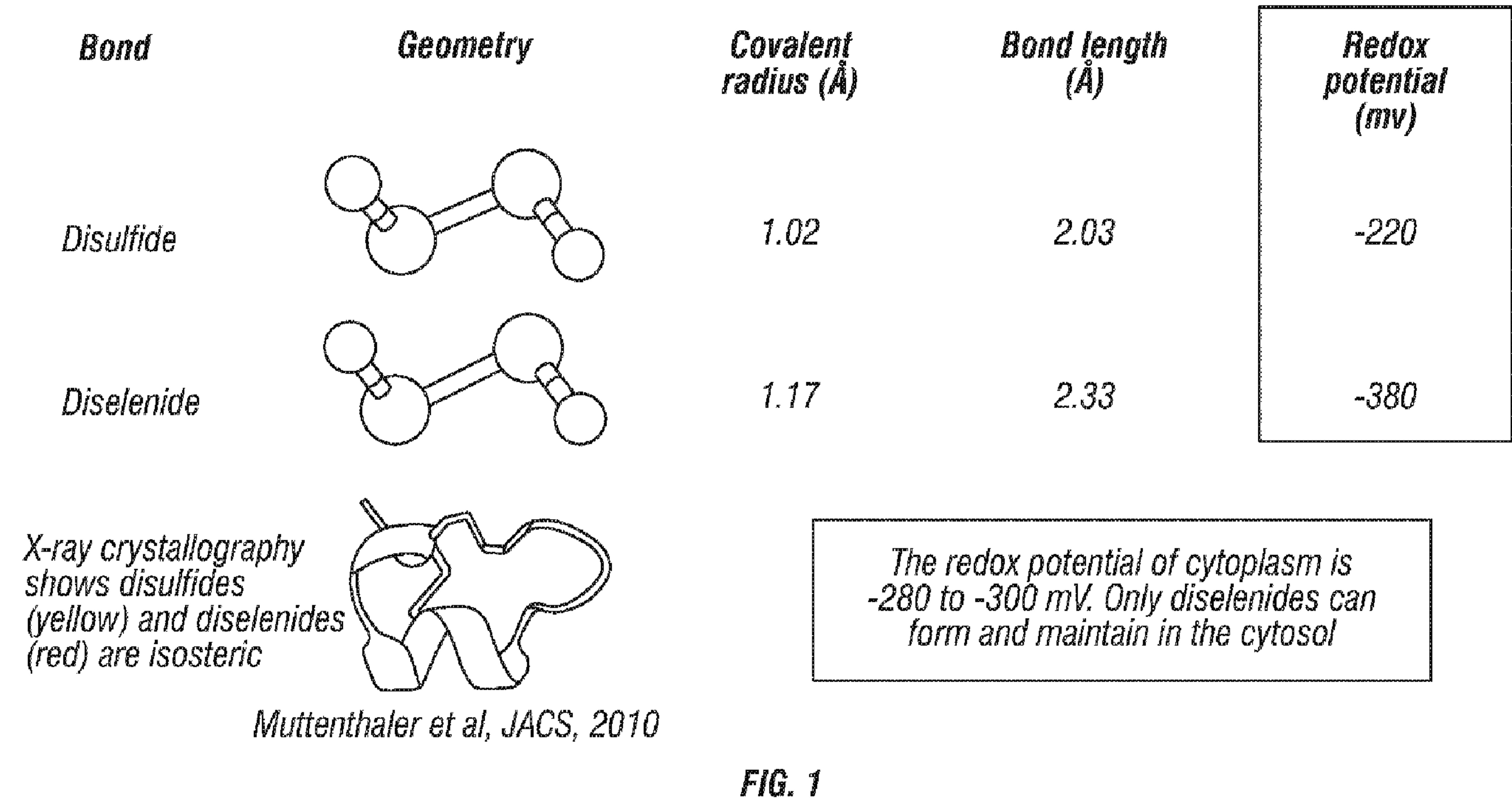




(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION
(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2017/08/30	(51) Cl.Int./Int.Cl. <i>A61K 39/395</i> (2006.01), <i>A61K 38/28</i> (2006.01), <i>C12N 15/09</i> (2006.01), <i>C12N 15/10</i> (2006.01), <i>C12N 15/11</i> (2006.01), <i>C12P 21/02</i> (2006.01), <i>C12P 21/08</i> (2006.01)
(87) Date publication PCT/PCT Publication Date: 2018/03/08	(71) Demandeurs/Applicants: BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM, US; GREGG, CHRISTOPHER J., US; PRESIDENT AND FELLOWS OF HARVARD COLLEGE, US
(85) Entrée phase nationale/National Entry: 2019/02/21	(72) Inventeurs/Inventors: GREGG, CHRISTOPHER J., US; MANDELL, DANIEL J., US; ...
(86) N° demande PCT/PCT Application No.: US 2017/049354	
(87) N° publication PCT/PCT Publication No.: 2018/045018	
(30) Priorités/Priorities: 2016/08/30 (US62/381,316); 2017/07/28 (US62/537,986)	

