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYG
QWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFS
SLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPI
IARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASS
YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVT
TQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVV
NSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

FIG. 9

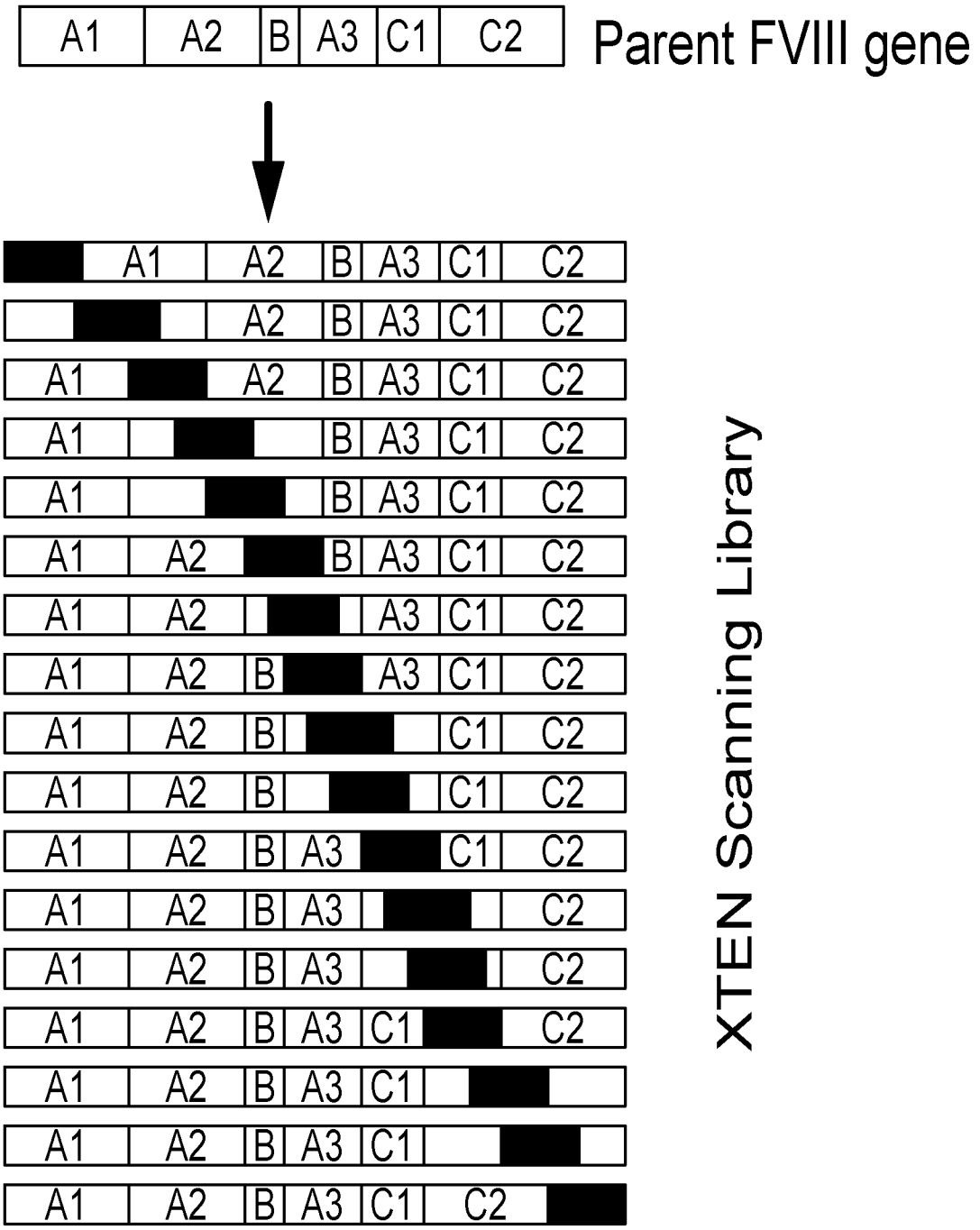


FIG. 10

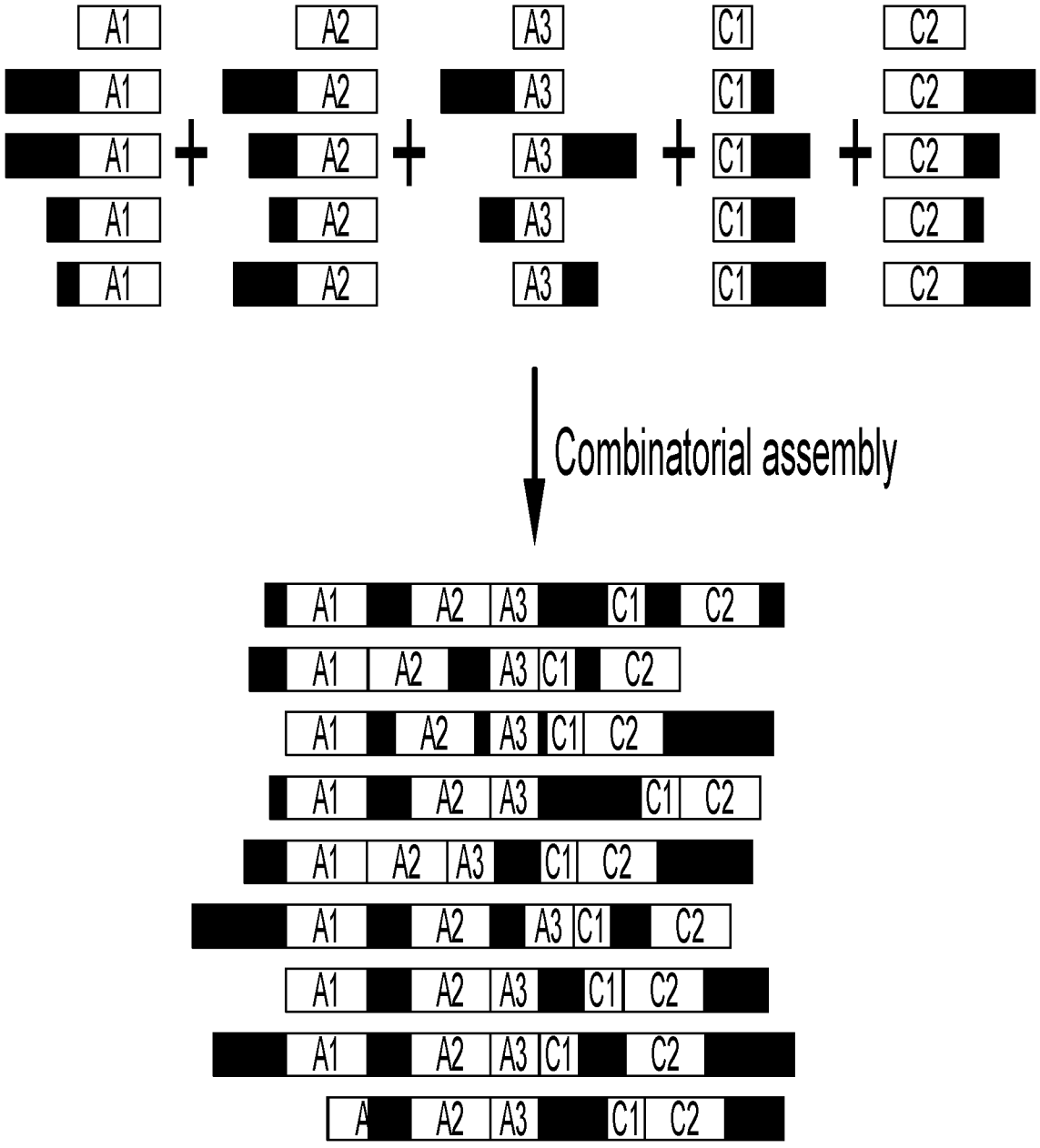


FIG. 11

FIG. 12A

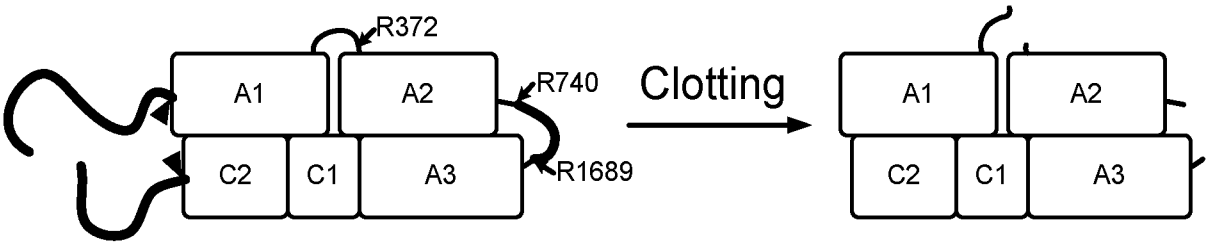


FIG. 12B

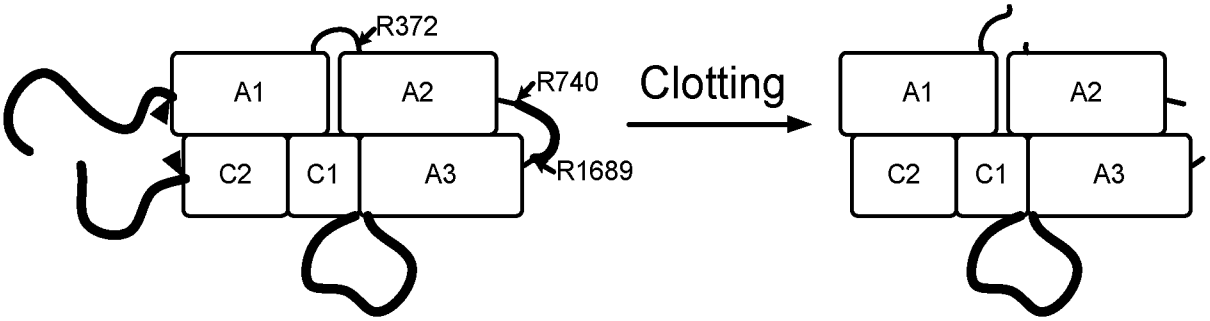


FIG. 12C

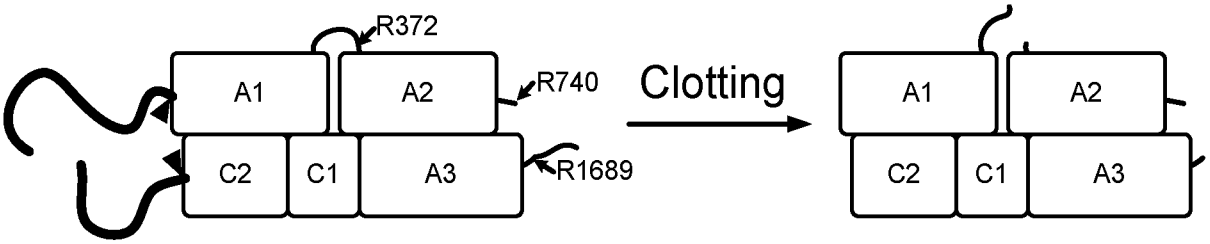


FIG. 12D

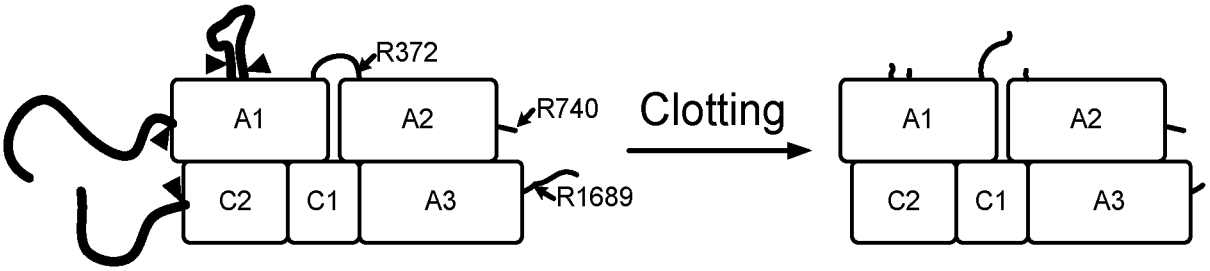


FIG. 12

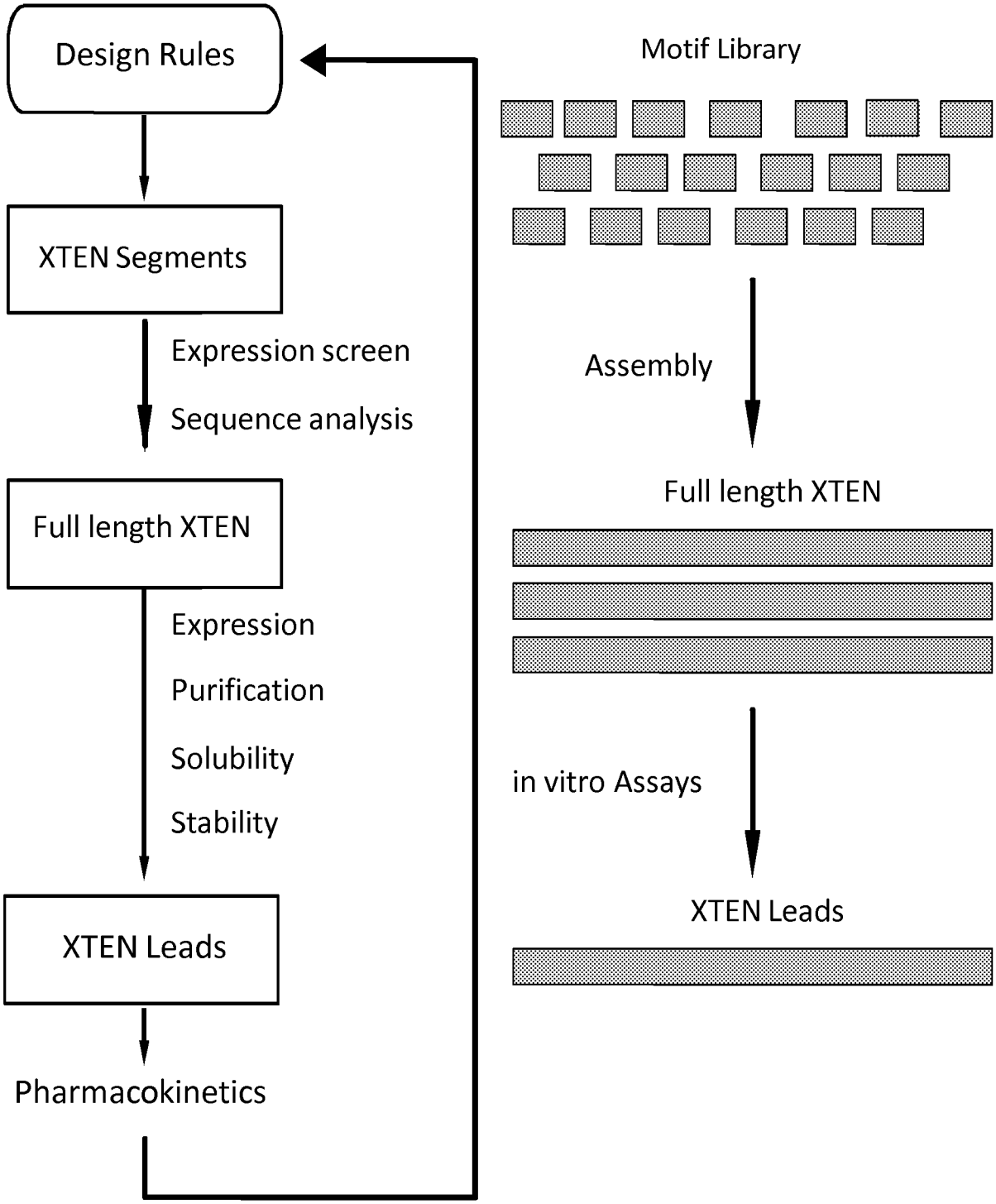

Design Rules
XTEN Segments
Expression screen
Sequence analysis
Full length XTEN
Expression
Purification
Solubility
Stability
XTEN Leads
Pharmacokinetics
Motif Library
Assembly
Full length XTEN
in vitro Assays
XTEN Leads

FIG. 13

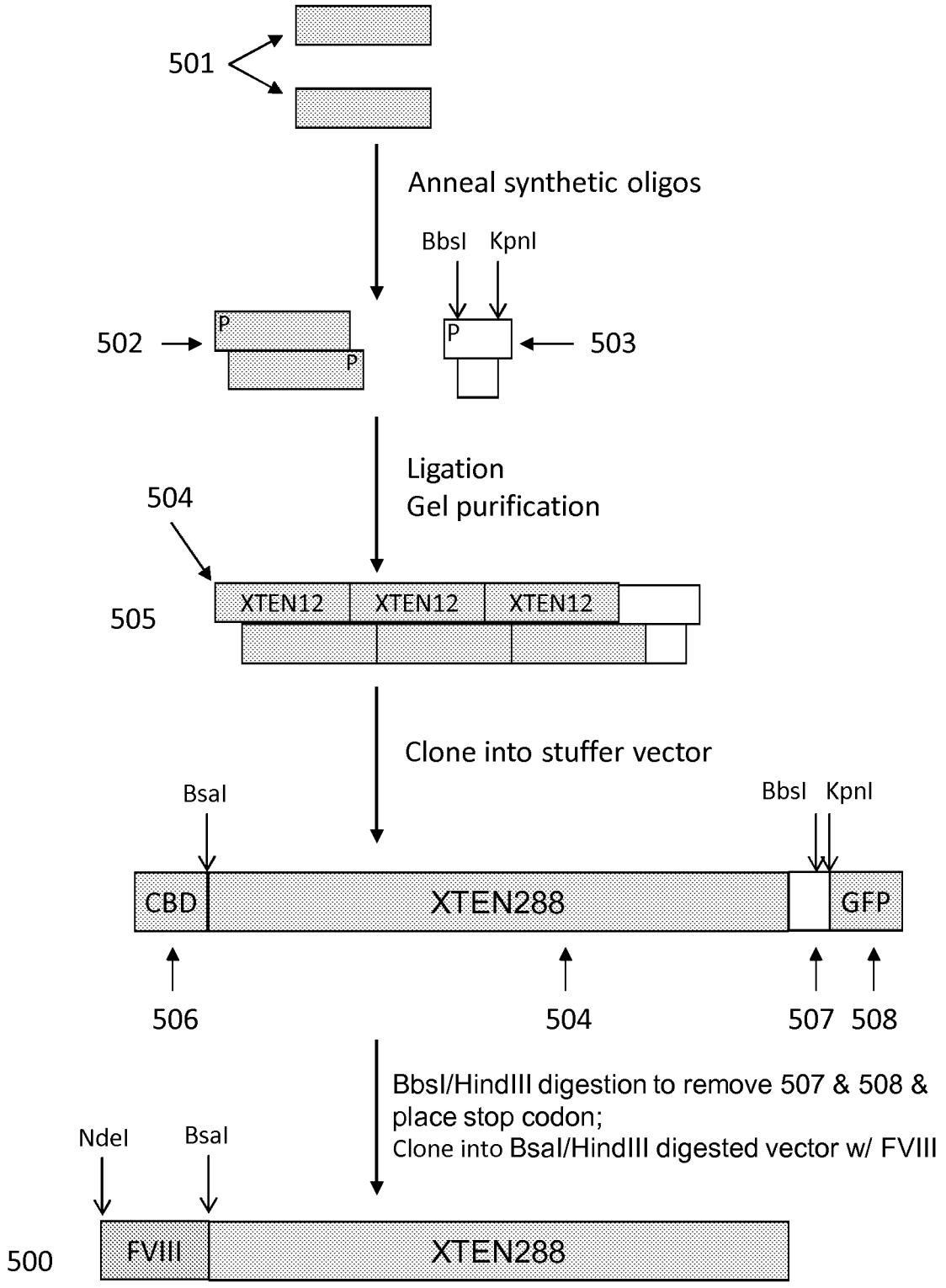

501
Anneal synthetic oligos
BbsI
KpnI
502
P
P
P
503
Ligation
Gel purification
504
505
XTEN12
XTEN12
XTEN12
Clone into stuffer vector
BsaI
BbsI
KpnI
CBD
XTEN288
GFP
506
504
507
508
BbsI/HindIII digestion to remove 507 & 508 &
place stop codon;
Clone into BsaI/HindIII digested vector w/ FVIII
NdeI
BsaI
500
FVIII
XTEN288

FIG. 14

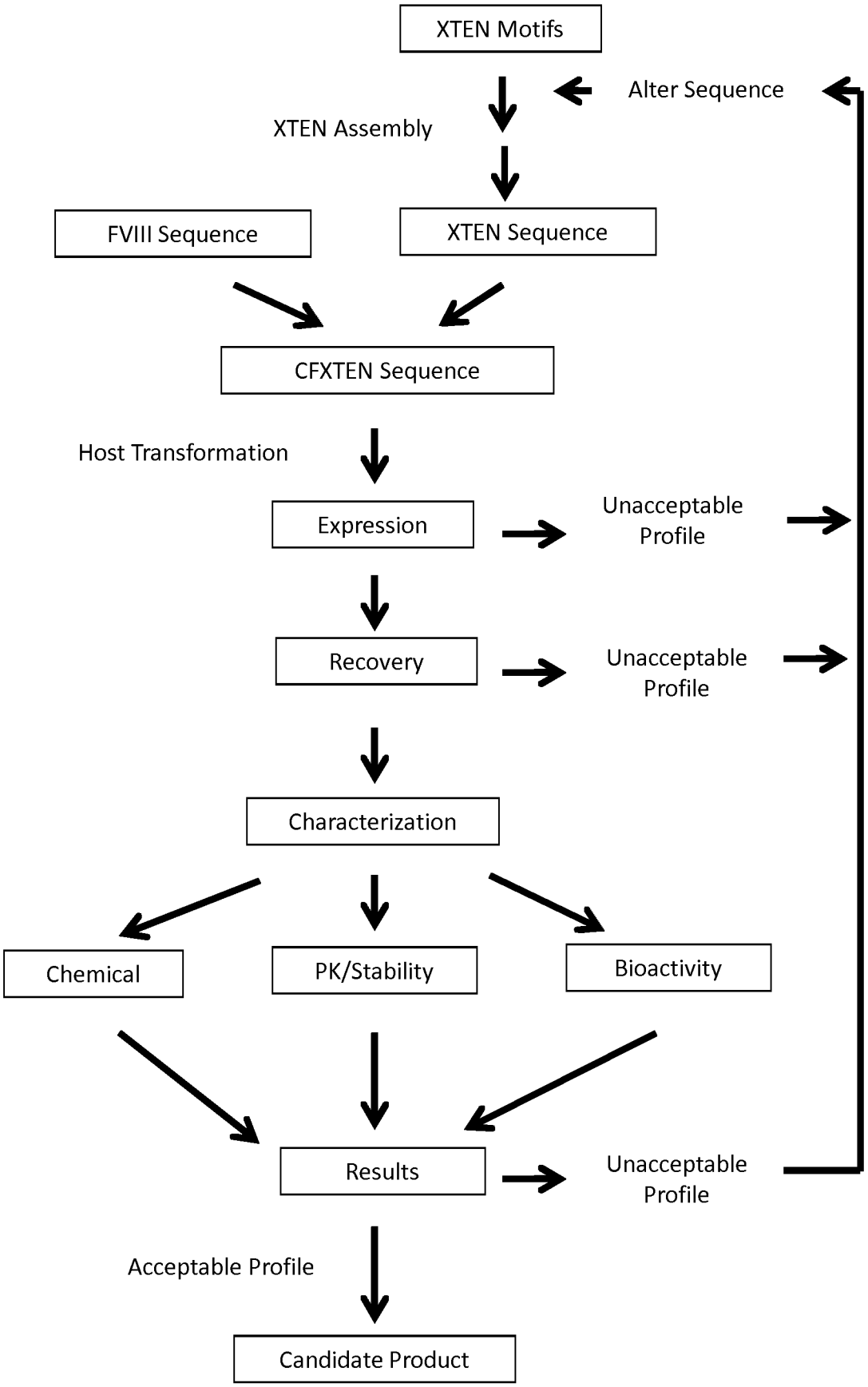

XTEN Motifs
Alter Sequence
XTEN Assembly
FVIII Sequence
XTEN Sequence
CFXTEN Sequence
Host Transformation
Expression
Unacceptable Profile
Recovery
Unacceptable Profile
Characterization
Chemical
PK/Stability
Bioactivity
Results
Unacceptable Profile
Acceptable Profile
Candidate Product

FIG. 15

AG864_2

GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGS
PGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS
SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSST
GSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGA
TGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGT
ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSG
TASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTP
SGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSST
PSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGAS
PGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGS
STPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPG
SSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSP
GSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGS
PGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASS
SPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSP

AG576

PGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGATG
SPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTAS
SSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTA
SSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTPSG
ATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSTPS
GATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGASPG
TSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSST
PSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSS
TPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGS
STPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGS

FIG. 16A

AG864_2

GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGS
PGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS
SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSST
GSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGA
TGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGT
ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSG
TASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTP
SGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSST
PSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGAS
PGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGS
STPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPG
SSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSP
GSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGS
PGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASS
SPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSP

AG288_1

PGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTA
SSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSSPS
ASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGS
SPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGS
PGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGA
TGS

FIG. 16B

AG864_2

GASPGTSSTGSPGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGS
PGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASS
SPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGASPGTSST
GSPGTPGSGTASSSPGSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGSSTPSGA
TGSPGSSTPSGATGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGT
ASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSPSASTGTGPGTPGSG
TASSSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGSSTPSGATGSPGSSTP
SGATGSPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGSSTPSGATGSPGSST
PSGATGSPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGAS
PGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGASPGTSSTGSPGTPGSGTASSSPGS
STPSGATGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGSSTPSGATGSPG
SSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGTPGSGTASSSP
GSSTPSGATGSPGSSPSASTGTGPGSSPSASTGTGPGASPGTSSTGSPGASPGTSSTGS
PGSSTPSGATGSPGSSPSASTGTGPGASPGTSSTGSPGSSPSASTGTGPGTPGSGTASS
SPGSSTPSGATGSPGSSTPSGATGSPGASPGTSSTGSP

AG144_2

PGSSPSASTGTGPGSSPSASTGTGPGTPGSGTASSSPGSSTPSGATGSPGSSPSAST
GTGPGASPGTSSTGSPGTPGSGTASSSPGSSTPSGATGSPGTPGSGTASSSPGASPG
TSSTGSPGASPGTSSTGSPGTPGSGTASSS

FIG. 16C

AE864

GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSA
PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTST
EEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG
SAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSG
SETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSP
TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEP
SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAG
SPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEP
ATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSP
AGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS
EPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG
SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEE
GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESG
PGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS
APGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP

AE576

GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG
SAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGS
PTSTEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTS
TEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGP
GSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPE
SGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEP
SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTS
TEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETP
GTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPE
SGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEP
SEGSAP

FIG. 16D

AE864

GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSA
PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTST
EEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEG
SAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEPATSG
SETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGSPAGSP
TSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEP
SEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAG
SPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEP
ATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSP
AGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGS
EPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG
SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEE
GSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGTSESATPESG
PGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGS
APGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP

AE288_2

GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPE
SGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESA
TPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTS
ESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEE
GTSTEPSEGSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEG
SAP

FIG. 16E

AE864

GSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGS
APGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPT
STEEGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEP
SEGSAPGTSTEPSEGSAPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSEP
ATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATPESGPGS
PAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP
GTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGS
APGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSESATP
ESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSPAGS
PTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSE
SATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGT
STEPSEGSAPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGP
GSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGTSESATPES
GPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSE
GSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAP

AE144_1A

SPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSA
PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTS
TEEGTSESATPESGPGTSTEPSEGSAPG

AE144_2B

TSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESG
PGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEG
SAPGTSESATPESGPGTSESATPESGPG

AE144_3A

SPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSA
PGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEG
SAPGSPAGSPTSTEEGTSTEPSEGSAPG

AE144_4B

TSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESG
PGTSTEPSEGSAPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTS
TEEGTSESATPESGPGTSTEPSEGSAPG

FIG. 16F

FIG. 17A

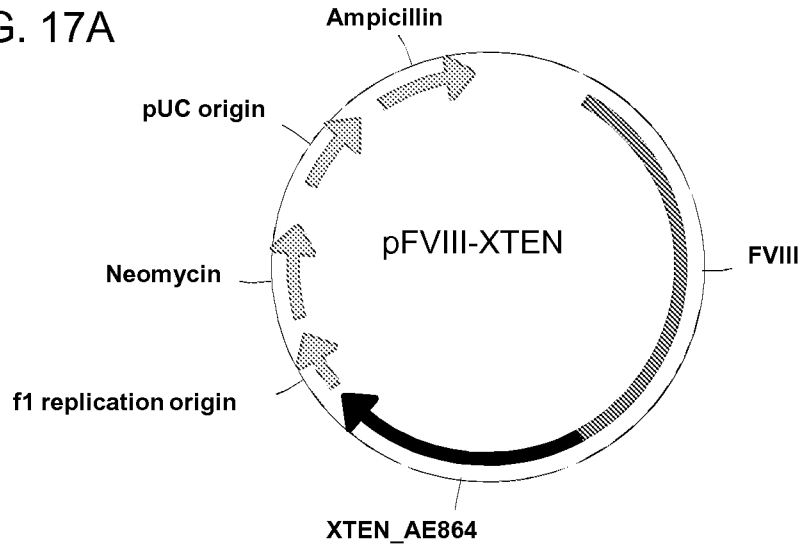


FIG. 17B

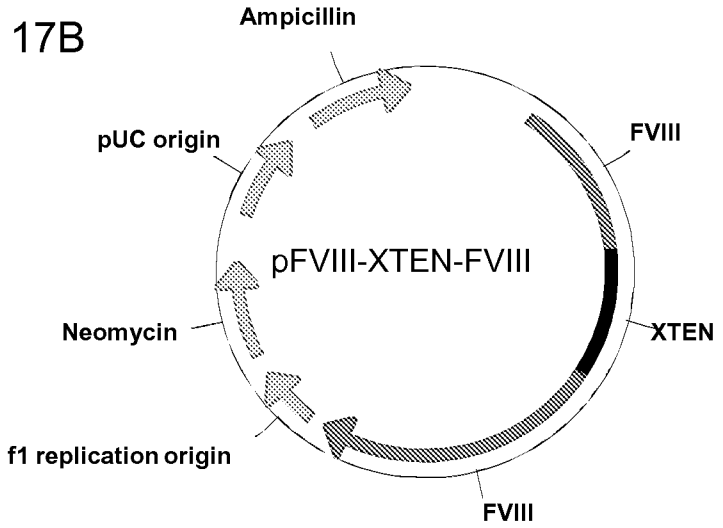


FIG. 17C

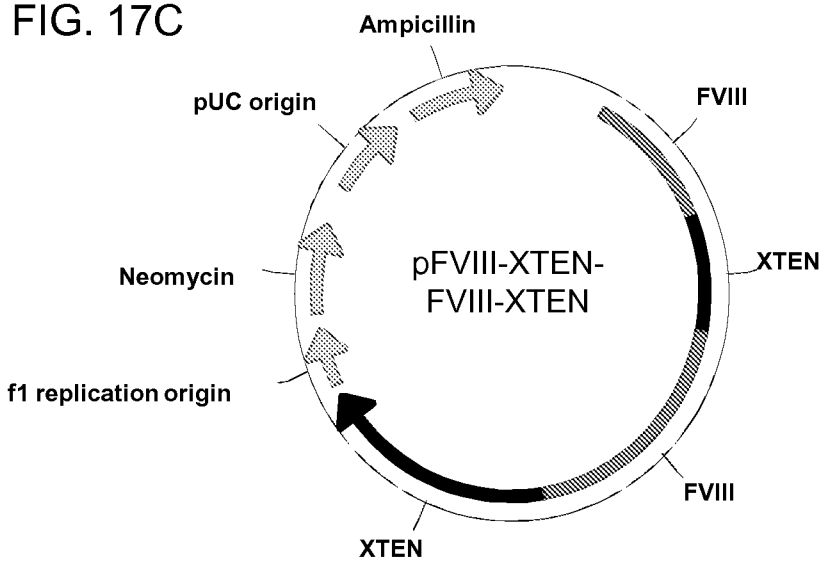


FIG. 17

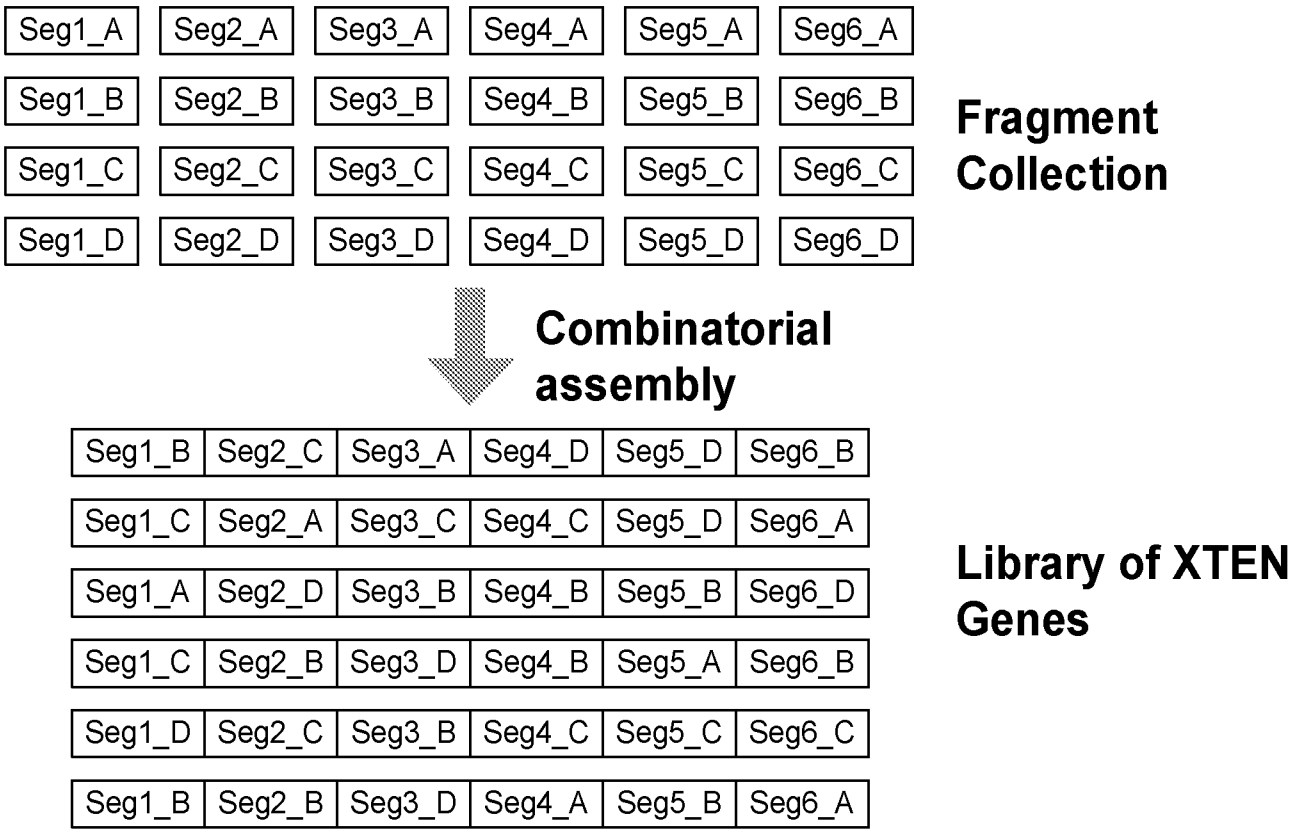

Seg1_A Seg2_A Seg3_A Seg4_A Seg5_A Seg6_A
Seg1_B Seg2_B Seg3_B Seg4_B Seg5_B Seg6_B
Seg1_C Seg2_C Seg3_C Seg4_C Seg5_C Seg6_C
Seg1_D Seg2_D Seg3_D Seg4_D Seg5_D Seg6_D
Fragment Collection
Combinatorial assembly
Seg1_B Seg2_C Seg3_A Seg4_D Seg5_D Seg6_B
Seg1_C Seg2_A Seg3_C Seg4_C Seg5_D Seg6_A
Seg1_A Seg2_D Seg3_B Seg4_B Seg5_B Seg6_D
Seg1_C Seg2_B Seg3_D Seg4_B Seg5_A Seg6_B
Seg1_D Seg2_C Seg3_B Seg4_C Seg5_C Seg6_C
Seg1_B Seg2_B Seg3_D Seg4_A Seg5_B Seg6_A
Library of XTEN Genes

FIG. 18

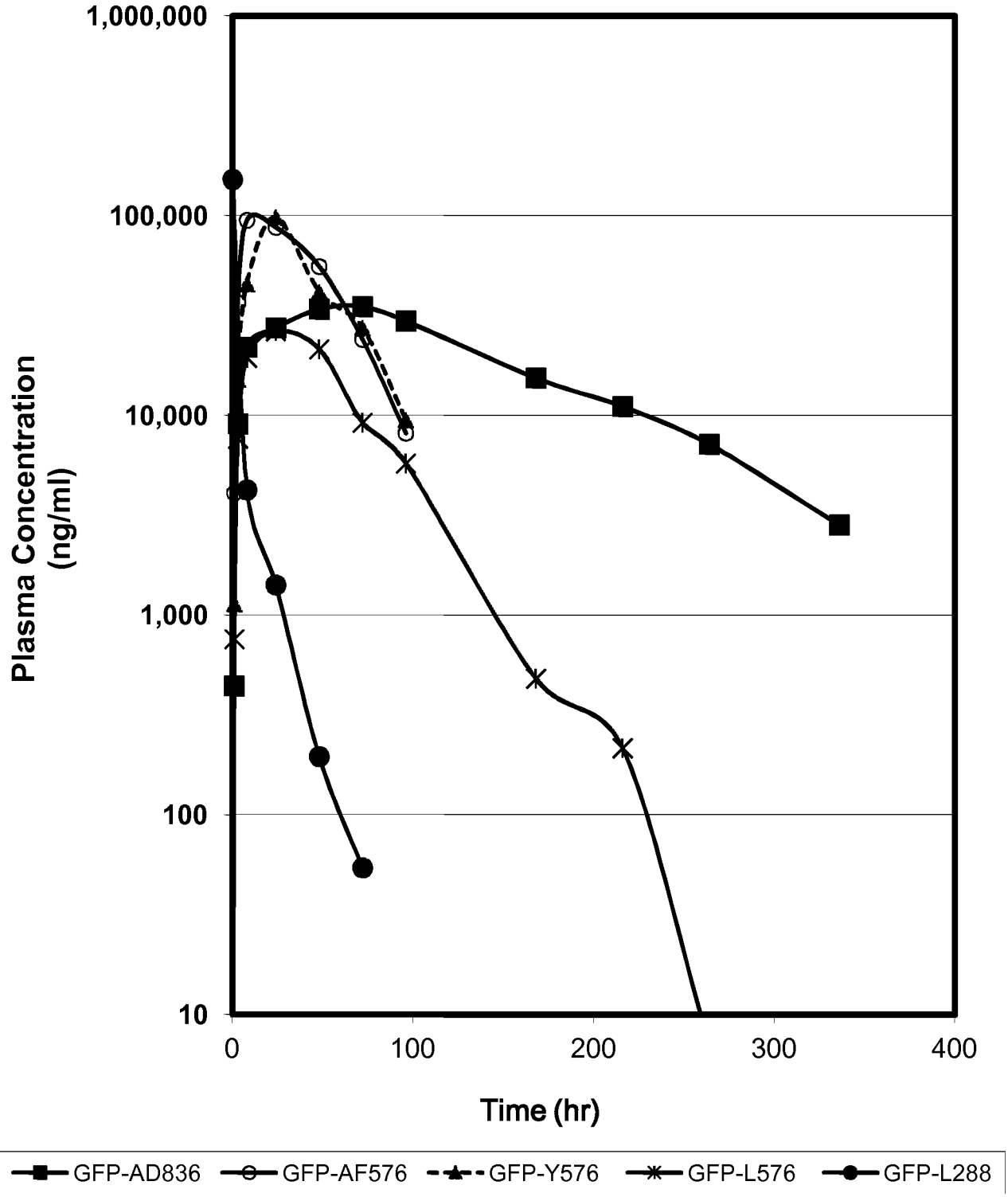

1,000,000
100,000
10,000
1,000
100
10
Plasma Concentration (ng/ml)
0
100
200
300
400
Time (hr)
GFP-AD836
GFP-AF576
GFP-Y576
GFP-L576
GFP-L288

FIG. 19

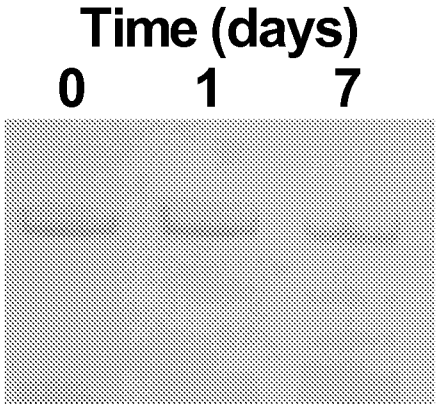


FIG. 18A

Monkey plasma

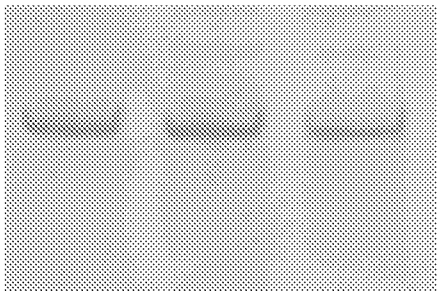


FIG. 18B

Monkey in vivo

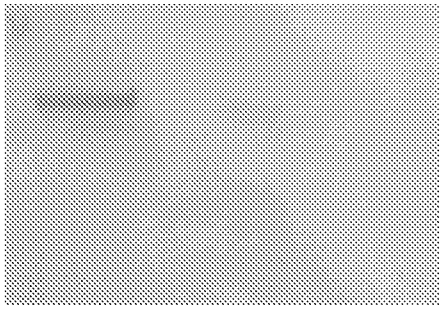
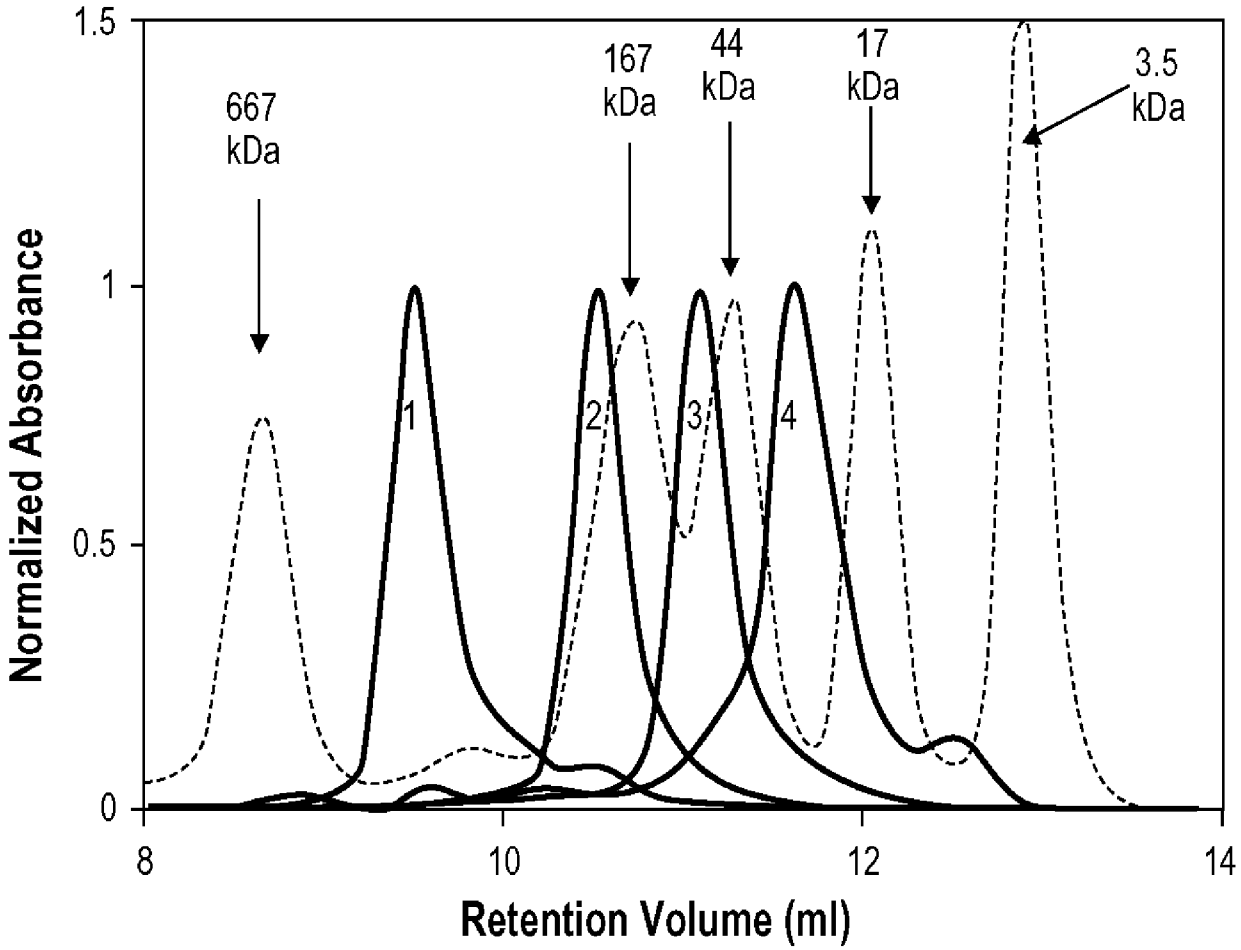