(54) Titre : PRODUCTION DE SELENO-BIOLOGIQUES DANS DES ORGANISMES GENETIQUEMENT RECODES
(54) Title: PRODUCTION OF SELENO-BIOLOGICS IN GENOMICALLY RECODED ORGANISMS



(57) **Abrégé/Abstract:**
Polypeptides that fold into biologies are stabilized by diselenide bonds between selenocysteine amino acids. Methods to produce such polypeptides in genomically recoded organisms (GRO) can be scaled up for industrial production. Since diselenides have the same geometric bond angles and torsions as disulfides, as well as very similar bond lengths, they can be substituted into polypeptides without disrupting the three dimensional structure of the polypeptides. Diselenides render the polypeptides resistant to reduction when they are exposed to blood serum or to reducing components of blood serum or to reducing components within cells.

(21) **3 034 701**

(13) **A1**

(72) **Inventeurs(suite)/Inventors(continued):** THYER, ROSS, US; ELLINGTON, ANDREW D., US;
STRANGES, PETER BENJAMIN, US

(74) **Agent:** BERESKIN & PARR LLP/S.E.N.C.R.L.,S.R.L.

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

CORRECTED VERSION

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
08 March 2018 (08.03.2018)



(10) International Publication Number
WO 2018/045018 A8

- (51) International Patent Classification:

A61K 39/395 (2006.01)

A61K 38/28 (2006.01)

C12P 21/02 (2006.01)

C12P 21/08 (2006.01)

C12N 15/09 (2006.01)

C12N 15/10 (2006.01)

C12N 15/11 (2006.01)

G06F 19/16 (2011.01)
- (21) International Application Number:

PCT/US2017/049354
- (22) International Filing Date:

30 August 2017 (30.08.2017)
- (25) Filing Language:

English
- (26) Publication Language:

English
- (30) Priority Data:

62/381,316

62/537,986

30 August 2016 (30.08.2016)

28 July 2017 (28.07.2017)

US

US
- (71) Applicants: BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM [US/US]; 210 West 7th Street, Austin, TX 78701 (US). PRESIDENT AND FELLOWS OF HARVARD COLLEGE [US/US]; 17 Quincy Street, Cambridge, MA 02138 (US).
- (72) Inventor; and

(71) Applicant: GREGG, Christopher, J. [US/US]; 28 Cumner Avenue, Melrose, MA 02176 (US).
- (72) Inventors: MANDELL, Daniel, J.; 131 Fuller Street, Unit 3, Brookline, MA 02446 (US). THYER, Ross; c/o University of Texas at Austin, 2500 Speedway MBB 3.424, Austin, TX 78712 (US). ELLINGTON, Andrew, D.; 2106 Elton Lane, Austin, TX 78703 (US). STRANGES, Peter Benjamin; 111 Woodstock Street, Apt. 5, Somerville, MA 02144 (US).

(54) Title: PRODUCTION OF SELENO-BIOLOGICS IN GENOMICALLY RECODED ORGANISMS

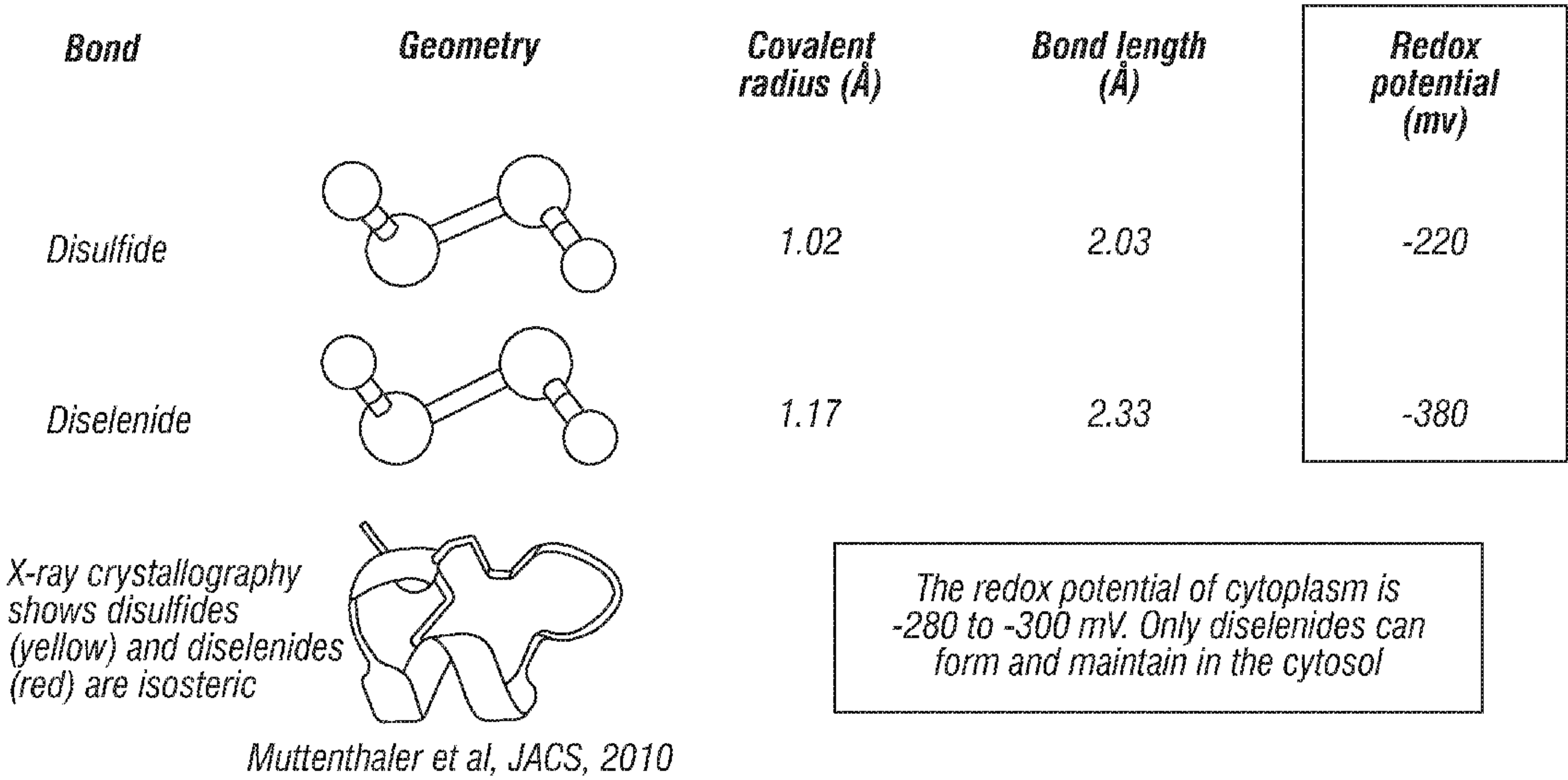


FIG. 1

(57) Abstract: Polypeptides that fold into biologies are stabilized by diselenide bonds between selenocysteine amino acids. Methods to produce such polypeptides in genomically recoded organisms (GRO) can be scaled up for industrial production. Since diselenides have the same geometric bond angles and torsions as disulfides, as well as very similar bond lengths, they can be substituted into polypeptides without disrupting the three dimensional structure of the polypeptides. Diselenides render the polypeptides resistant to reduction when they are exposed to blood serum or to reducing components of blood serum or to reducing components components within cells.

WO 2018/045018 A8



(74) Agent: BYRD, Marshall, P.; Parker Highlander PLLC,
1120 S. Capital of Texas Highway, Bldg. One, Suite 200,
Austin, TX 78746 (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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, ST, 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, KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*

(48) Date of publication of this corrected version:

11 May 2018 (11.05.2018)

(15) Information about Correction:

see Notice of 11 May 2018 (11.05.2018)

DESCRIPTION

PRODUCTION OF SELENO-BIOLOGICS IN GENOMICALLY RECODED ORGANISMS

[01] This application claims the benefit of United States Provisional Patent Application
5 Nos. 62/381,316, filed August 30, 2016, and 62/537,986, filed July 28, 2017, the
entirety of which are incorporated herein by reference.