FIG. 18C

Kidney homogenate

FIG. 20



1. Glucagon-Y288
2. Glucagon-Y144
3. Glucagon-Y72
4. Glucagon-Y36

----- = Standards

FIG. 21

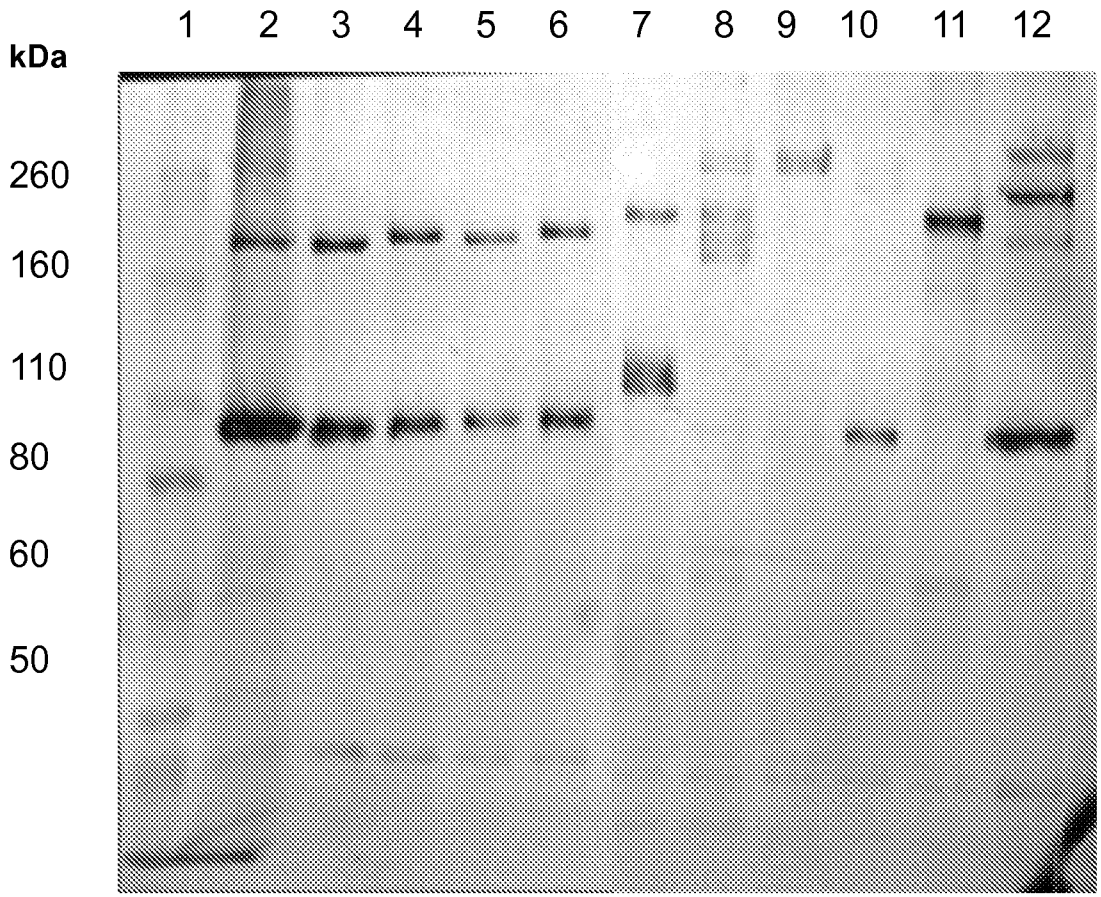


FIG. 22

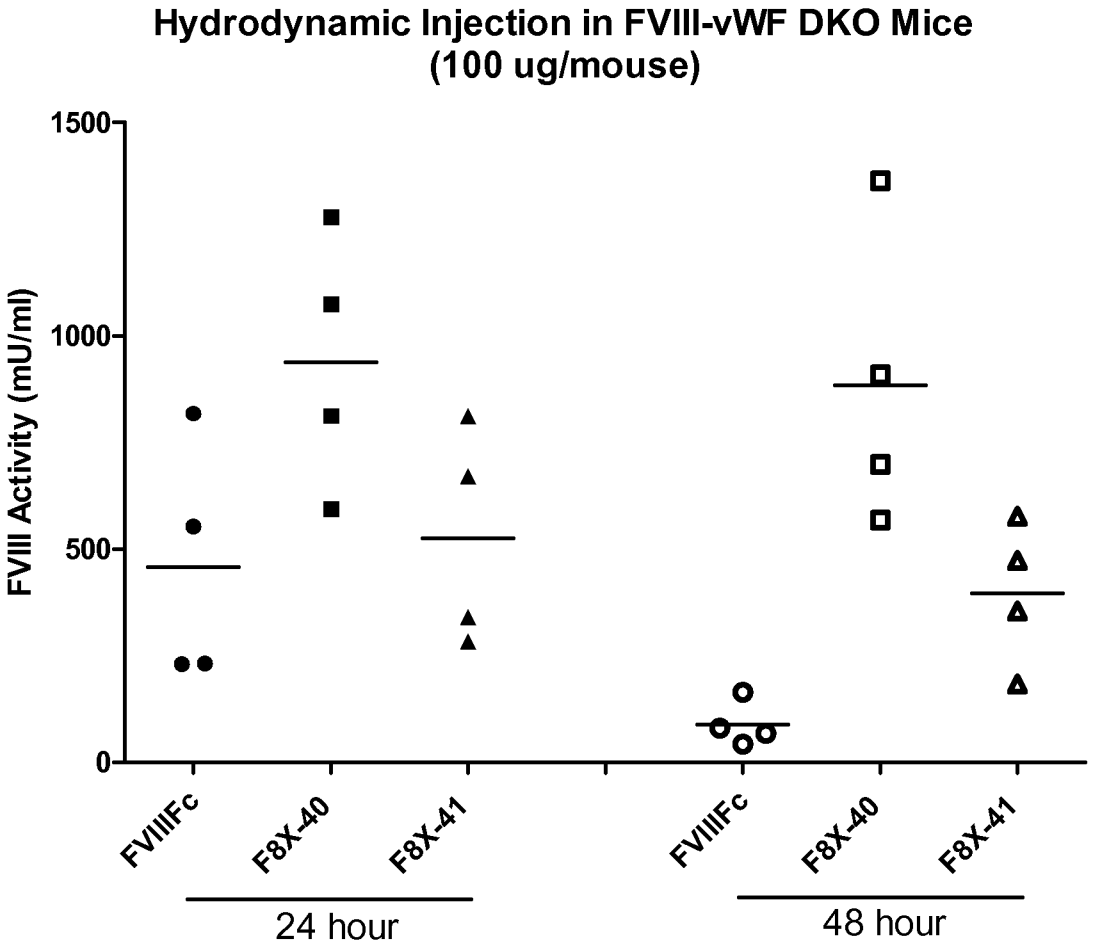

Hydrodynamic Injection in FVIII-vWF DKO Mice
(100 ug/mouse)
FVIII Activity (mU/ml)
1500
1000
500
0
FVIIIFc
F8X-40
F8X-41
FVIIIFc
F8X-40
F8X-41
24 hour
48 hour

FIG. 23

Fig. 24A

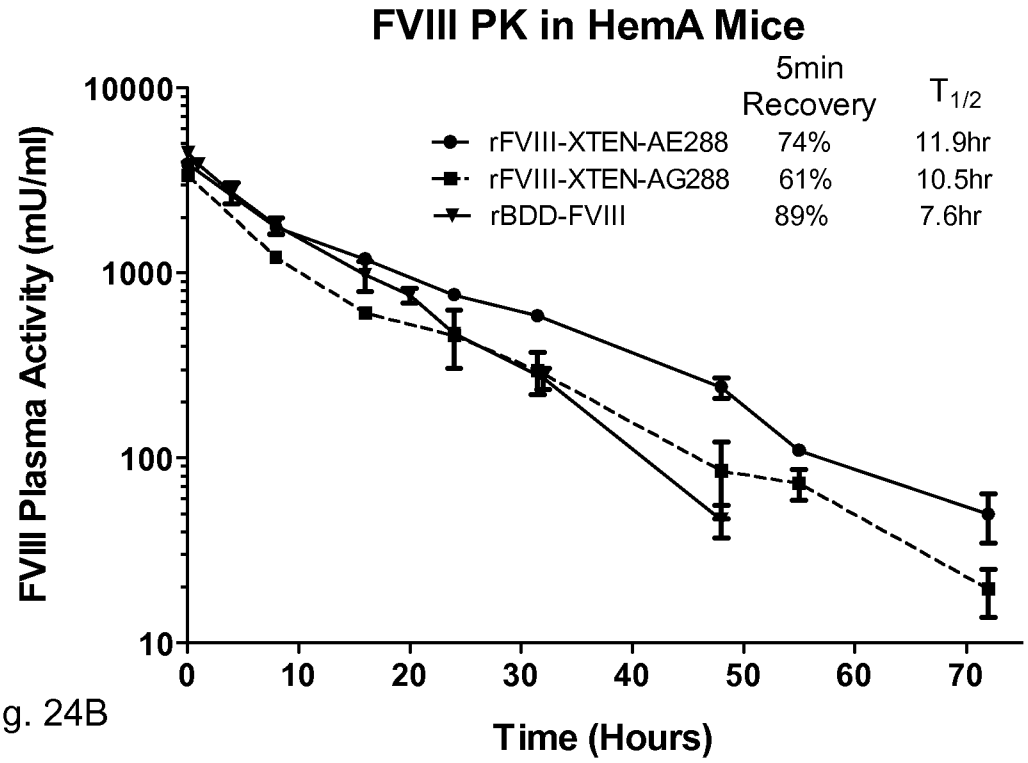


Fig. 24B

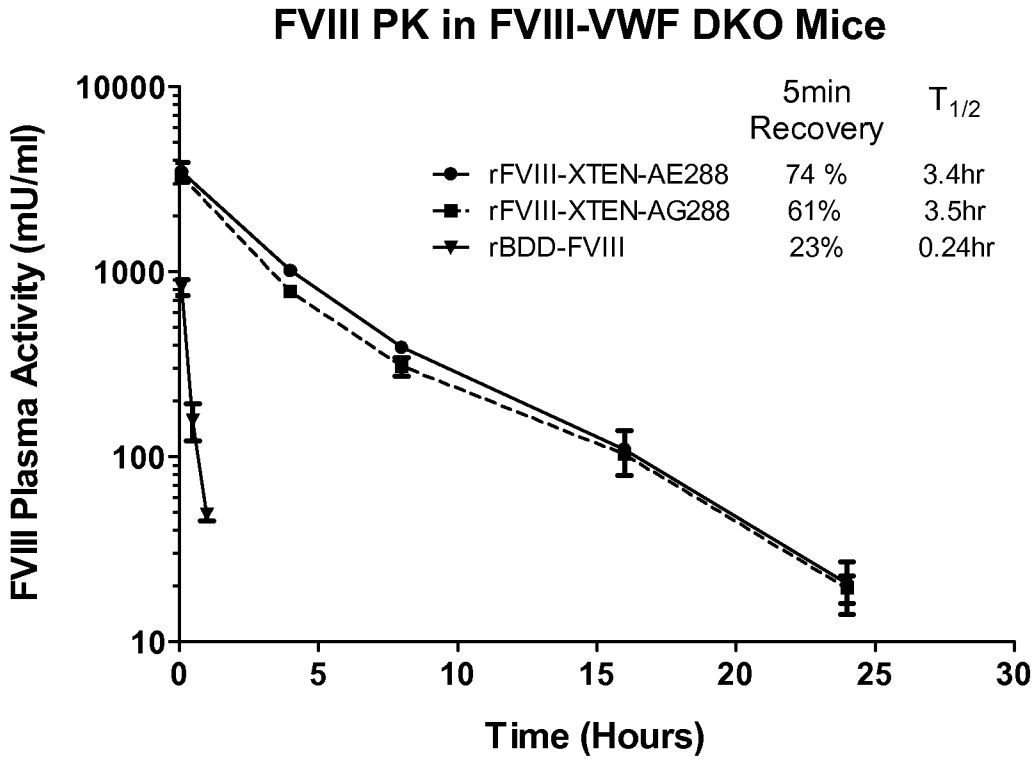


FIG. 24

Fig. 25A

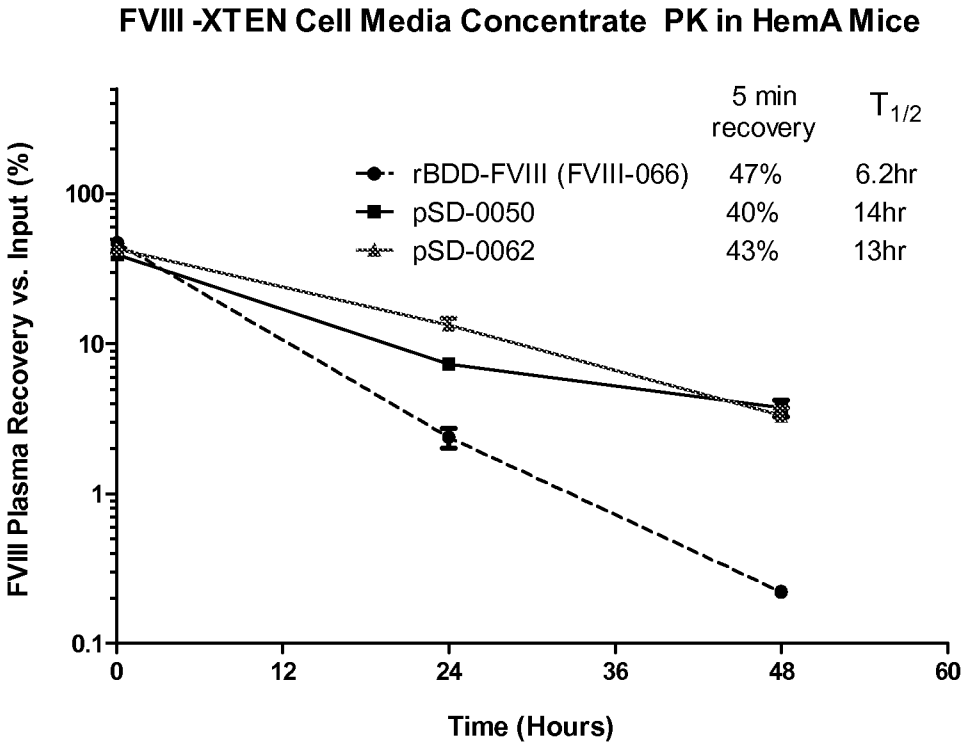


Fig. 25B

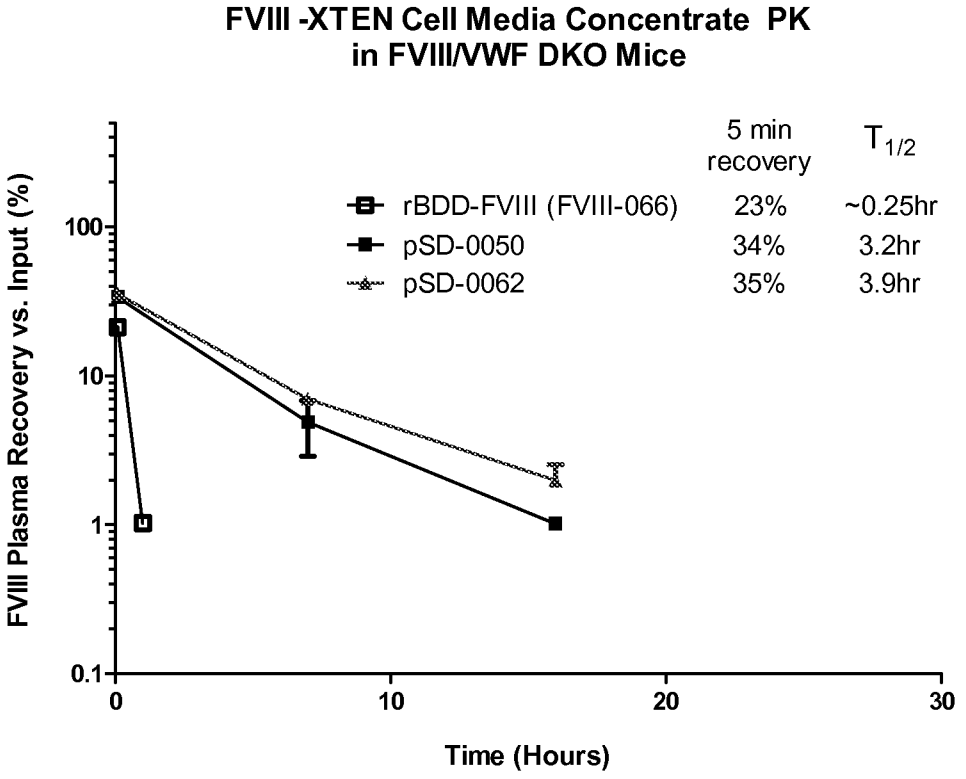


FIG. 25

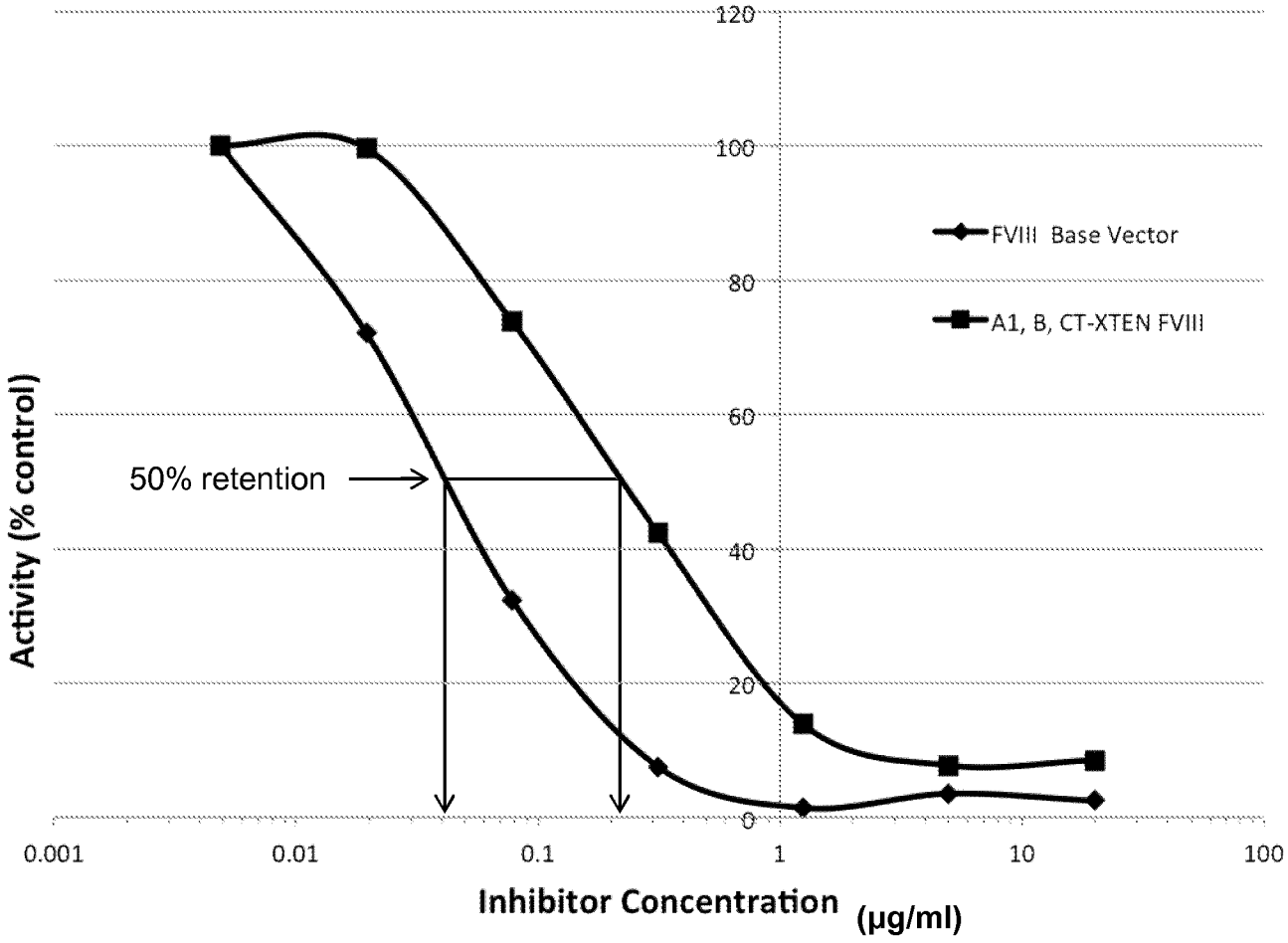

120
100
80
60
40
20
0
Activity (% control)
50% retention
FVIII Base Vector
A1, B, CT-XTEN FVIII
0.001
0.01
0.1
1
10
100
Inhibitor Concentration (µg/ml)

FIG. 26

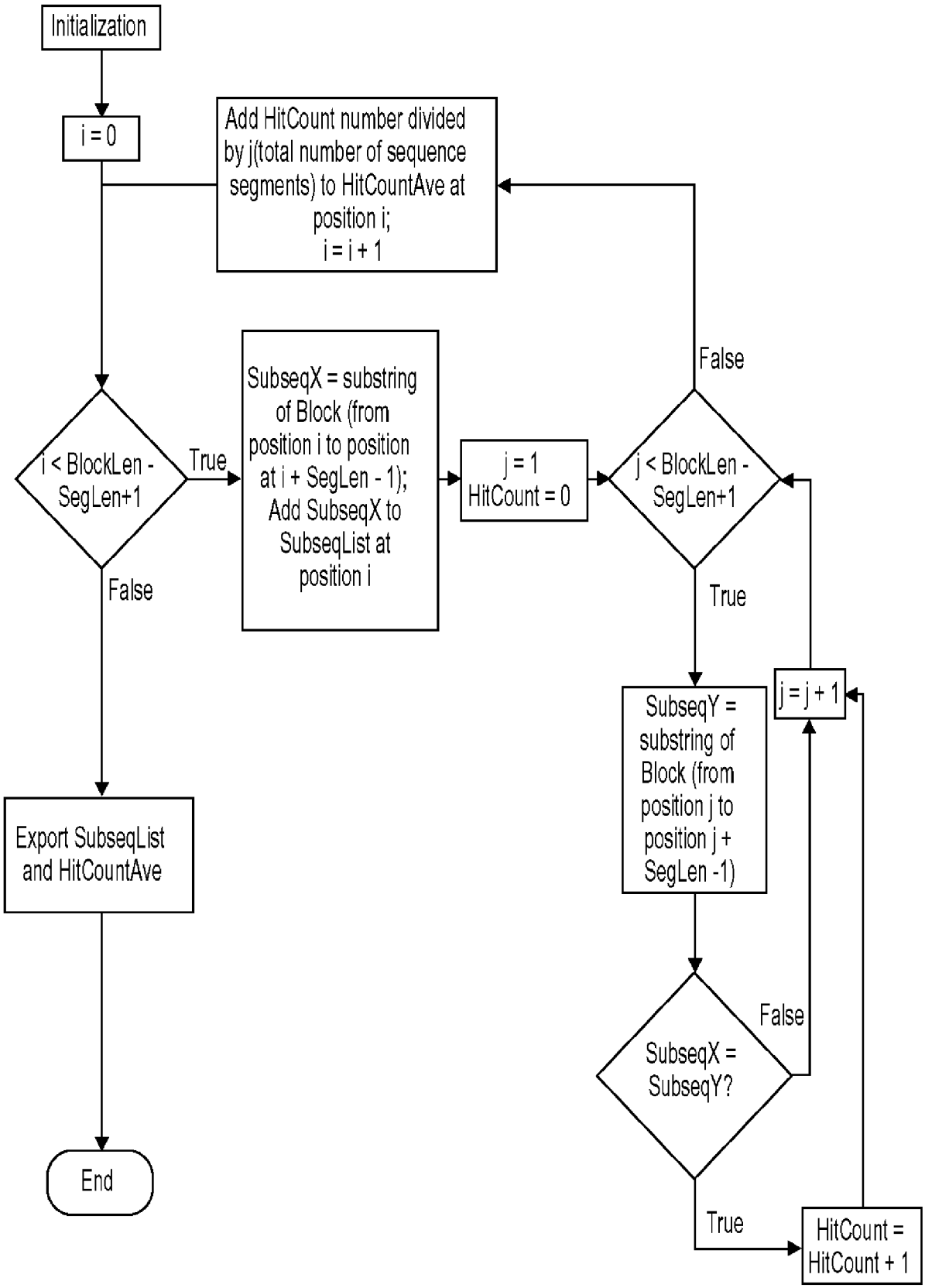

Initialization
i = 0
Add HitCount number divided by j(total number of sequence segments) to HitCountAve at position i;
i = i + 1
i < BlockLen - SegLen+1
True
False
SubseqX = substring of Block (from position i to position at i + SegLen - 1); Add SubseqX to SubseqList at position i
j = 1
HitCount = 0
j < BlockLen - SegLen+1
False
True
SubseqY = substring of Block (from position j to position j + SegLen -1)
j = j + 1
SubseqX = SubseqY?
False
True
HitCount = HitCount + 1
Export SubseqList and HitCountAve
End