[02] The sequence listing that is contained in the file named “UTFB_P1174WO.txt”, which
is 3305 KB (as measured in Microsoft Windows) and was created on August 30,
2017, is filed herewith by electronic submission and is incorporated by reference
10 herein.

TECHNICAL FIELD OF THE INVENTION

[03] This invention was made with government support under Grant no. FA9550-10-1-
0169 awarded by Office of Naval Research and Grant no. N66001-14-2-4051
awarded by the Space and Naval Warfare Systems Center, Pacific. The government
15 has certain rights in the invention.

[04] This invention is related to the area of protein biologics. In particular, it relates to
methods to make protein biologics with beneficial properties.

BACKGROUND OF THE INVENTION

[05] Protein biologics, which include monoclonal antibodies, protein therapeutics and
20 vaccines, are effective therapies for cancer, infection, and immunological and other
diseases, comprising a multi-billion dollar market. But most biologics must be made in
expensive mammalian cell systems in large part because they are stabilized by
chemical bonds – called disulfides – that cannot form in microbial cells. Further,
because disulfide bonds are prone to chemical reduction, they can break in
25 environments like blood serum and inside cells, so that many biologics have poor half-
lives.

[06] The vast majority of biologics – including 98% of therapeutic monoclonal antibodies – are produced in mammalian systems like Chinese hamster ovary (CHO) cells. CHO cells have the advantage of producing human-like glycosylation, and possess the necessary folding chaperones and redox machinery to produce disulfide-bearing proteins at high yield. However, CHO cells require substantially higher capital expenditures than microbial facilities, longer production times (weeks to months vs. days), many-fold higher consumables cost (*e.g.*, media), and additional purification steps to remove endogenous retrovirus-like particles.

[07] There are two existing approaches to producing disulfide-bearing biologics in bacteria. The first is to change the chemical environment inside the cell to become oxidizing and permissive to disulfide formation, but this change dramatically impacts central metabolism and may impair fermentation performance as well as protein production. The second approach is to secrete the nascent biologic to a smaller oxidizing compartment called the periplasm. This approach produces poor yields due to energetic losses during export, and also is highly susceptible to product loss through aggregation. Due to these shortcomings, neither approach has gained widespread adoption for large-scale production of disulfide-bearing biologics, so that more expensive mammalian bio-processes dominate. Importantly, neither approach solves the intrinsic problem that disulfide bonds are sensitive to reducing agents in environments like the cytoplasm or blood.

[08] There is a continuing need in the art for methods of producing protein biologics in a cost effective way. There is a continuing need in the art for method of producing protein biologics that are stable in blood and cytoplasm.

SUMMARY OF THE INVENTION

[09] According to one aspect of the invention a protein is provided which comprises at least two selenocysteine residues with at least one diselenide bond between the at least two selenocysteine residues. The protein comprises a sequence selected from the group consisting of: SEQ ID NO: 1-1631 and sequences that are at least 80% identical to said sequence. The protein may comprise a sequence selected from the

group consisting of: SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence. The protein may comprise a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence.

5 [10] According to another aspect of the invention an antibody molecule is provided which comprises (i) a heavy chain sequence selected from the group consisting of SEQ ID NO: 472-801 and sequences that are at least 80% identical to said sequence; and (ii) a light chain sequence selected from the group consisting of SEQ ID NO: 812-1131 and sequences that are at least 80% identical to said sequence. From 0-15 pairs of selenocysteine residues in the antibody molecule are replaced with from 0-15 pairs of cysteine residues. The antibody may contain at least one diselenide bond between the heavy chain and the light chain. The antibody may contain at least one intra-chain diselenide bond within the heavy chain and/or within the light chain.

10 [11] According to yet another aspect of the invention a modified insulin protein is provided. It comprises (i) an A chain having a sequence selected from SEQ ID NO: 1142, 1144, 1146, 1148, 1150, 1152, and 1154 and sequences that are at least 80% identical to said sequence; and (ii) a B chain having a sequence selected from SEQ ID NO: 1143, 1145, 1147, 1149, 1151, 1153, and 1155 and sequences that are at least 80% identical to said sequence. The protein comprises at least one diselenide bond between the A chain and the B chain.

20 [12] According to one aspect of the invention a bacterial host is provided that has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon. The bacterial host contains a constructed DNA sequence encoding a variant form of a protein in which a pair of cysteine codons has been replaced with the unassigned codon. The variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1631 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence selected from the group consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence.

[13] According to another aspect of the invention a bacterial host is provided. The bacterial host has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon. The bacterial host contains a constructed DNA sequence encoding a variant form of a monoclonal antibody in which a pair of cysteine codons has been replaced with the unassigned codon. The monoclonal antibody comprises a VL chain, and positions 23 and 88 of the VL chain are substituted with selenocysteine residues. The monoclonal antibody comprises a VH chain, and positions 22 and 92 of the VH chain are substituted with selenocysteine residues.

10 [14] According to yet another aspect of the invention a method of producing a variant form of a protein is provided. The protein has a disulfide bond between cysteine residues, and the variant form has a diselenide bond between selenocysteine residues. A bacterial host that has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon is cultured. The bacterial host contains a constructed DNA sequence encoding the variant form of the protein in which a pair of cysteine codons has been replaced with the unassigned codon. The culturing is performed under conditions in which the variant form of the protein is expressed. The variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1631 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence selected from the group consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence.

25 [15] According to one aspect of the invention a method of producing a variant form of a monoclonal antibody is provided. The monoclonal antibody has a disulfide bond between cysteine residues, and the variant form has a diselenide bond between selenocysteine residues. A bacterial host that has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon is cultured. The bacterial host contains a constructed DNA sequence encoding the variant form of the monoclonal antibody in which a pair of cysteine codons has been replaced with the unassigned codon. The culturing is performed

under conditions in which the variant form of the monoclonal antibody is expressed. The monoclonal antibody comprises a VL chain, and positions 23 and 88 of the VL chain are substituted with selenocysteine residues. The monoclonal antibody comprises a VH chain and positions 22 and 92 of the VH chain are substituted with selenocysteine residues.

[16] According to another aspect of the invention a method of treating a mammal with a variant form of a protein is provided. The variant form of the protein is administered to the mammal. The protein has a disulfide bond between cysteine residues. The variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues. The variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1631 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence selected from the group consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence.

[17] According to yet another aspect of the invention a method is provided of treating a mammal with a variant form of a monoclonal antibody. The variant form of the monoclonal antibody is administered to the mammal. The monoclonal antibody has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues. The monoclonal antibody comprises a VL chain, and positions 23 and 88 of the VL chain are substituted with selenocysteine residues. The monoclonal antibody comprises a VH chain and positions 22 and 92 of the VH chain are substituted with selenocysteine residues.

[18] According to one aspect of the invention an *in vitro* translation system comprising a constructed nucleic acid molecule and a tRNA molecule are provided. The constructed nucleic acid molecule encodes a variant form of a protein in which a pair of cysteine residues has been replaced with a pair of selenocysteine residues. The variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1631 and sequences that are at least 80% identical to said sequence. The variant

form may comprise a sequence selected from the group consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence. The nucleic acid molecule comprises an unassigned codon or leverages suppression of cognate codons to incorporate selenocysteine(s) into a growing polypeptide. The tRNA molecule is a selenocysteine tRNA with an anticodon that corresponds to the unassigned codon.

[19] According to another aspect of the invention an *in vitro* translation system comprising a constructed nucleic acid molecule and a tRNA molecule is provided. The constructed nucleic acid molecule encodes a variant form of a monoclonal antibody which comprises a VL chain. Positions 23 and 88 of the VL chain are substituted with selenocysteine residues. The monoclonal antibody comprises a VH chain, and positions 22 and 92 of the VH chain are substituted with selenocysteine residues. The constructed nucleic acid molecule comprises at least a pair of cysteine codons that has been replaced with an unassigned codon in the nucleic acid. The tRNA molecule is a selenocysteine tRNA with an anticodon that is complementary to the unassigned codon.

[20] According to yet another aspect of the invention a method is provided for producing a variant form of a protein. The protein has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues. A constructed nucleic acid sequence encoding the variant form of the protein is translated. The constructed nucleic acid sequence comprises a pair of cysteine codons that has been replaced with an unassigned codon. The variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1631 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence selected from the group consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence. The step of translating employs a selenocysteine tRNA with a corresponding anticodon to the unassigned codon.

[21] According to one aspect of the invention a method of testing a sample is provided. A sample is contacted with a variant form of a protein to form a reaction mixture. The protein has a disulfide bond between cysteine residues. The variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues. The variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1631 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence selected from the group consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence. The variant form may comprise a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence. The reaction mixture is tested to detect a specific binding partner bound to the variant form of the protein.