FIG. 27

XTEN Sequence	1	2	3	4	5	6	7	8	9	10	11
	A	S	T	G	E	T	G	E	T	G	E

XTEN Length: N=11

Subsequence length: S=3

		Seg 1	Seg 2	Seg 3	Seg 4	Seg 5	Seg 6	Seg 7	Seg 8	Seg 9
		AST	STG	TGE	GET	ETG	TGE	GET	ETG	TGE
Seg 1	AST	x								
Seg 2	STG		x							
Seg 3	TGE			x			x			
Seg 4	GET				x			x		
Seg 5	ETG					x			x	
Seg 6	TGE			x			x			x
Seg 7	GET				x			x		
Seg 8	ETG					x			x	
Seg 9	TGE						x			x
		1	1	2	2	2	3	2	2	2

Subsequence Score = 1.89

FIG. 28

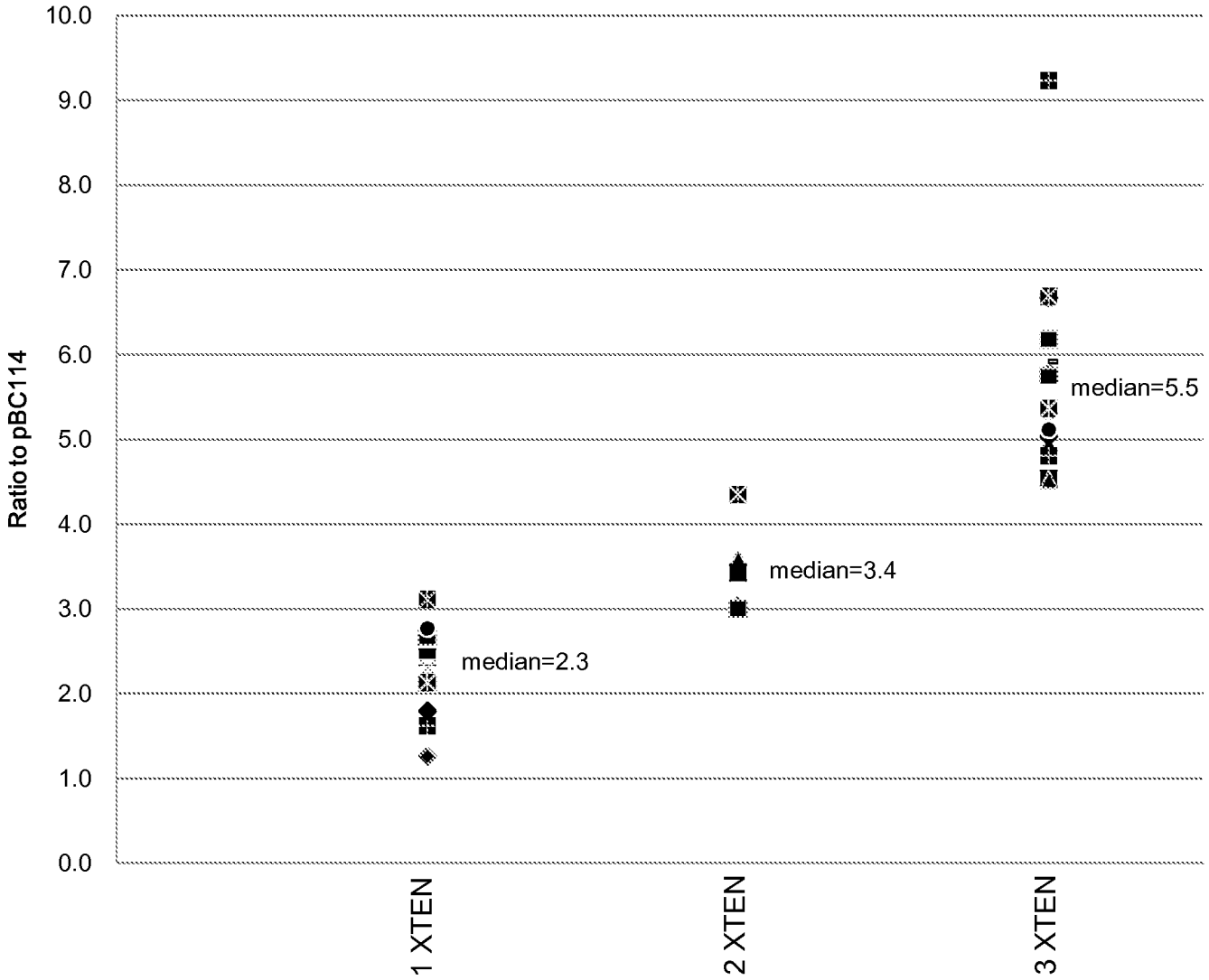

10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
Ratio to pBC114
median=2.3
median=3.4
median=5.5
1 XTEN
2 XTEN
3 XTEN

FIG. 29

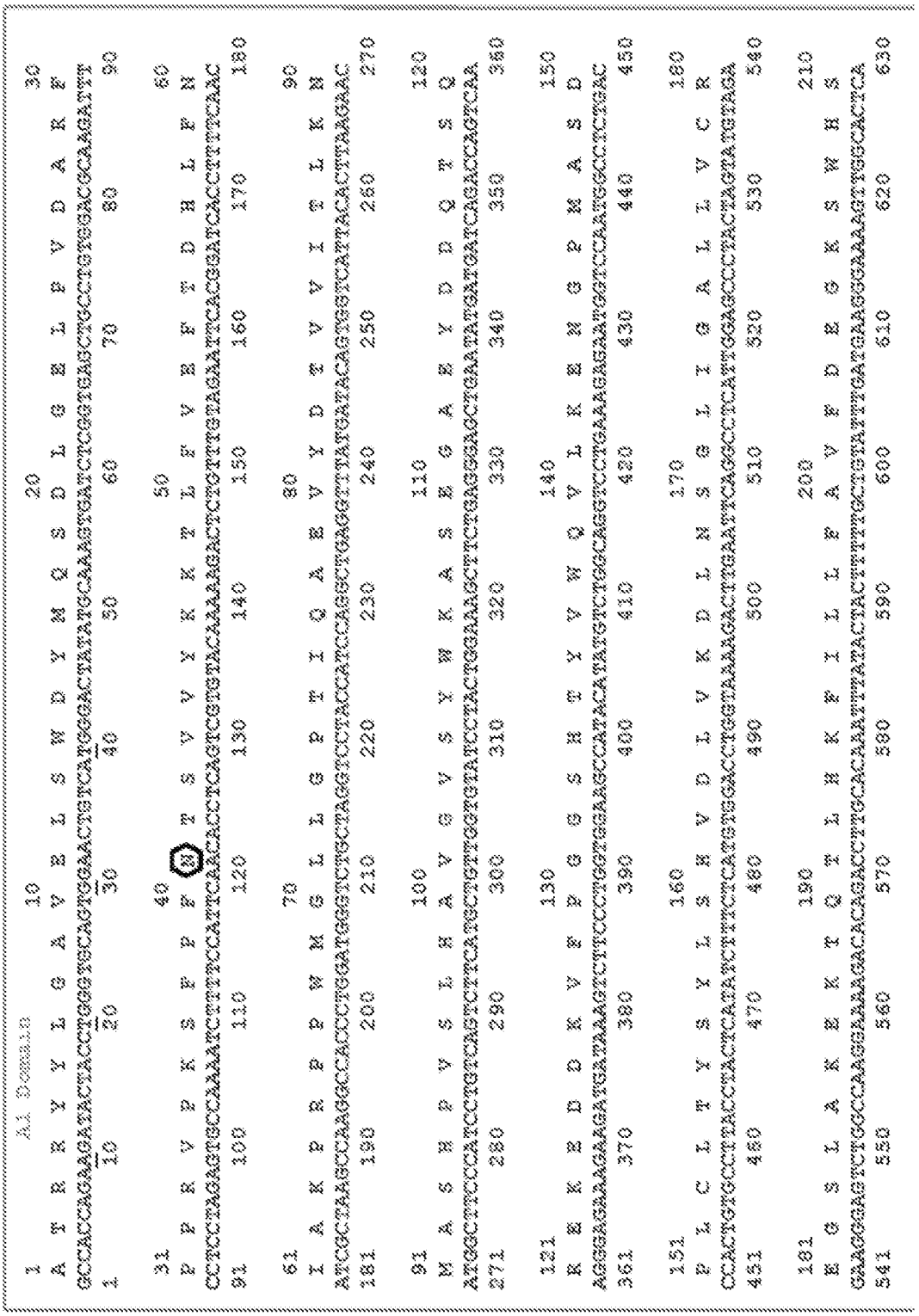

A1 Domain
1 10 20 30
A T E R Y Y L G A Y E L S W D Y M Q S D L G E L P V D A R F
GCCACCAGAAGATACTACCTGGGTGCAGTGGAACTGTCATGGACTATATGCAAAGTGATCTCGGTGAGCTGCCTGTGGACGCAAGATTT
1 10 20 30 40 50 60 70 80 90
31 40 50
P P R V P K S F P F N T S V V Y K K T L F V E F T D H L F N
CCTCCTAGAGTGCCAAAATCTTTCCATTCAACACCTCAGTCGTGTACAAAAAGACTCTGTTTGTAGAATTCACGGATCACCTTTTCAAC
91 100 110 120 130 140 150 160 170 180
61 70 80 90
I A K P R P P W M G L L G P T I Q A E V Y D T V V I T L K N
ATCGCTAAGCCAAGGCCACCCTGGATGGTCTGCTAGGTCCTACCATCCAGGCTGAGGTTTATGATACAGTGGTCATTACACTTAAGAAC
181 190 200 210 220 230 240 250 260 270
91 100 110
M A S H P V S L H A V G V S Y W K A S E G A E Y D D Q T S Q
ATGGCTTCCCATCCTGTCAGTCTTCATGCTGTTGGTGTATCCTACTGGAAAGCTTCTGAGGGAGCTGAATATGATCAGACCAGTCAA
271 280 290 300 310 320 330 340 350 360
121 130 140 150
R E K D D K V F P G G S H T Y V W Q V L K E N G P M A S D
AGGGAGAAAGAAGATGATAAAGTCTTCCCTGGTGGAAGCCATACATATGTCTGGCAGGTCCTGAAAGAGAATGGTCCAATGGCCTCTGAC
361 370 380 390 400 410 420 430 440 450
151 160 170 180
P L C L T Y S Y L S K V D L V K D L N S G L I G A L L V C R
CCACTGTGCCTTACCTACTCATACTTTCTCATGTGGACCTGGTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGTACTGAGA
451 460 470 480 490 500 510 520 530 540
181 190 200 210
E G S L A K K E K T Q T L H K F I L L F A V F D E G K S W H S
GAAGGGAGTCTTGGCCAAGGAAAAGACACAGACCTTGCACAAATTTATACTACTTTTTGCTGTATTTGATGAAGGGAAAAGTTGGCACTCA
541 550 560 570 580 590 600 610 620 630