[22] According to another aspect of the invention a method is provided for testing a sample. A sample is contacted with a variant form of a monoclonal antibody to form a reaction mixture. The monoclonal antibody has a disulfide bond between cysteine residues. The variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues. The monoclonal antibody comprises a VL chain, and positions 23 and 88 of the VL chain are substituted with selenocysteine residues. The monoclonal antibody comprises a VH chain and positions 22 and 92 of the VH chain are substituted with selenocysteine residues. The reaction mixture is tested to detect a specific binding partner bound to the variant form of the protein.

[23] According to yet another aspect of the invention a modified antibody is provided. It comprises at least one diselenide bond.

[24] According to one aspect of the invention a modified insulin protein is provided. It comprises an A chain having residues 90-110 of SEQ ID NO: 1630 with at least one cysteine to selenocysteine substitution; a B chain having residues 25-54 of SEQ ID NO: 1630 with at least one cysteine to selenocysteine substitution; and at least one diselenide bridge between the A chain and the B chain.

[25] Another aspect of the invention is a modified human insulin protein comprising at least two selenocysteine residues replacing two cysteine residues in a native human

insulin protein. It has at least one diselenide bond between the at least two selenocysteine residues.

- [26] These and other aspects which will be apparent to those of skill in the art upon reading the specification provide the art with methods and products for treating human and animal diseases and for making and using therapeutic, diagnostic, analytical, and prophylactic agents.

BRIEF DESCRIPTION OF THE DRAWINGS

- [27] Fig. 1 shows a comparison of properties of diselenide and disulfide bonds.
- [28] Fig. 2 shows that diselenide biologics show dramatic improvements to serum half-life over disulfides.
- [29] Fig. 3A-3B shows zoomed-in deconvoluted mass spectrum (Fig. 3A) and fragmentation sequence map (Fig. 3B) of 10 μ M anti-MS2 scFv. Fragmentation spectra were deconvoluted using Xtract with a S/N threshold of 3 and analyzed using ProSight Lite. Formation of diselenide bonds is indicated by the lack of fragment ions between the marked selenocysteine (U) residues. No mass peaks corresponding to serine incorporation were detected.
- [30] Fig. 4 shows protein expression for 6 different selenobiologics. Sequences 1-6 on the gel correspond to SEQ ID NO: 1620-1625 in the Sequence Listing. Diselenide bonds are at: SEQ ID NO: 1620 SeCys23-SeCys97 and SeCys138-SeCys192; SEQ ID NO: 1621 SeCys23-SeCys97, SeCys138-SeCys192, SeCys267-SeCys327, and SeCys373-SeCys431; SEQ ID NO: 1622 SeCys24-SeCys98; SEQ ID NO: 1623 SeCys24-SeCys98, SeCys164-SeCys224, and SeCys270-SeCys328; SEQ ID NO: 1624 SeCys32-SeCys92 and SeCys138-SeCys196; and SEQ ID NO: 1625 SeCys4-SeCys28 and SeCys86-SeCys92.
- [31] Fig. 5 shows protein expression for 2 different selenobiologics. Sequences 7-8 on the gel correspond to SEQ ID NO: 1626-1627 in the Sequence Listing. Diselenide bonds are at SEQ ID NO: 1626 SeCys183-SeCys190, SeCys210-SeCys234, and

SeCys298; and SEQ ID NO: 1627 SeCys183-SeCys190, SeCys238-SeCys298, and SeCys344-SeCys402

[32] Fig. 6 shows protein expression for 2 different selenobiologics. Sequences 9-10 on the gel correspond to SEQ ID NO: 1628-1629 in the Sequence Listing. Diselenide bonds are at SEQ ID NO: 1628 SeCys54-SeCys166 and SeCys183-SeCys190; and SEQ ID NO: 1629 SeCys24-SeCys89 and SeCys149-SeCys244.

[33] Fig. 7A-7B shows ELISA data for wild-type (disulfide) and seleno (diselenide) anti-RCA scFv in different redox conditions. Wild-type (**A**) and seleno (SEQ ID NO: 1631, **B**) -anti-RCA scFv were diluted to 1 μ M in TBS pH 7.5 with 0, 1, 10 or 50 mM DTT and incubated for 16 hours at 37 °C. Ricin A-chain (Sigma L9514) was diluted to 10 μ g/mL in PBS and bound to a 96-well plate overnight at 4 °C. Control plates were bound with PBS containing 5% w/v skim milk. Plates were washed with PBS and blocked with PBS 5% w/v skim milk for two hours. Following incubation with DTT, scFvs were diluted in PBS with 5% w/v skim milk and 0.05% Tween 20 to 0, 0.25, 0.5, 1, 1.5, 2, 4, 6, 8, 10, 15 and 20 nM and bound in quadruplicate for two hours at 25 °C. Secondary antibody (anti-polyhistidine-HRP, Sigma A7058) was diluted 1/10,000 in PBS with 5% w/v skim milk and 0.05% Tween 20 and bound for two hours at 25 °C. Plates were incubated with TMB for 10 minutes at 25 °C with agitation and color development terminated by the addition of 2M H₂SO₄. Absorbance was immediately measured at 450 nm. All binding, blocking and enzymatic steps were performed on an orbital shaker at 450 rpm, and all washes were performed with PBS or PBS with 0.05% Tween 20. Only wild-type anti-RCA scFv showed substantial loss of binding at 1 and 10 mM DTT, and loss of binding was substantially abrogated at all concentrations for seleno-anti-RCA compared to WT scFv.

[34] Fig. 8A-8B shows thermal shift data for wild-type (disulfide) and seleno (diselenide) variants of 2 different biologics. Thermal denaturation of purified wild-type and seleno (SEQ ID NO: 1132) hGH (**A**) and wild-type and seleno (SEQ ID NO: 1629) Herceptin scFv (**B**) in different reducing conditions was monitored using the Applied Biosystems Protein Thermal Shift dye on a QuantStudio 3 Real Time PCR system.

The reducing agent DTT was added to triplicates of 5 μ g of hGH-Cys/SeCys and 2 μ g of Herceptin scFv-Cys/SeCys at concentrations of 0 M and 10 mM followed by incubation for 10 minutes at room temperature. The proteins were denatured in a linear ramp of 0.03°C/s from 25°C to 98°C. Melting temperature (T_m) was determined by finding the minimum of dF/dT with Protein Thermal Shift Software (ThermoFisher). The hGH samples (**A**) showed a single minimum at both concentrations of DTT. The T_m for hGH-Cys decreased from 77 °C without DTT to 60 °C in the presence of 10 mM DTT. The SeCys variant showed a marked stabilization under reducing conditions compared to the Cys variant, with a T_m of 75 °C without DTT and a T_m of 66 °C in 10 mM DTT. For Herceptin (**B**) the Cys variant showed two minima of dF/dT in both conditions. With addition of 10 mM DTT the dominant T_m minima shifted from the minima at a higher temperature (70 °C) to the lower temperature (58 °C). In contrast, Herceptin-SeCys only had one minimum in both concentrations of DTT. Again, the SeCys variant showed stabilization under reducing conditions, with the T_m decreasing from 68 C without DTT to 64 C in 10 mM DTT.

DETAILED DESCRIPTION OF THE INVENTION

[35] A sequence listing forms part of the disclosure of this application and is incorporated as part of the disclosure.

[36] The inventors have developed polypeptides and methods to produce polypeptides in genomically recoded organisms (GRO) that fold into biologics that are stabilized by diselenide bonds between selenocysteine amino acids. Whereas disulfide bonds between cysteine amino acids have a redox potential of about -220 mV, diselenide bonds have a redox potential of about -380 mV. Since the bacterial cytosol typically has a redox potential of about -280 to -300 mV, diselenides but not disulfides avoid reduction so that they form and persist in the cytosol. Since diselenides have the same geometric bond angles and torsions as disulfides, as well as very similar bond lengths, they can be substituted into polypeptides without disrupting the three dimensional structure of the polypeptide (Figure 1). Further, since intended *in vivo* environments like blood contain reducing agents like glutathione, albumin, and thioredoxin,

disulfides in polypeptides can be reduced, causing the polypeptide to unfold and, in the case of multiple disulfides, “scramble” the disulfides so that incorrect cysteines are bonded to each other. Both of these result in abrogation of the intended biological activity of the polypeptide. The lower redox potential of diselenides renders them resistant to reduction when exposed to blood serum or purified reducing components of blood serum (Figure 2), endowing them with a longer blood serum half-life than disulfide-bearing counterparts (Muttenhaler et al.).