FIG. 30A

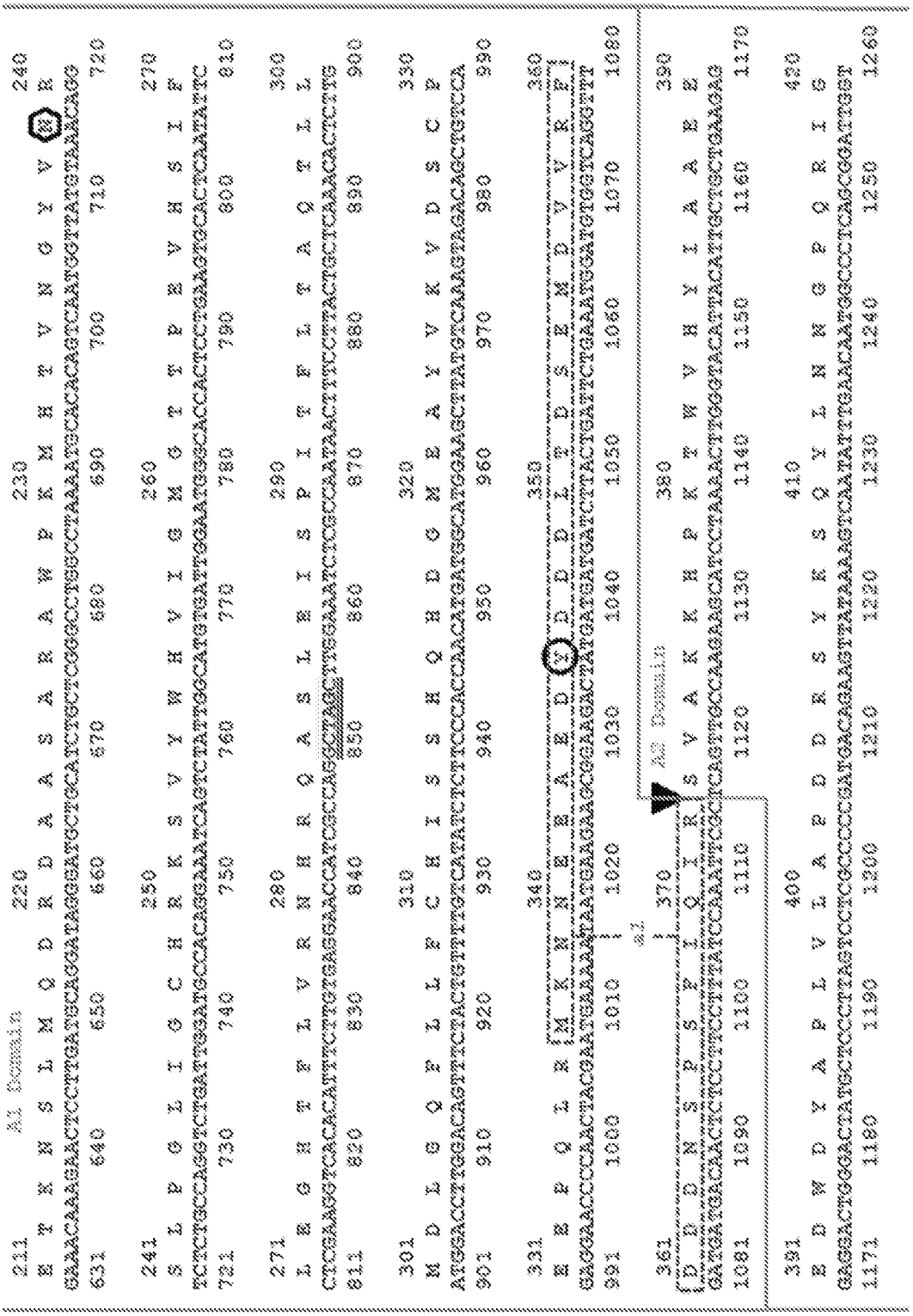


FIG. 30B

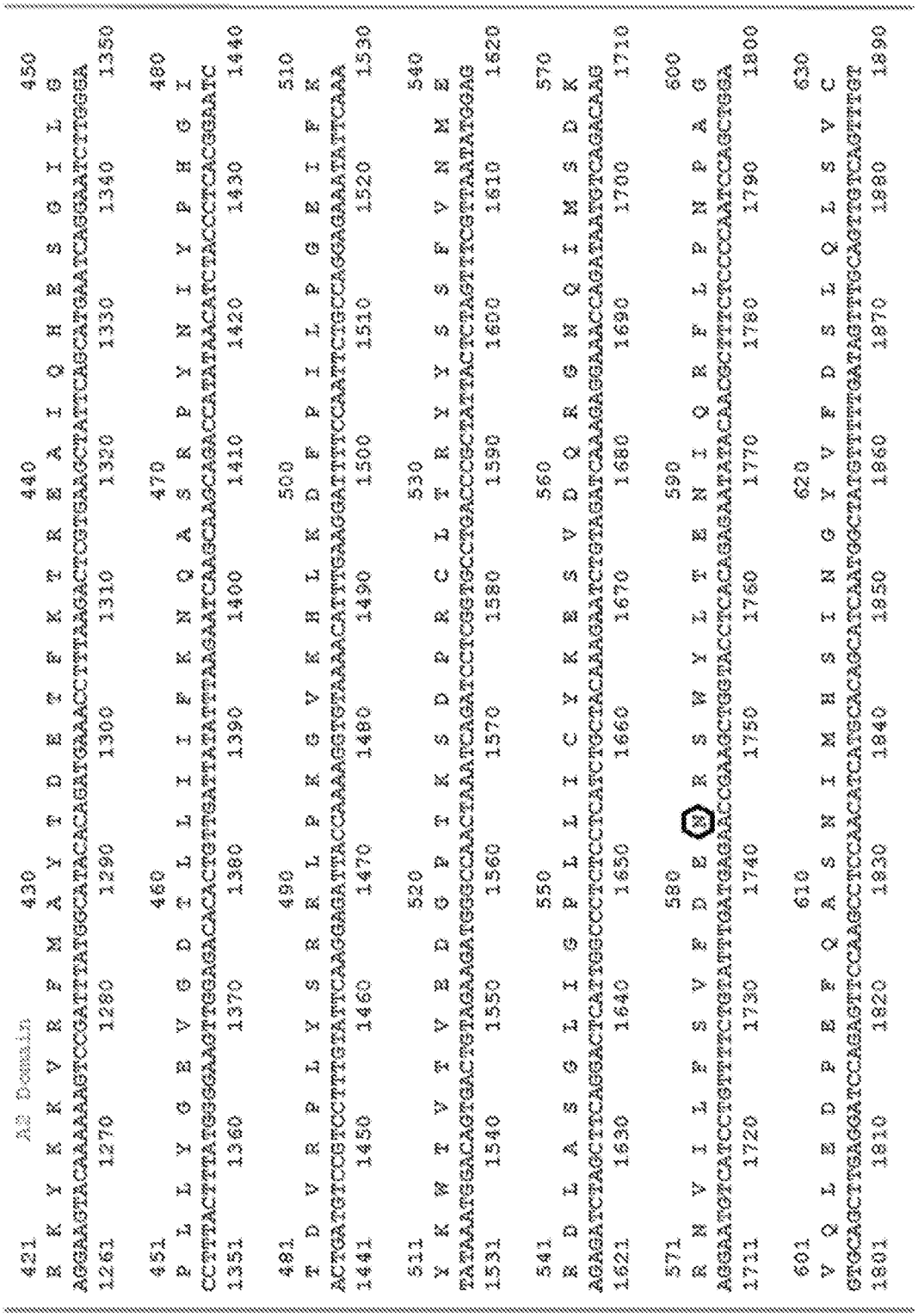


FIG. 30C

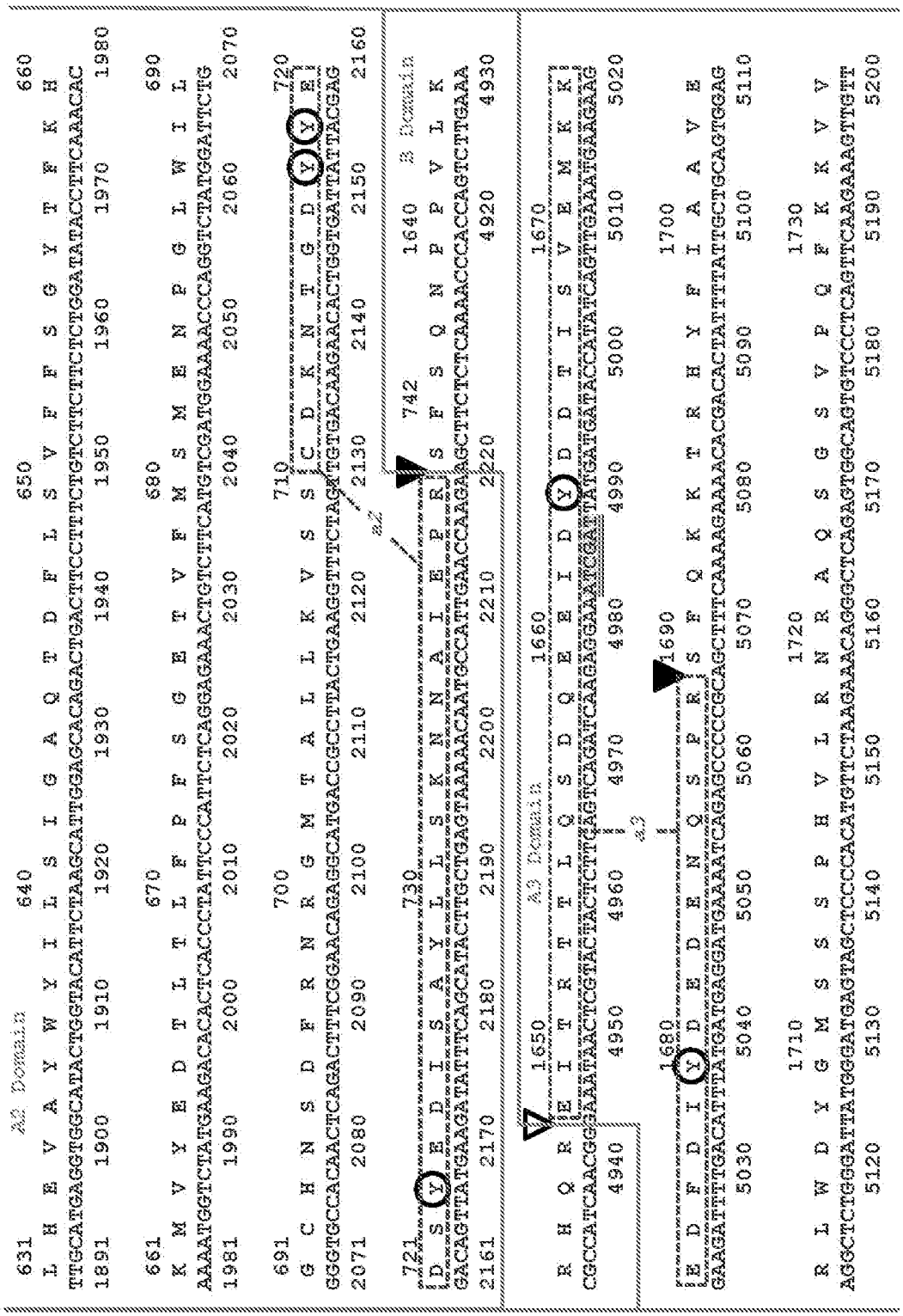


FIG. 30D

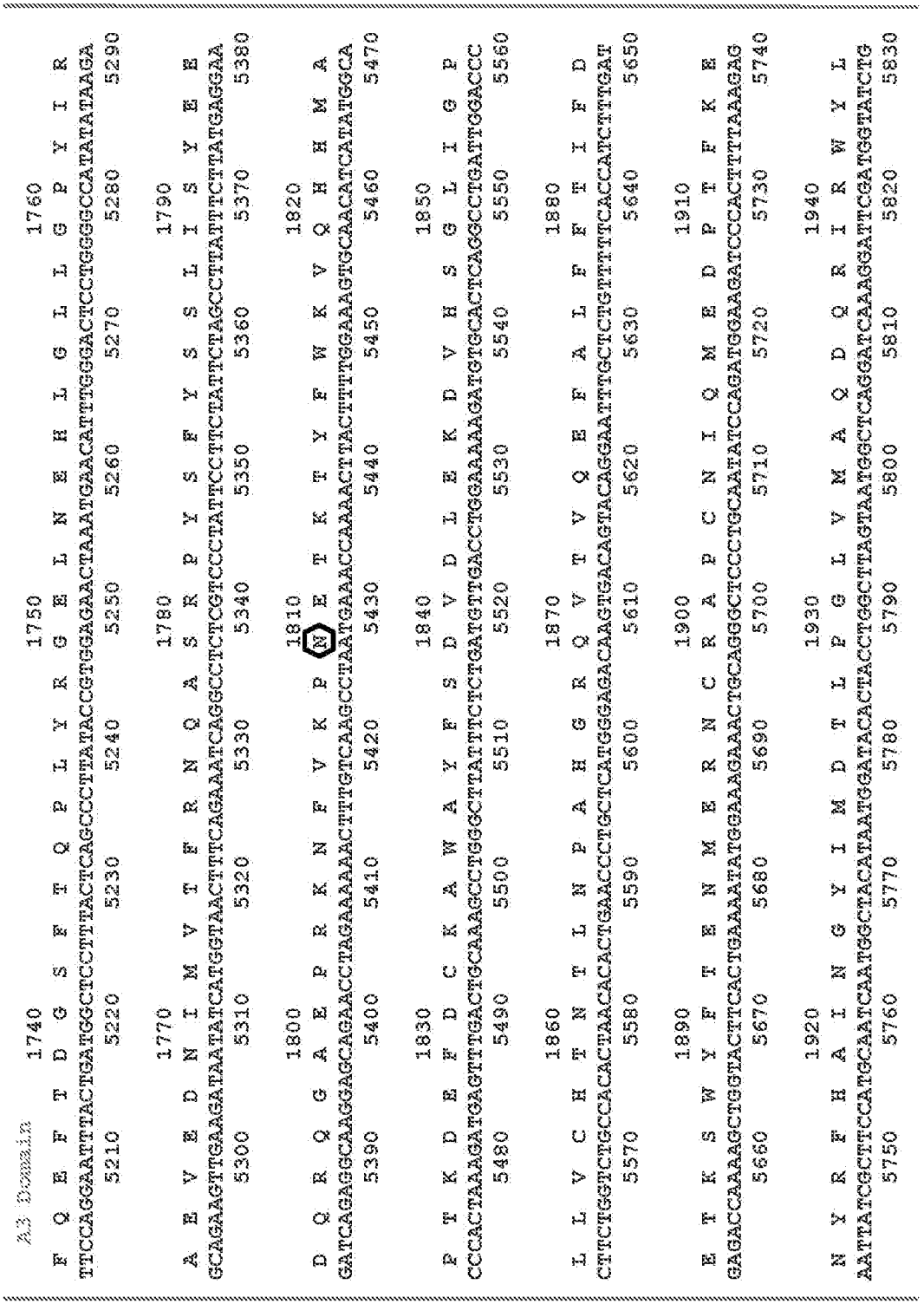


FIG. 30E

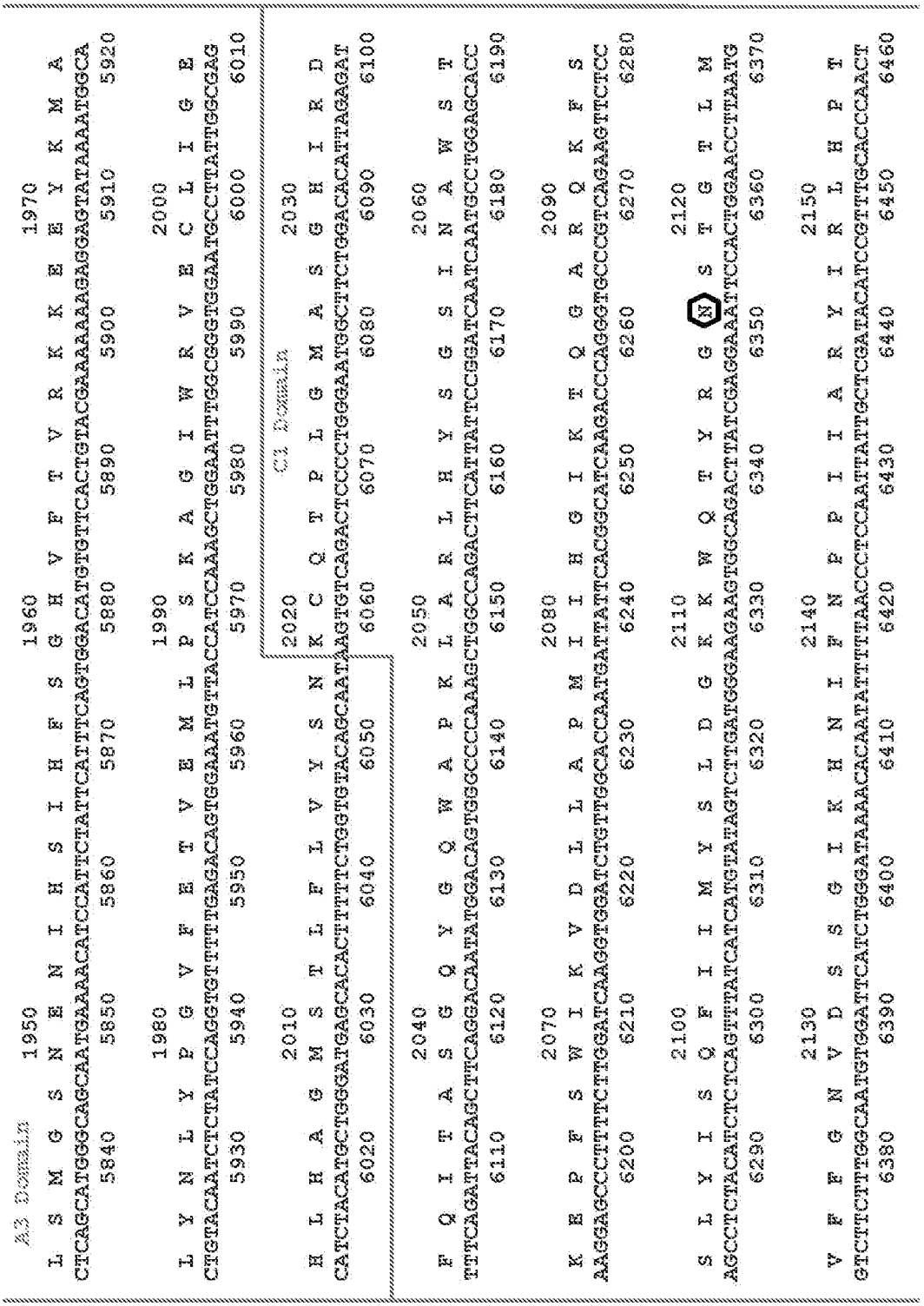


FIG. 30F

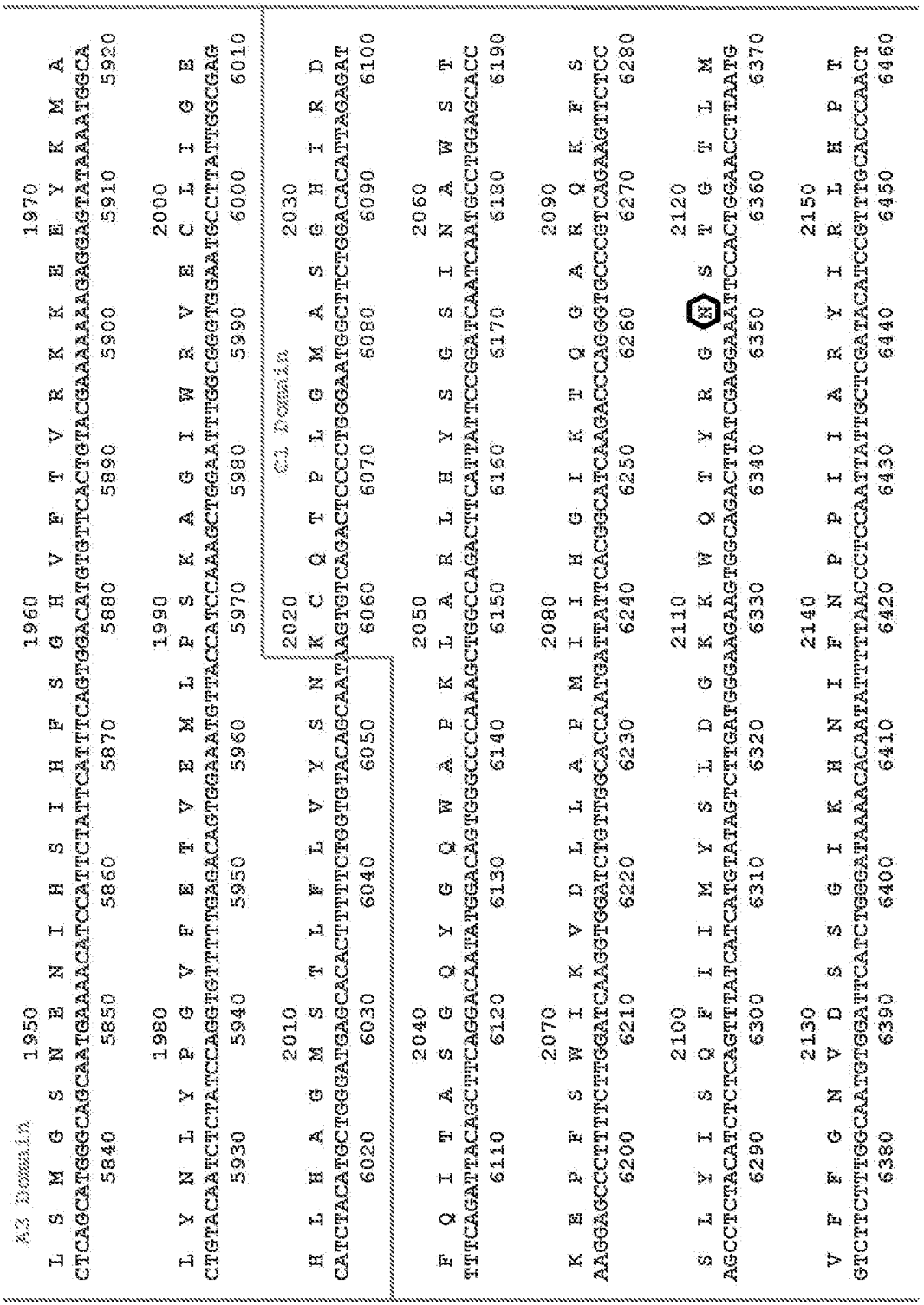


FIG. 30G

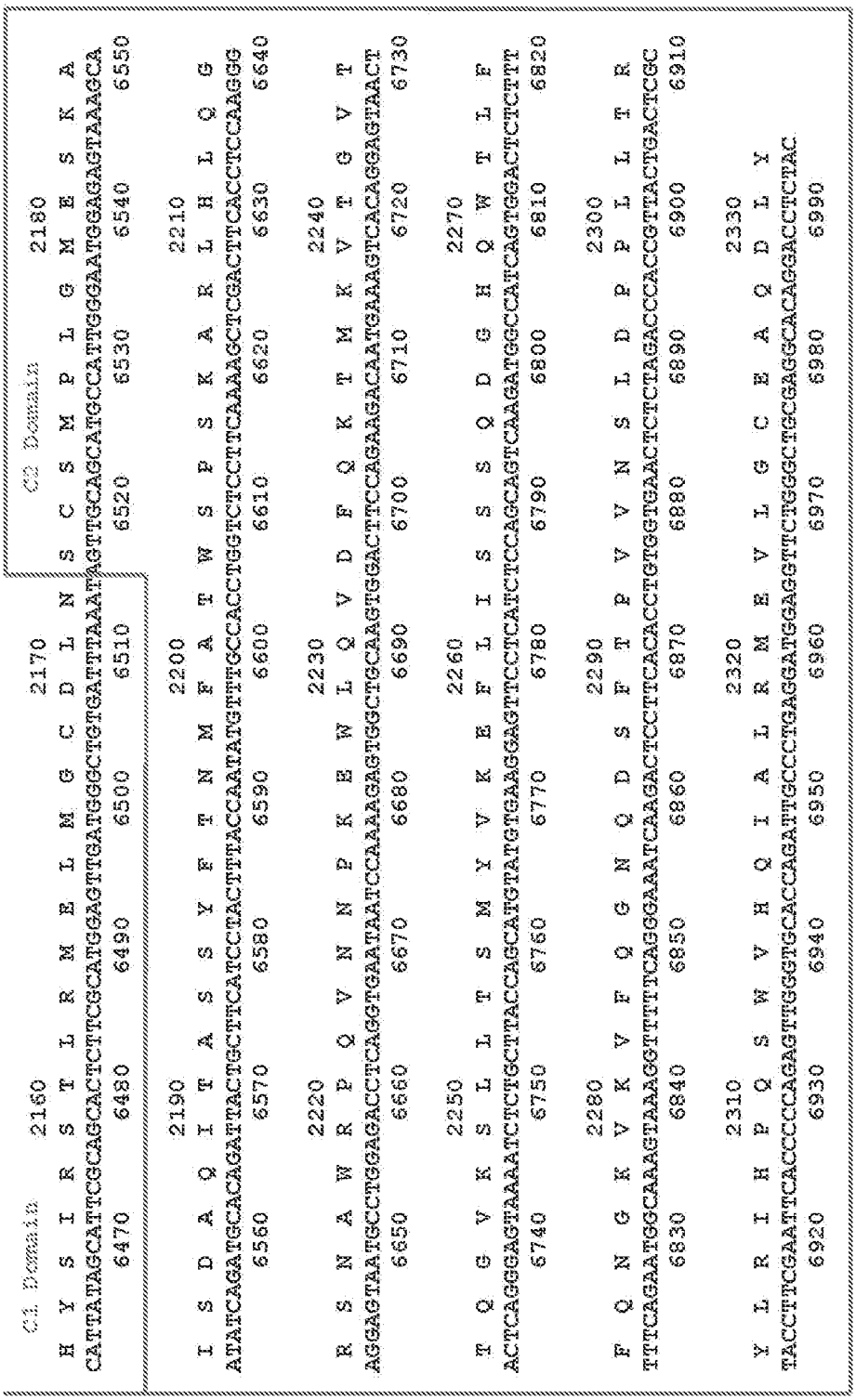

C1 Domain
C2 Domain
2160
2170
2180
2190
2200
2210
2220
2230
2240
2250
2260
2270
2280
2290
2300
2310
2320
2330

FIG. 30G

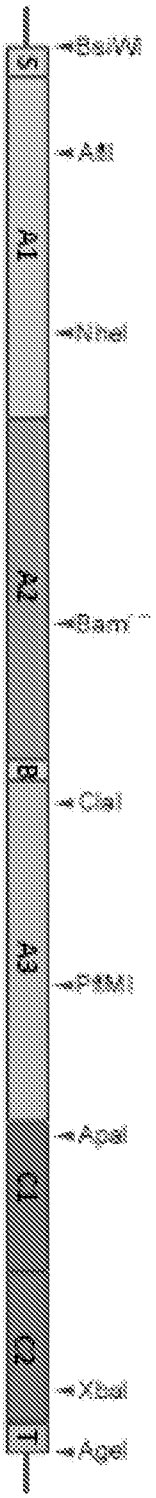

S
A1
A2
B
A3
T3
T2
T
BsaWI
AflII
NheI
BamHI
ClaI
PflMI
ApaI
XbaI
AgeI

FIG. 31

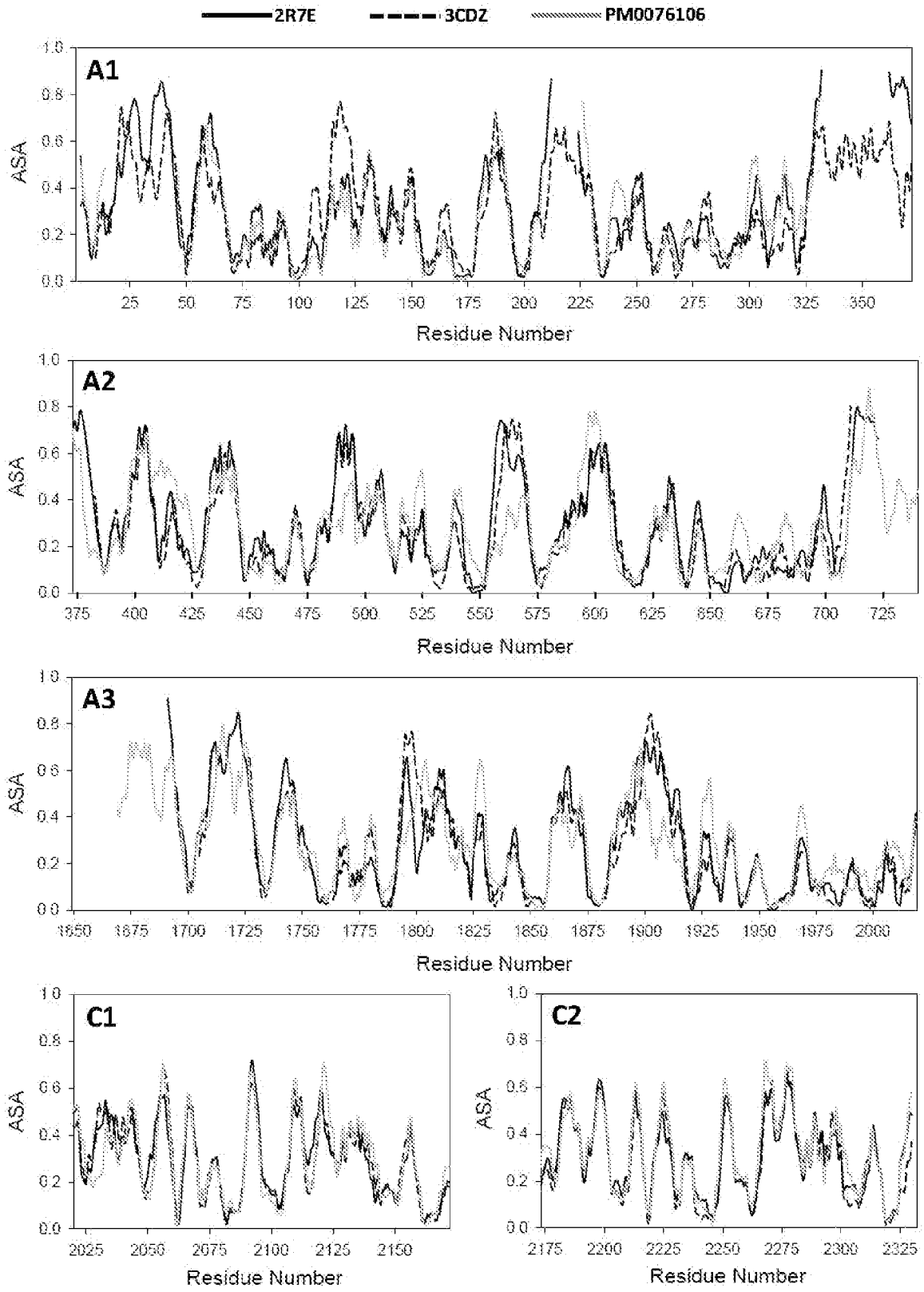


FIG. 32

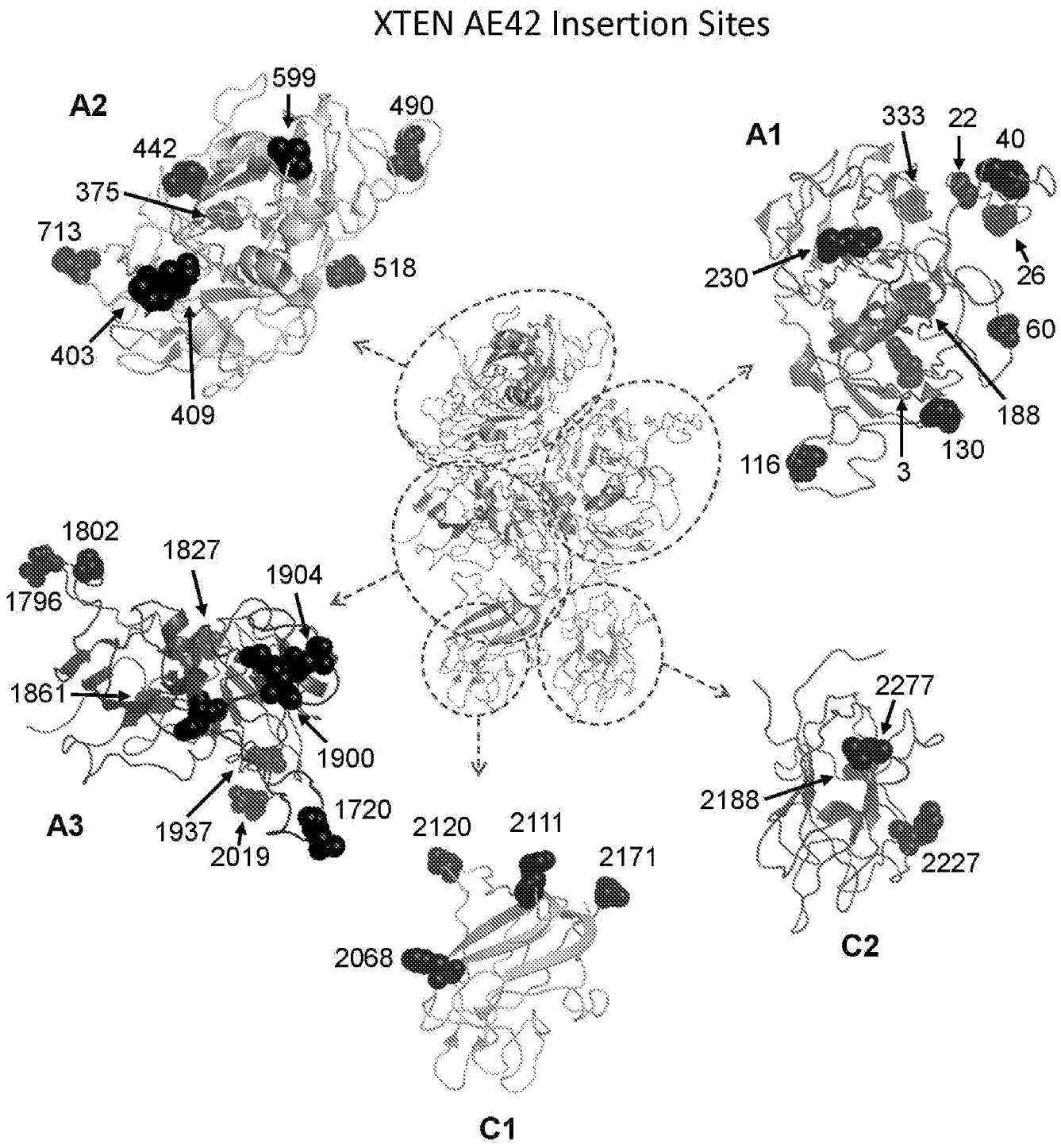

XTEN AE42 Insertion Sites
A2
599
490
442
375
713
518
403
409
A1
333
22
40
230
26
60
188
116
3
130
A3
1802
1827
1904
1796
1861
1900
1720
1937
2019
C1
2120
2111
2171
2068
C2
2277
2188
2227

FIG. 33

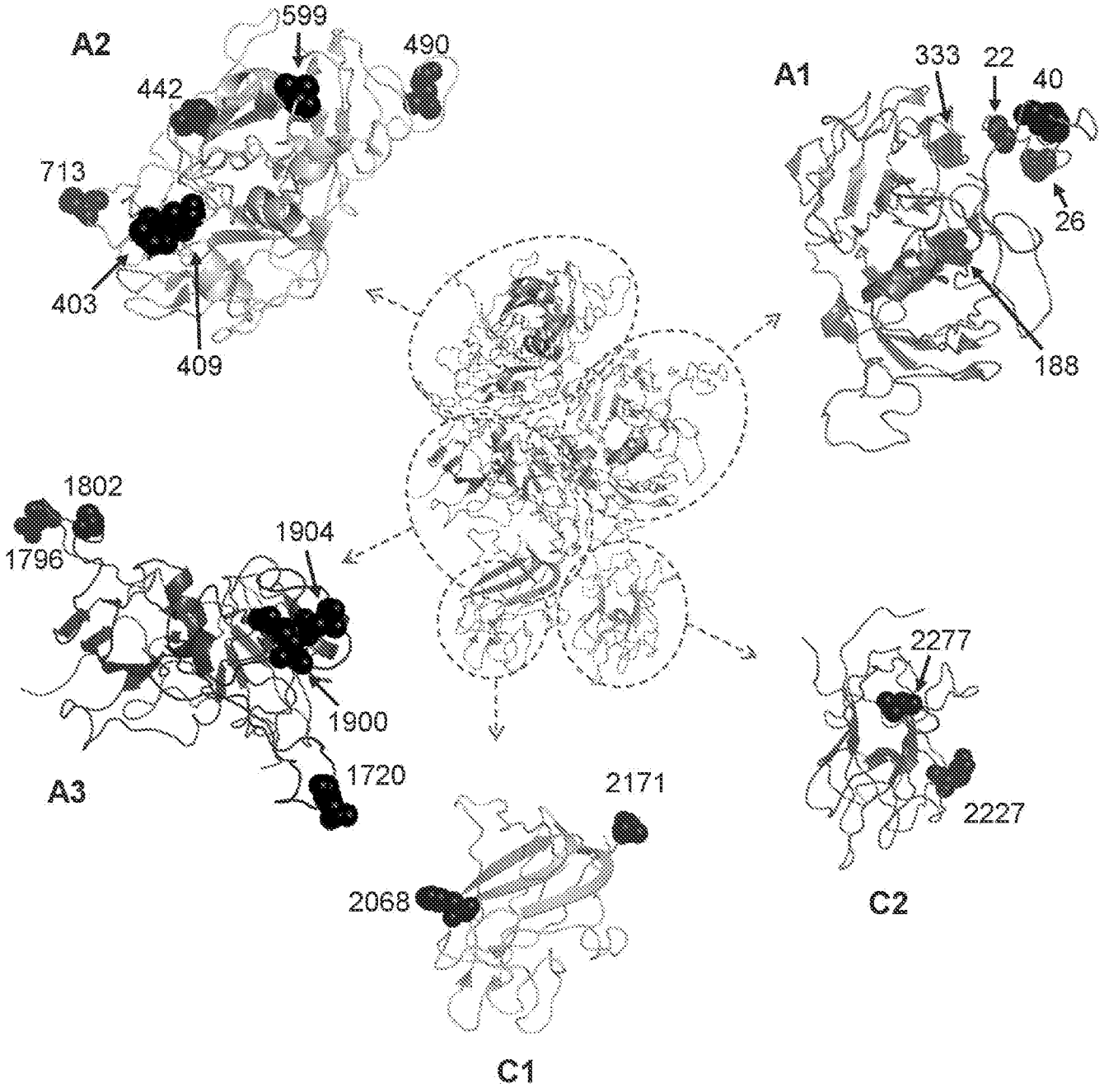

XTEN AE42 Insertion Sites with Activity
A2
599
490
442
713
403
409
A1
333
22
40
26
188
1802
1796
1904
1900
1720
A3
2277
2227
C2
2171
2068
C1

FIG. 34

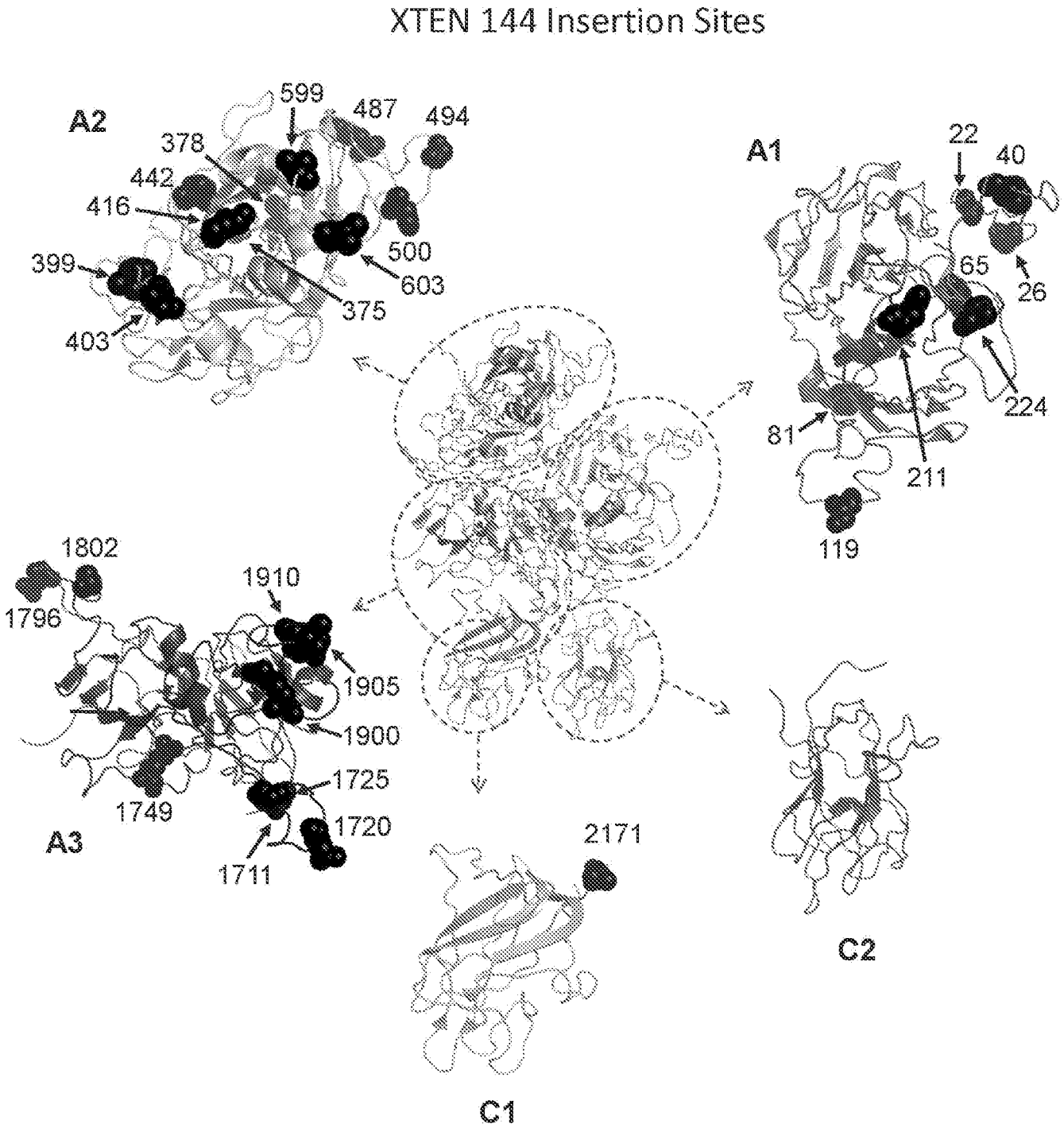

XTEN 144 Insertion Sites
A2
599
487
494
378
442
416
399
500
603
375
403
A1
22
40
65
26
224
81
211
119
1802
1796
1910
1905
1900
1725
1720
1749
1711
A3
2171
C1
C2

FIG. 35

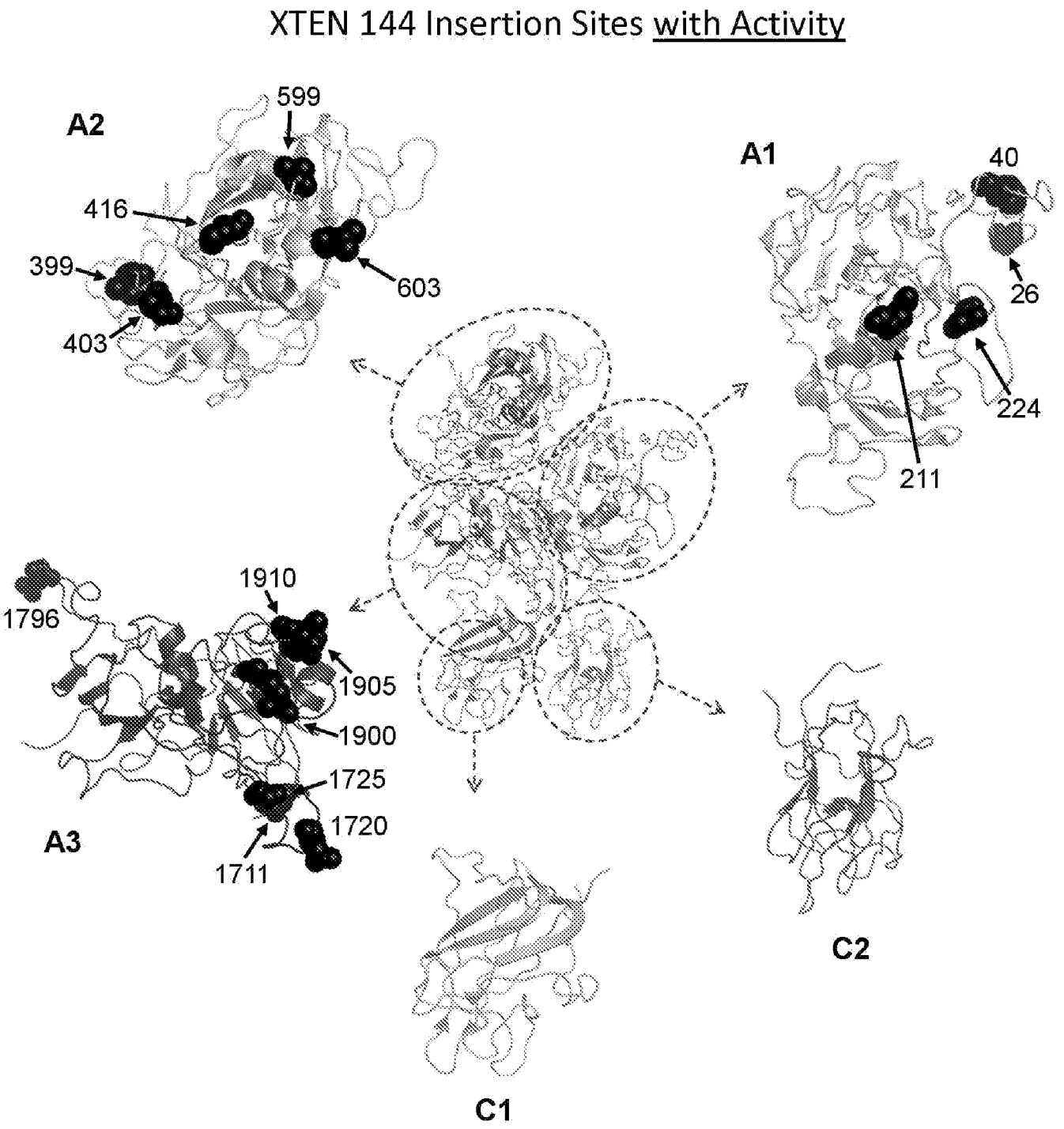

XTEN 144 Insertion Sites with Activity
A2
599
416
399
603
403
A1
40
26
224
211
1910
1796
1905
1900
1725
1720
A3
1711
C2
C1

FIG. 36

```
A1 (1)                                                          ATRRYYLGAVELSWD
A2 (373                                                       SVAKKHPKTWVHYIAAEEEDWD
A3 (1649)   EITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWD
                                                                *  :::.* *  **

A1 (16)     YMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFTDHLFNIAKPR---PPWMGLL
A2 (395)    YAPLVL--APDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAI---QHESGIL
A3 (1709)   YG---------MSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLL
            *                    :         :**. *  :**  *.             *:*

A1 (73)     GPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGG
A2 (450)    GPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLY--SRRLPKGVKHLKDFPILPGE
A3 (1760)   GPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQR--QGAEPR------KNFVKPNE
            ** : .** *.::: ::* **:* .::.  ::   .     .           .  : *.

A1 (133)    SHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAK--EKTQ
A2 (508)    IFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIM
A3 (1812)   TKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQV
              .* * *  . .*  .:  * :  * * *:: :*: *****.**:*   ::     :

A1 (191)    TLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWP------KMHTVNGYVNRSLPG
A2 (568)    SDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQ-
A3 (1872)   TVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPG
            : :.  ::*::*** :**:   : .     .     *       :*::***:  :*

A1 (245)    LIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRN---HRQASLEISPITFLTAQTLLM
A2 (627)    LSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSG---ETVFMSME
A3 (1932)   LVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPS
            *  .  :   *:::.:*:  :. *:.:.*:.*  :       :   :      *.

A1 (302)    DLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFD
A2 (684)    NPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPR
A3 (1992)   KAGIWRVECLIGEHLHAGMSTLFLVYSN
            . * : : *  ... : ** : . * *

A1 (362)    DDNSPSFIQIR


C1 (2020)   KCQTPLGMASGHIRDFQITASGQYG----QWAPKLARLHYSGSINAWS--TKEPFSWIKV
C2 (2173)   SCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQV
            .*. **** *  * * *****. :      *:*. **** .*  ***   .::* .*::*

C1 (2074)   DLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSS
C2 (2233)   DFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQN--GKVKVFQGNQDSF
            *:  .* : *:.***.:. ::*:*:.:*:*  * **::*  :  *  *.: ** ** **

C1 (2134)   GIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLN---
C2 (2291)   TPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY
                * ::**:::**:*:**  :  : :*****::**: :
```

FIG. 37

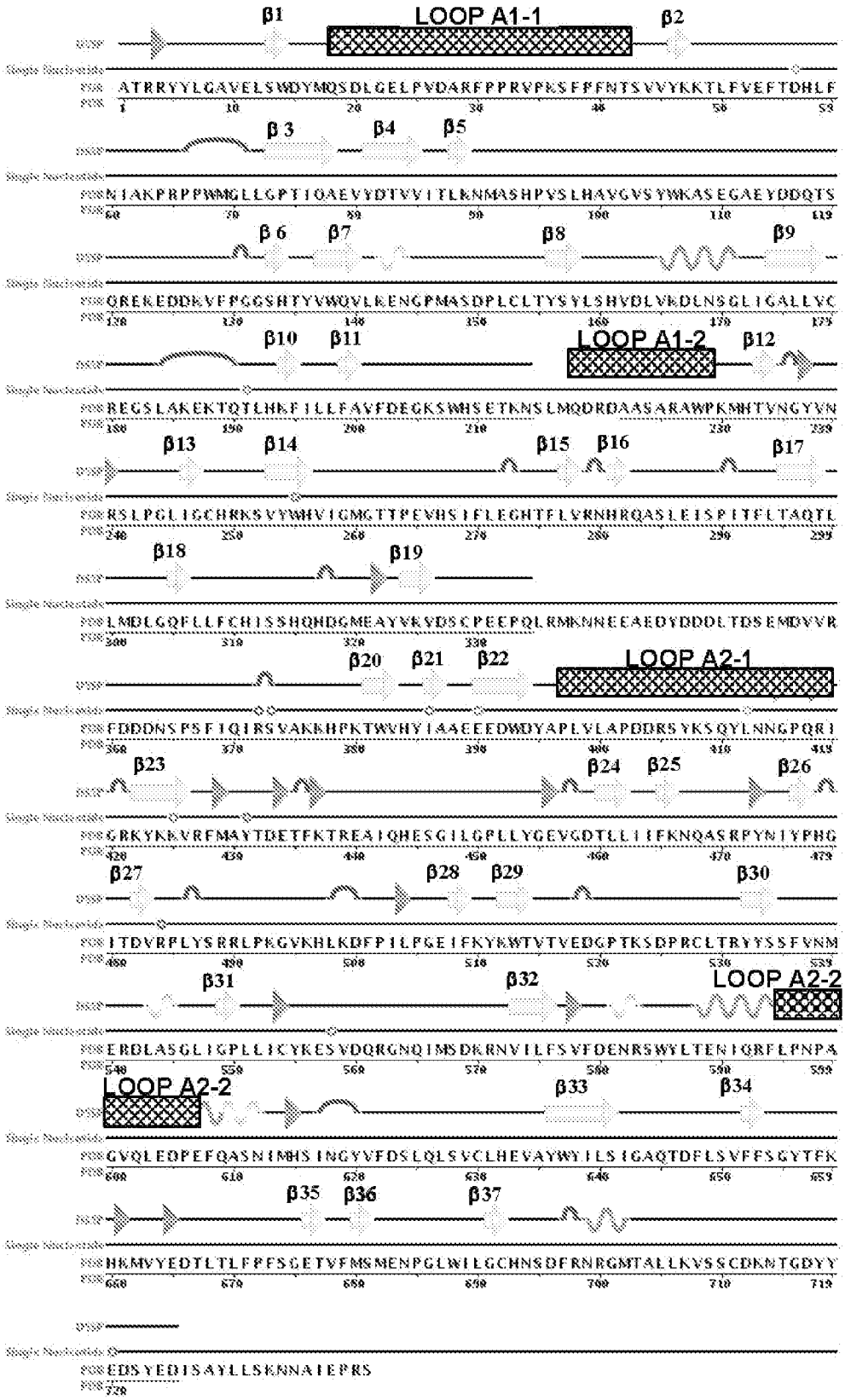

β1
LOOP A1-1
β2
β3
β4
β5
β6
β7
β8
β9
β10
β11
LOOP A1-2
β12
β13
β14
β15
β16
β17
β18
β19
β20
β21
β22
LOOP A2-1
β23
β24
β25
β26
β27
β28
β29
β30
β31
β32
LOOP A2-2
LOOP A2-2
β33
β34
β35
β36
β37

FIG. 38

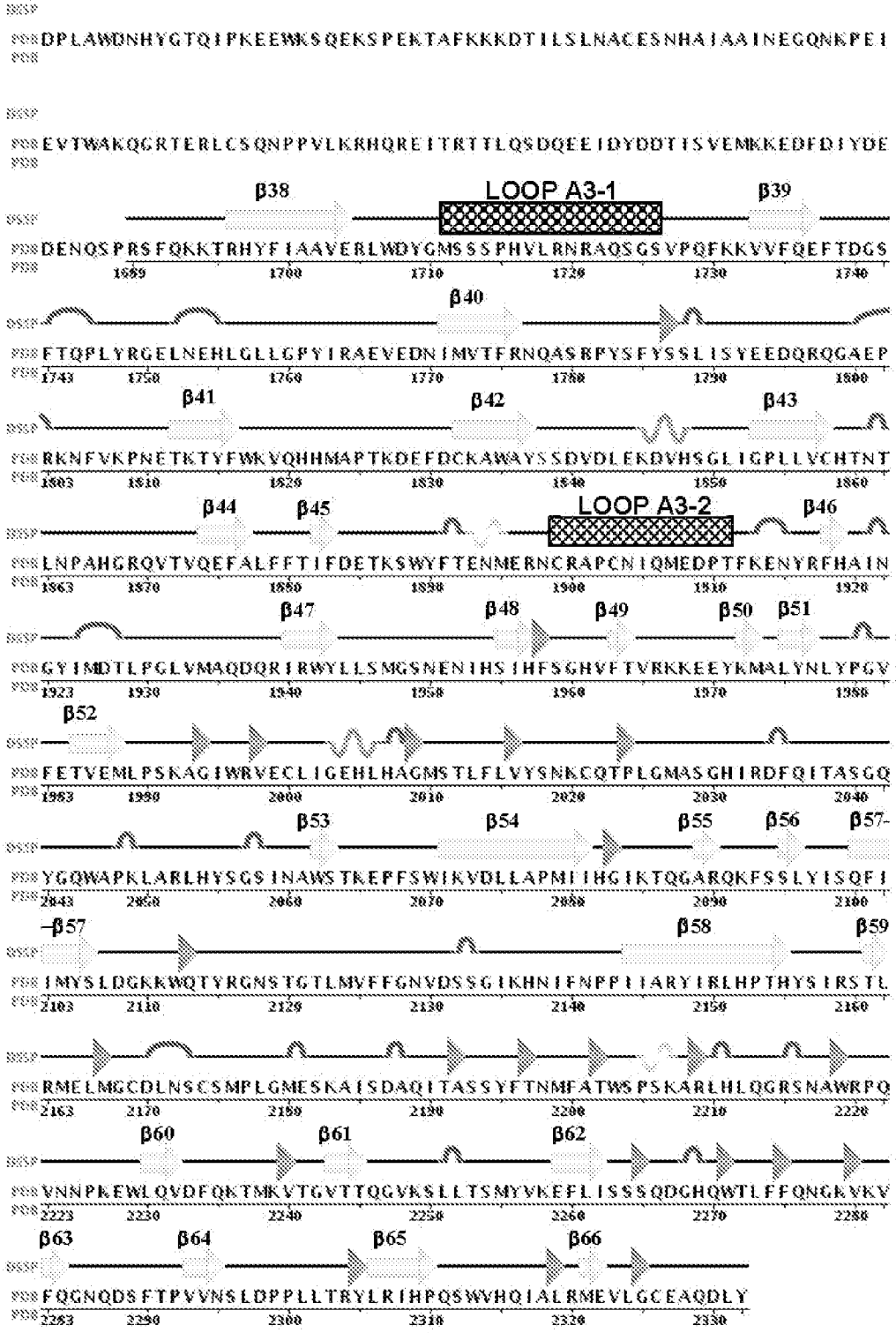

β38
LOOP A3-1
β39
β40
β41
β42
β43
β44
β45
LOOP A3-2
β46
β47
β48
β49
β50
β51
β52
β53
β54
β55
β56
β57
β58
β59
β60
β61
β62
β63
β64
β65
β66

FIG. 39

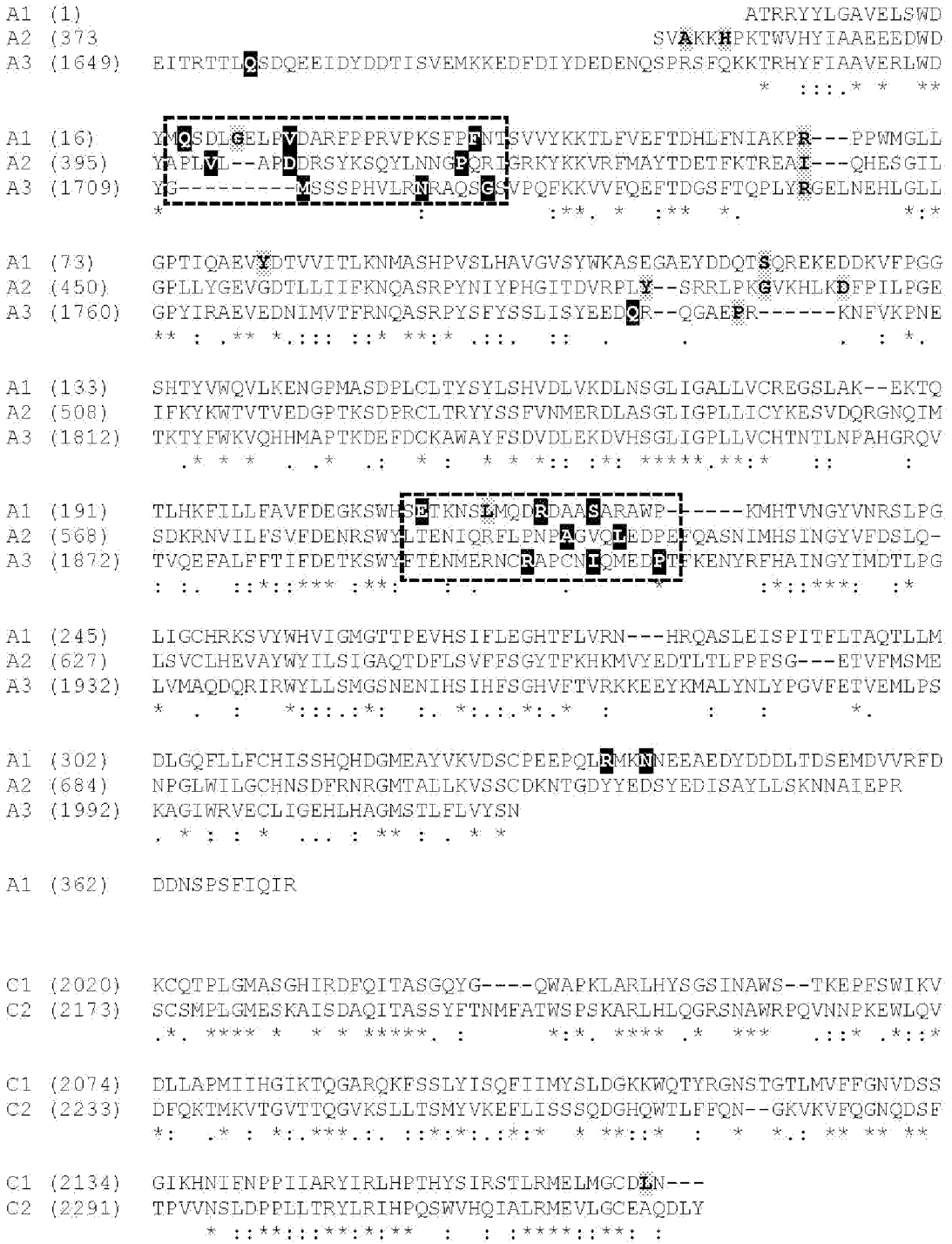

A1 (1) ATRRYYLGAVELSWD
A2 (373 SVAKKHPKTWVHYIAAEEEDWD
A3 (1649) EITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWD
A1 (16) YMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTLFVEFTDHLFNIAKPR---PPWMGLL
A2 (395) YAPLVL--APDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAI---QHESGIL
A3 (1709) YG---------MSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLL
A1 (73) GPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGG
A2 (450) GPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLY--SRRLPKGVKHLKDFPILPGE
A3 (1760) GPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQR--QGAEPR------KNFVKPNE
A1 (133) SHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAK--EKTQ
A2 (508) IFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIM
A3 (1812) TKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQV
A1 (191) TLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWP-L----KMHTVNGYVNRSLPG
A2 (568) SDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQ-
A3 (1872) TVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPG
A1 (245) LIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRN---HRQASLEISPITFLTAQTLLM
A2 (627) LSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSG---ETVFMSME
A3 (1932) LVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPS
A1 (302) DLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFD
A2 (684) NPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPR
A3 (1992) KAGIWRVECLIGEHLHAGMSTLFLVYSN
A1 (362) DDNSPSFIQIR
C1 (2020) KCQTPLGMASGHIRDFQITASGQYG----QWAPKLARLHYSGSINAWS--TKEPFSWIKV
C2 (2173) SCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQV
C1 (2074) DLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSS
C2 (2233) DFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQN--GKVKVFQGNQDSF
C1 (2134) GIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLN---
C2 (2291) TPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

FIG. 40

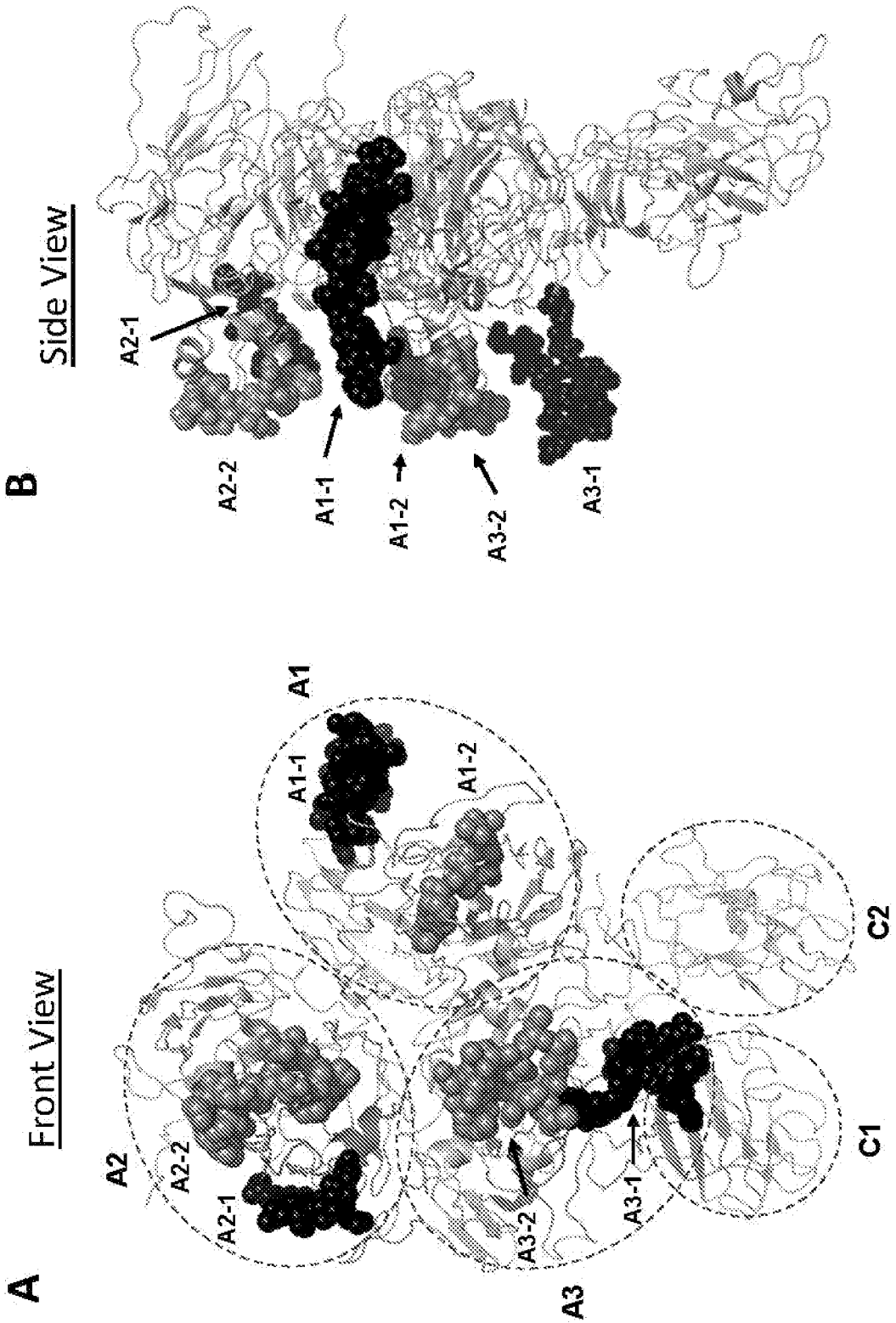

A
Front View
A1
A1-1
A1-2
A2
A2-2
A2-1
A3
A3-2
A3-1
C1
C2
B
Side View
A2-1
A2-2
A1-1
A1-2
A3-2
A3-1

FIG. 41

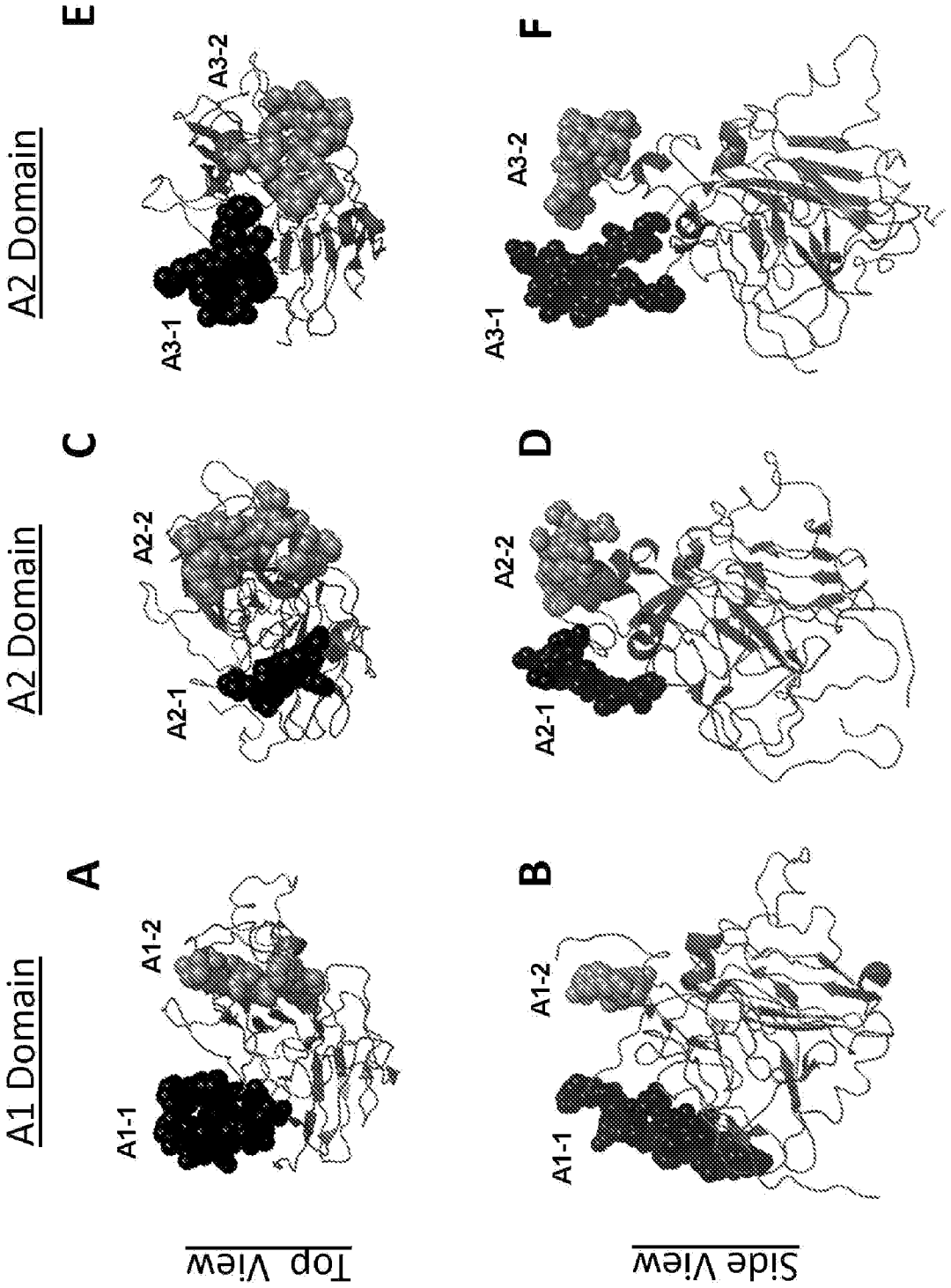

A1 Domain
A2 Domain
A2 Domain
A
B
C
D
E
F
A1-1
A1-2
A2-1
A2-2
A3-1
A3-2
Top View
Side View

FIG. 42

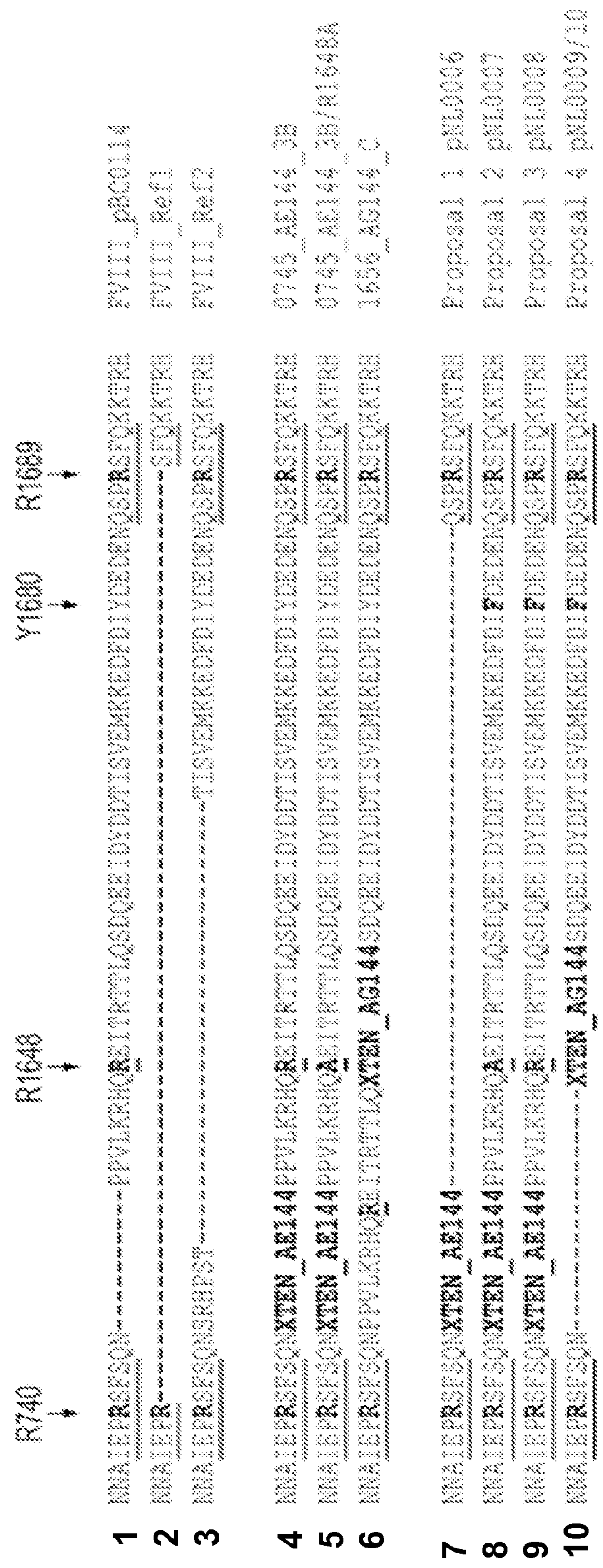

R740
R1648
Y1680
R1689
FVIII_pBC0114
FVIII_Ref1
FVIII_Ref2
0745_AE144_3B
0745_AE144_3B/R1648A
1656_AG144_C
Proposal 1 pNL0006
Proposal 2 pNL0007
Proposal 3 pNL0008
Proposal 4 pNL0009/10

FIG. 43

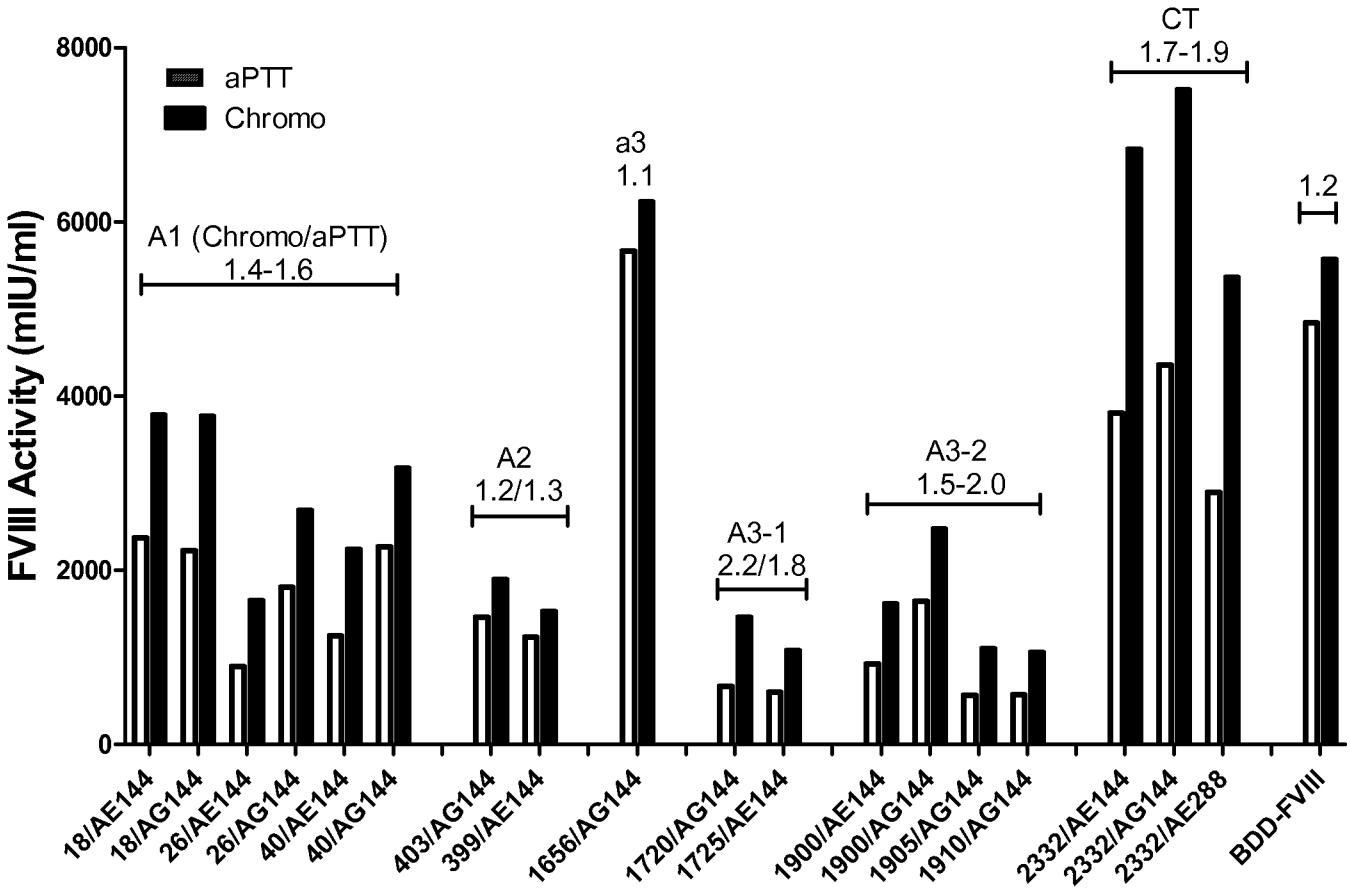

FVIII Activity (mIU/ml)
8000
6000
4000
2000
0
aPTT
Chromo
A1 (Chromo/aPTT)
1.4-1.6
A2
1.2/1.3
a3
1.1
A3-1
2.2/1.8
A3-2
1.5-2.0
CT
1.7-1.9
1.2
18/AE144
18/AG144
26/AE144
26/AG144
40/AE144
40/AG144
403/AG144
399/AE144
1656/AG144
1720/AG144
1725/AE144
1900/AE144
1900/AG144
1905/AG144
1910/AG144
2332/AE144
2332/AG144
2332/AE288
BDD-FVIII

FIG. 44

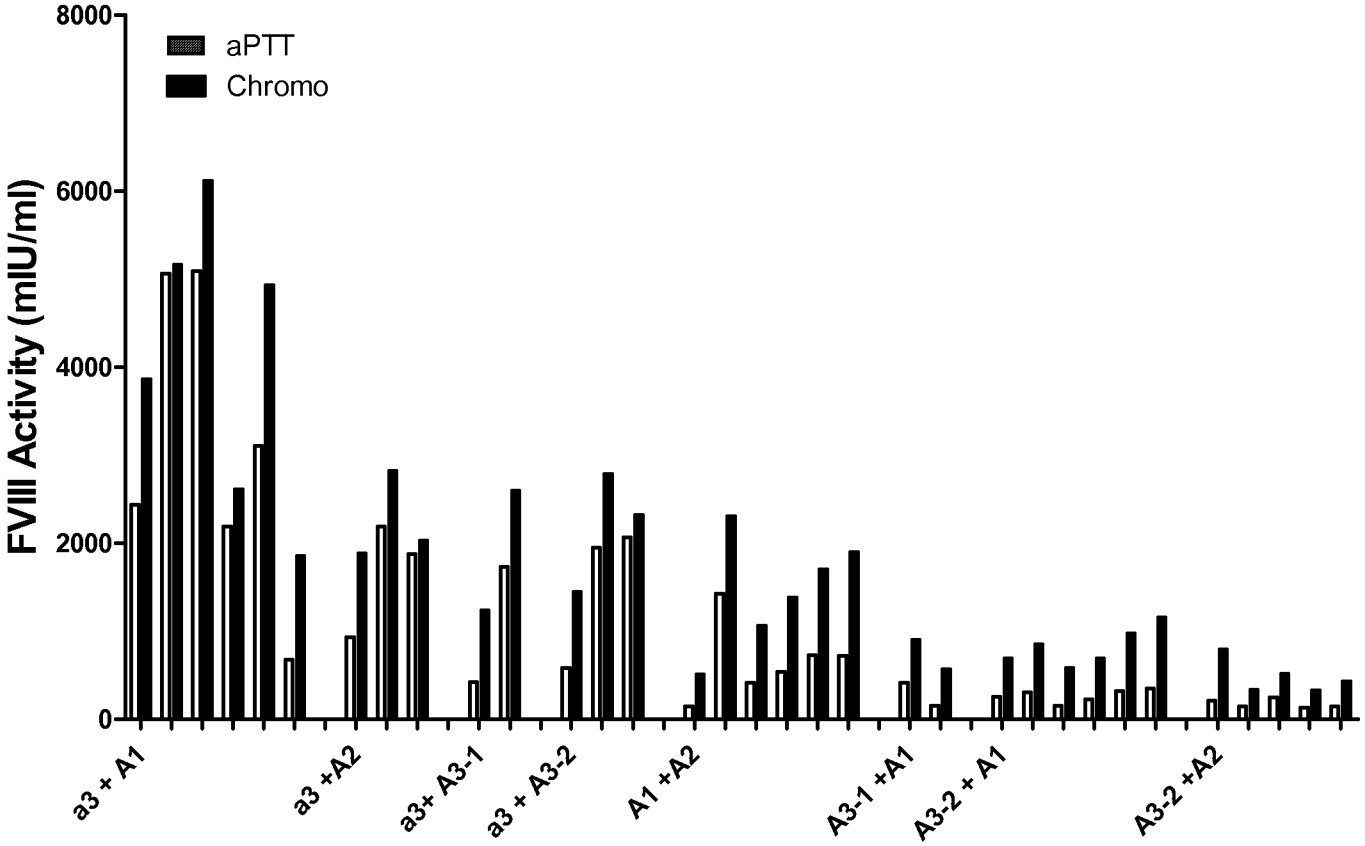

8000
6000
4000
2000
0
aPTT
Chromo
FVIII Activity (mIU/ml)
a3 + A1
a3 +A2
a3+ A3-1
a3 + A3-2
A1 +A2
A3-1 +A1
A3-2 + A1
A3-2 +A2

FIG. 45

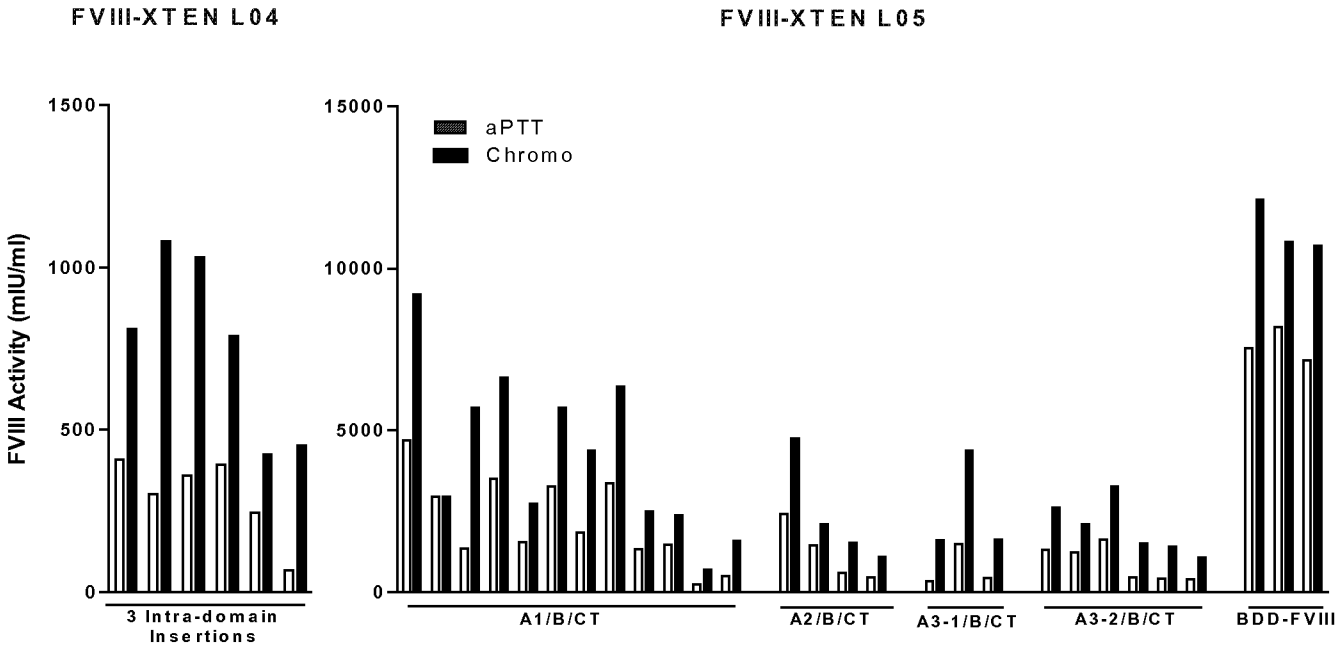

FVIII-XTEN L04
FVIII-XTEN L05
FVIII Activity (mIU/ml)
1500
1000
500
0
15000
10000
5000
0
aPTT
Chromo
3 Intra-domain Insertions
A1/B/CT
A2/B/CT
A3-1/B/CT
A3-2/B/CT
BDD-FVIII

FIG. 46

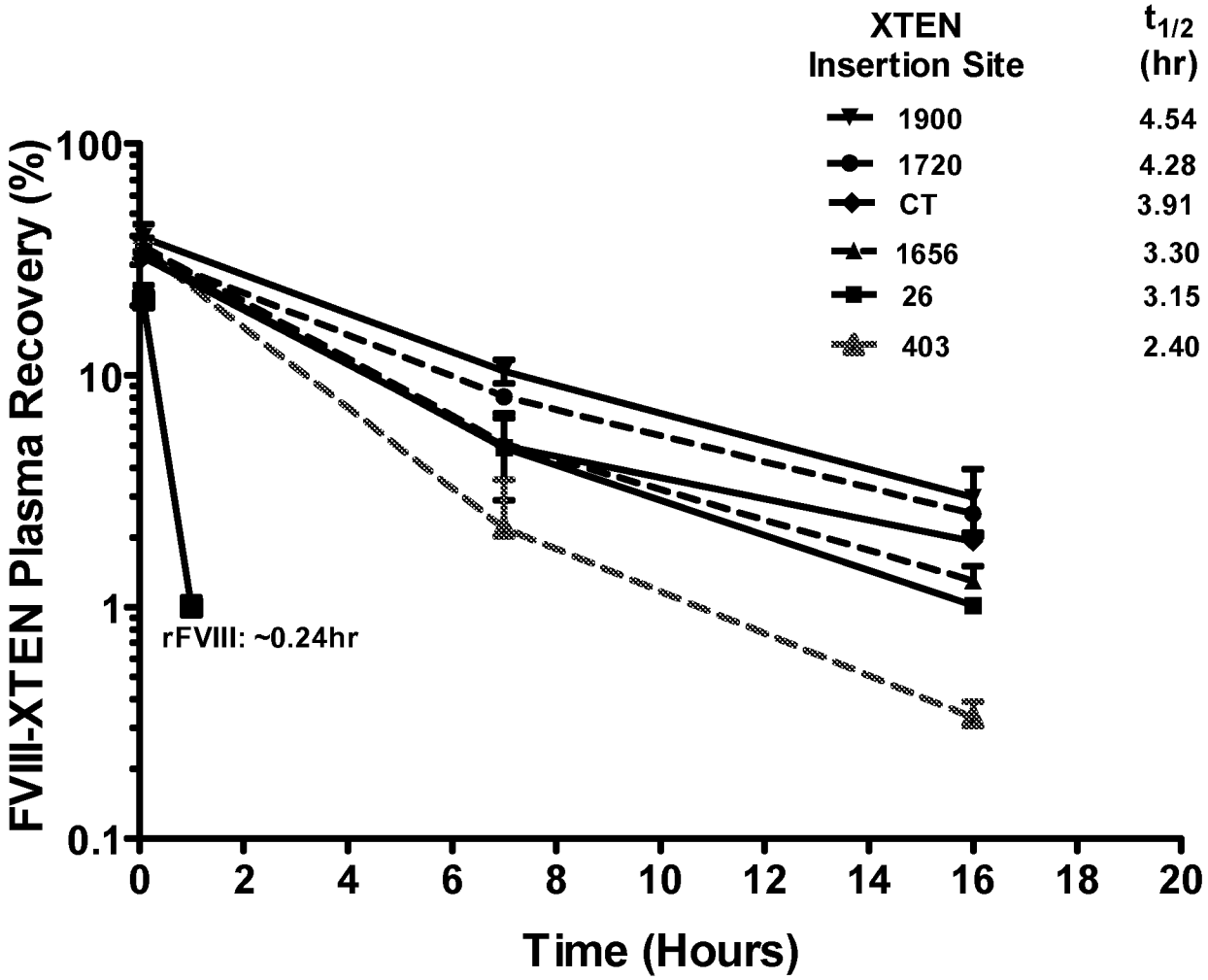

XTEN Insertion Site
t1/2 (hr)
1900 4.54
1720 4.28
CT 3.91
1656 3.30
26 3.15
403 2.40
FVIII-XTEN Plasma Recovery (%)
100
10
1
0.1
rFVIII: ~0.24hr
0 2 4 6 8 10 12 14 16 18 20
Time (Hours)

FIG. 47

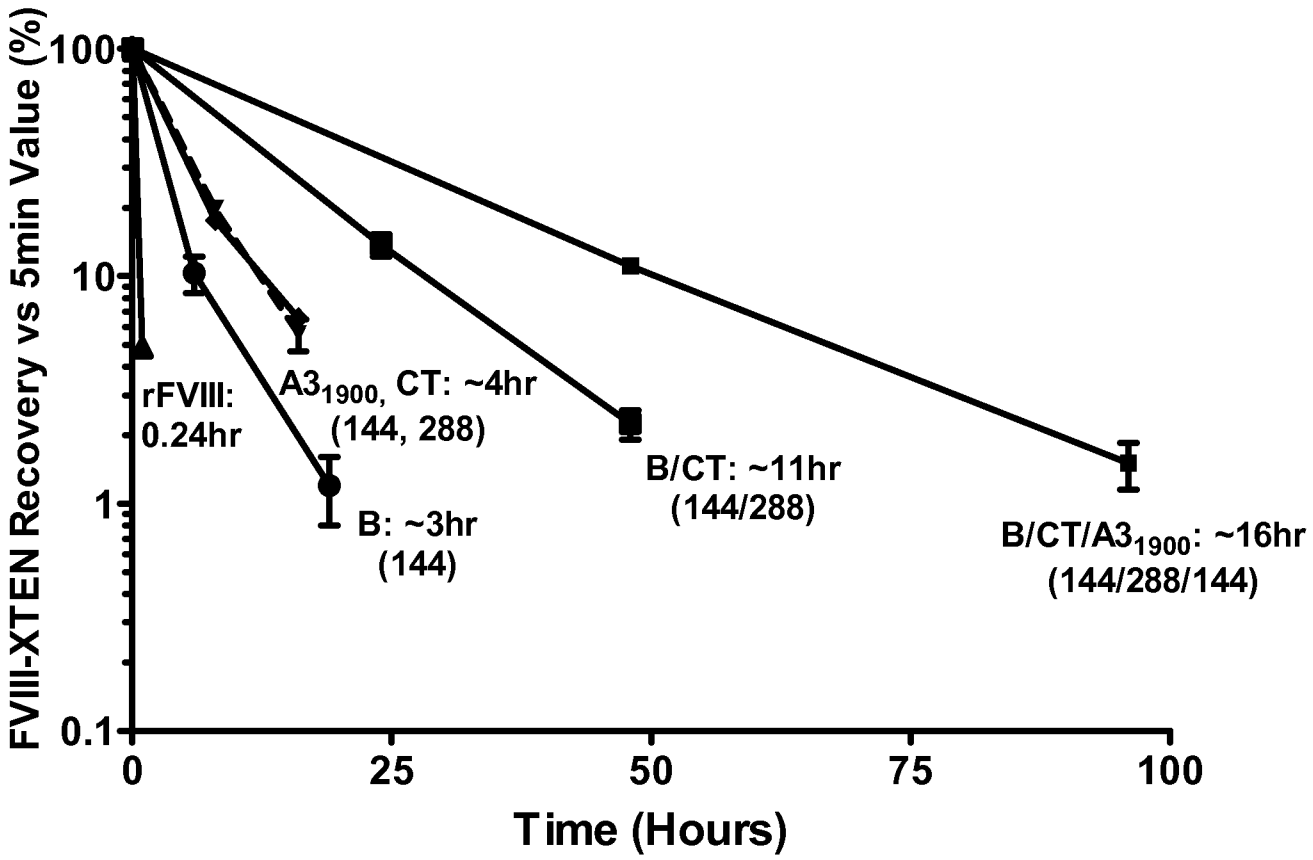
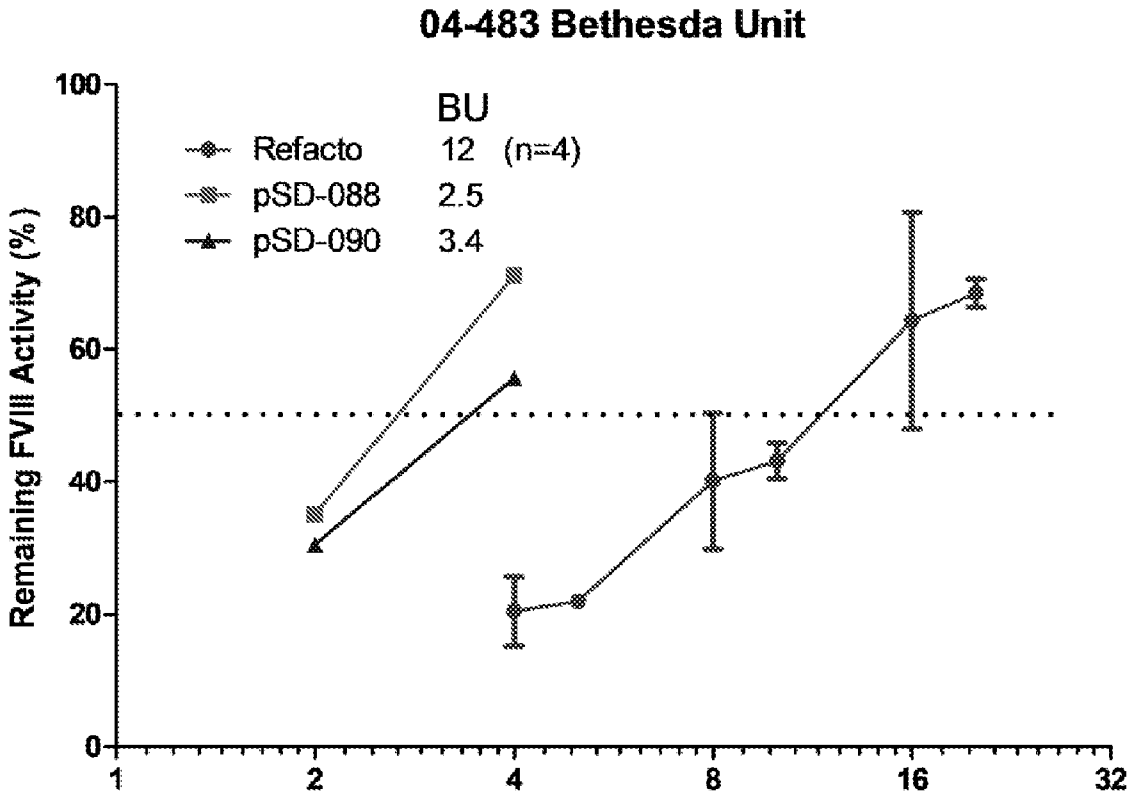

FVIII-XTEN Recovery vs 5min Value (%)
100
10
1
0.1
0
25
50
75
100
Time (Hours)
rFVIII:
0.24hr
A3$_{1900}$, CT: ~4hr
(144, 288)
B: ~3hr
(144)
B/CT: ~11hr
(144/288)
B/CT/A3$_{1900}$: ~16hr
(144/288/144)

FIG. 48

A


04-483 Bethesda Unit
BU
Refacto 12 (n=4)
pSD-088 2.5
pSD-090 3.4
Remaining FVIII Activity (%)
100
80
60
40
20
0
1
2
4
8
16
32

B

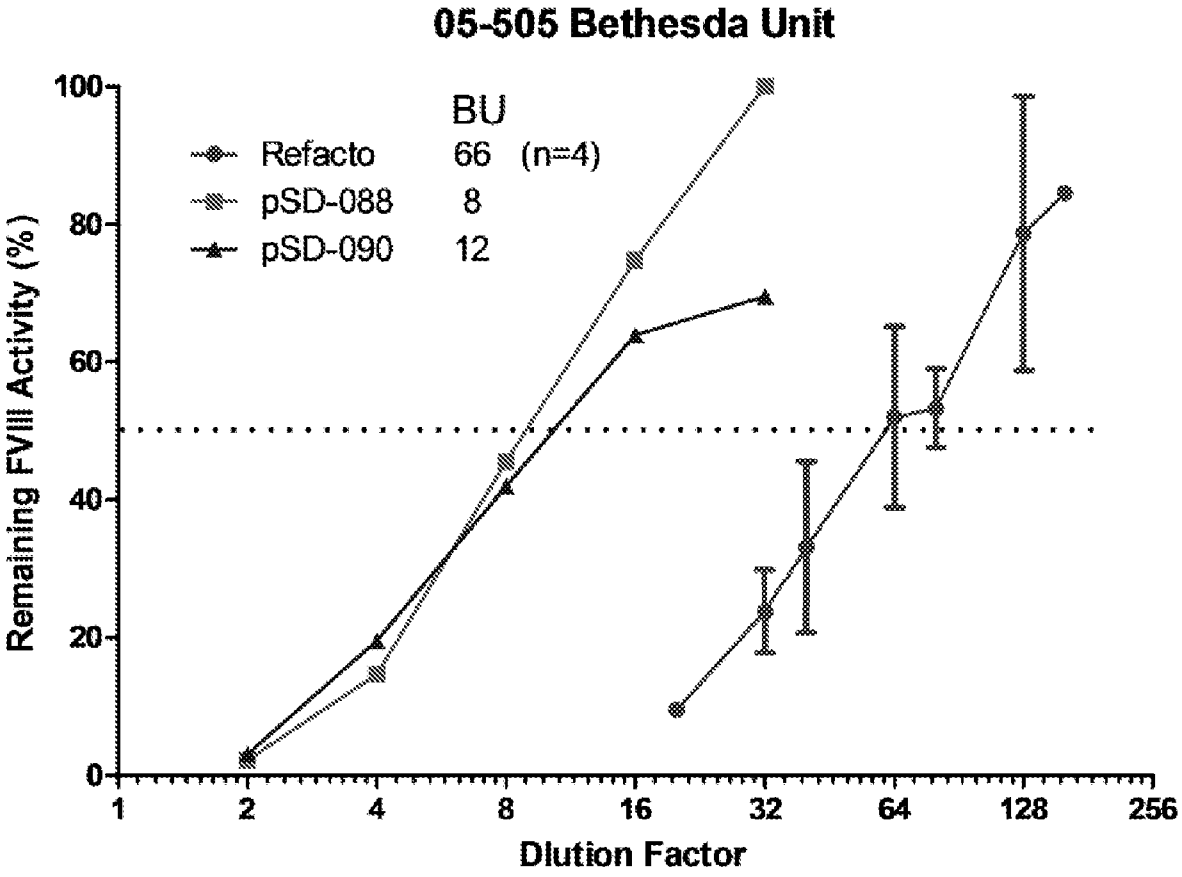
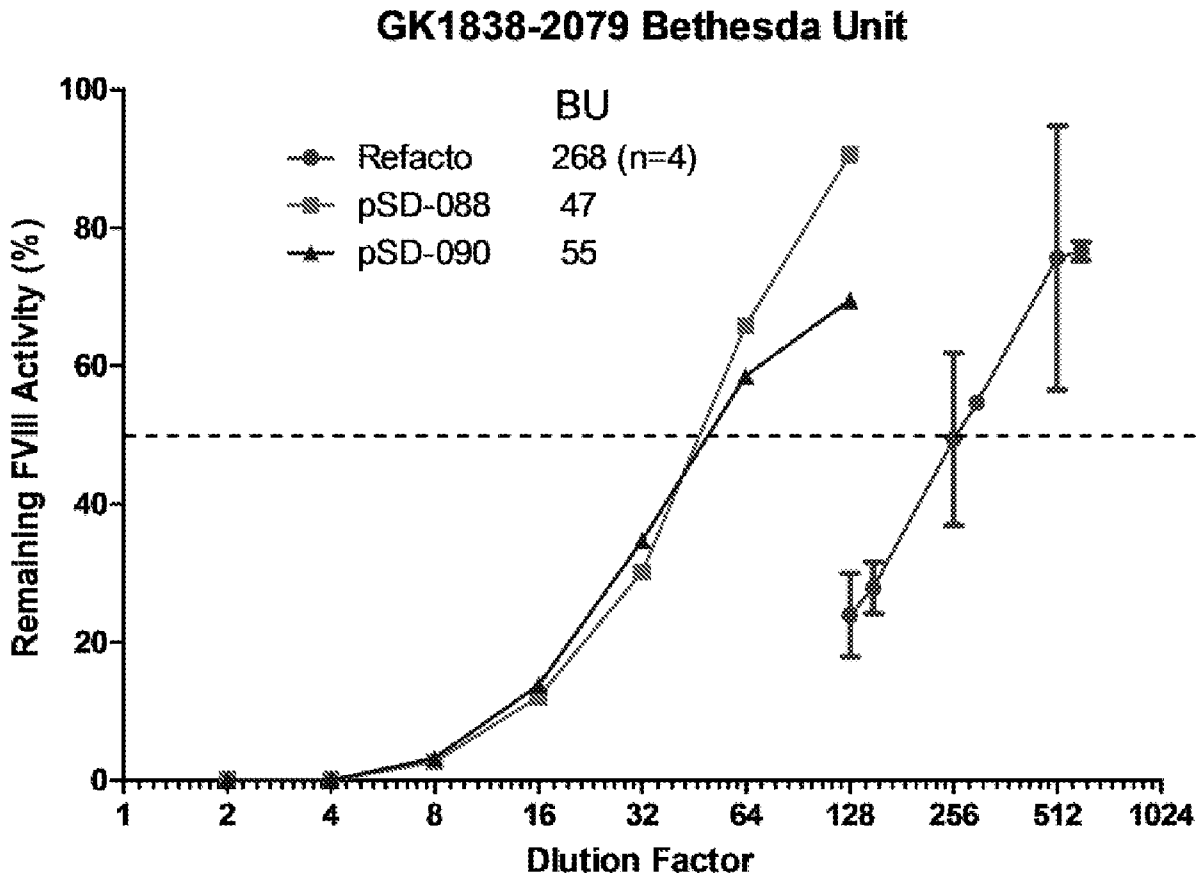

05-505 Bethesda Unit
BU
Refacto 66 (n=4)
pSD-088 8
pSD-090 12
Remaining FVIII Activity (%)
100
80
60
40
20
0
1
2
4
8
16
32
64
128
256
Dlution Factor

FIG. 49

C


GK1838-2079 Bethesda Unit
BU
Refacto 268 (n=4)
pSD-088 47
pSD-090 55
Remaining FVIII Activity (%)
100
80
60
40
20
0
1
2
4
8
16
32
64
128
256
512
1024
Dilution Factor

D

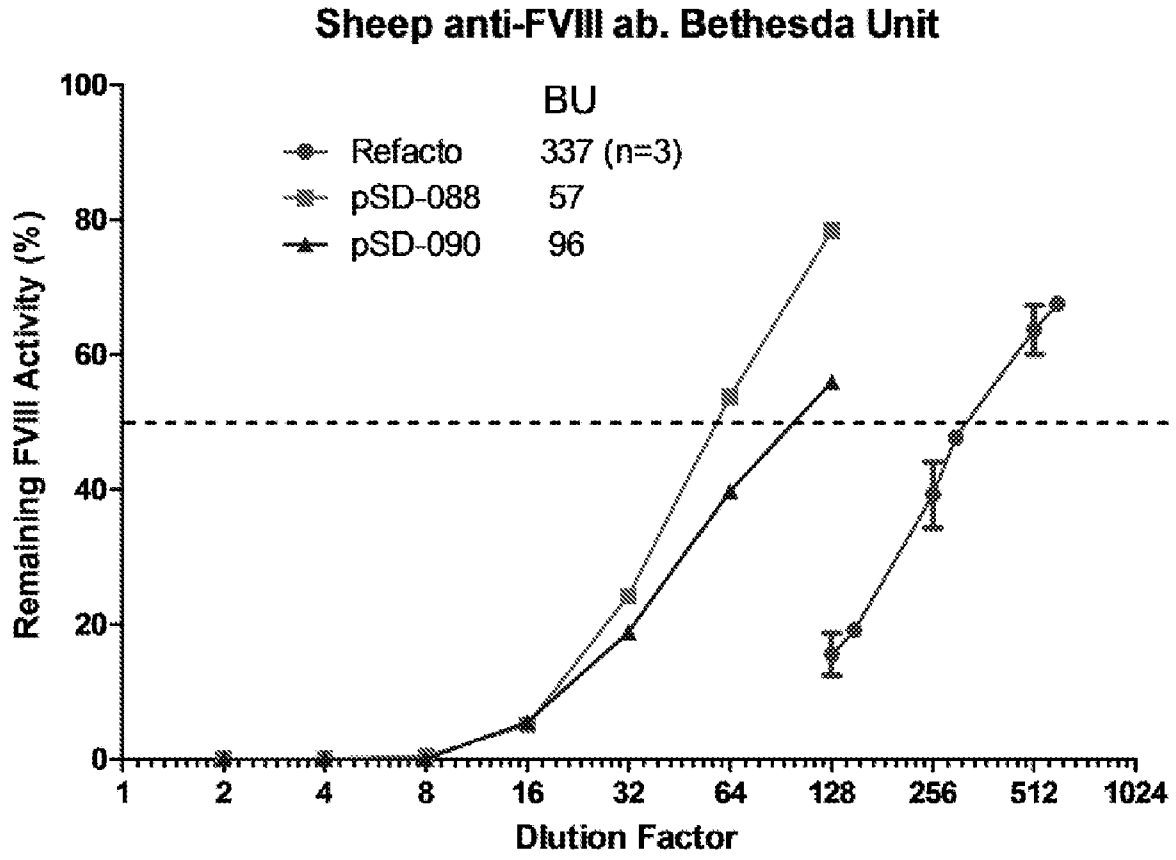

Sheep anti-FVIII ab. Bethesda Unit
BU
Refacto 337 (n=3)
pSD-088 57
pSD-090 96
Remaining FVIII Activity (%)
100
80
60
40
20
0
1
2
4
8
16
32
64
128
256
512
1024
Dilution Factor

FIG. 49