[37] While peptides bearing diselenide-forming selenocysteines may be produced *in vitro* by solid phase peptide synthesis (Armishaw et al., Safavi-Hemami et al.), the process does not scale tractably to the yields necessary for therapeutic applications, particularly for proteins. However, *in vivo* production of recombinant seleno-proteins is limited by strict sequence requirements on where selenocysteine may appear in proteins. In particular, a selenocysteine insertion sequence (SECIS) element must appear in the coding DNA sequence at the selenocysteine incorporation site in order to recruit endogenous selenocysteine translation machinery, composed of specialized enzymes (selA, selD), tRNA (selC), and an elongation factor (selB). Instead, we use a recoded strain of *E. coli*, which has an unassigned codon, such as an amber stop codon, together with an engineered selenocysteine tRNA with an anti-amber anticodon that permits targeted placement of selenocysteine into polypeptides by introduction of the amber stop codon into the corresponding DNA coding sequence. Typically, recoding entails removing all naturally occurring instances of the unassigned codon from the genome, as well as removing or otherwise deactivating the translational machinery associated with the unassigned codon (*e.g.*, for amber, removing or deactivating Release Factor 1). The modified tRNA interacts with the endogenous elongation factor EF-Tu. Other codons can be recoded, typically rare codons, as is known in the art. See, Ostrov et al., *Science*, 353:819-822, 2016. A codon on an mRNA and an anti-codon on a tRNA are typically triplets of complementary base sequences.

[38] Recoded proteins may be synthesized in bacteria, such as *E. coli* cells, or *in vitro*, in translation or linked transcription-translation systems. Genes or mRNA encoding

such recoded proteins are non-naturally occurring, and are variants of naturally occurring coding sequences.

[39] Although many of the proteins that we show in the associated sequence listing have all cysteine residues which participate in disulfide bonds replaced with selenocysteine residues, all cysteine residues need not be replaced to gain the benefits of the substitution. Even one diselenide bond may improve the stability of a protein. Any number of diselenide bonds (selenocysteine pairs) may be substituted for disulfide bonds in the proteins. If a protein has N disulfide bonds, the protein may have anywhere from N, N minus 1, N minus 2, N minus 3, N minus 4,down to 1 such bond. Proteins with different combinations of diselenide bonds and disulfides are shown in SEQ ID NO: 1220 to 1631. It is also possible to form a bond between cysteine and selenocysteine residues called a selenylsulfide. This bond has a lower redox potential (~270 mv) than a disulfide (-220 mv) but not than bacterial cytoplasm (-280 mv). The selenylsulfide bond may be used to increase resistance to reduction in certain redox environments. Selenylsulfides may be used in place of diselenides using methods described here by substituting selenocysteine for a single disulfide bonded cysteine, or by substituting cysteine for a single diselenide bonded selenocysteine.

[40] Sequences of disulfide-stabilized biologics with substituted selenocysteines can be produced in the cytosol of *E. coli* using our method at the mg/L scale in standard laboratory shaker flasks, and scaled to g/L production in microbial fermenters. These sequences include therapeutic proteins, vaccines, and monoclonal antibodies, antibody antibody fragments, antibody conjugates, and other immunoglobulin-like molecules. The Fv domains may be combined with any human Fc domains. Further, any Fv domain may be produced in our system by substituting selenocysteine for cysteine at positions 23 and 88 using Kabat numbering in the variable light chain and/or by substituting selenocysteine for cysteine at positions 22 and 92 using Kabat numbering in the variable heavy chain. Immunoglobulin classes IgG, IgA, IgM, IgD and IgE (including IgA1, IgA2, IgG1, IgG2, IgG3, and IgG4 subclasses) may be produced by these methods as typified classes, with any constant variant therein combined with either the Ig kappa or lambda constant light chain variants known to those in the art. Any selenocysteine in the disclosed sequences may be maintained as a cysteine so

long as the presence of the cysteine does not interfere with the expression, folding, or intended function of the polypeptide. We disclose methods below for producing and verifying the presence of selenocysteines participating in the intended diselenide bonds for antibody single chain variable fragments (scFv, Figures 3, 6), human growth hormone (hGH, Figure 6), PD-L1 (Figure 4), variants of human Fc (Figure 4), and fusions of PD-L1 and hGH to variants of human Fc (Figures 4 and 5), but these methods also apply to full monoclonal antibodies, other antibody fragments, antibody conjugates, other therapeutic proteins, and vaccines.

[41] Therapeutic, diagnostic, biosensing, and prophylactic biologics may be made and used according to the invention with diselenide bonds between two selenocysteine residues. This technique and modification may be useful for any protein biologic, including but not limited to hormones (*e.g.*, human growth hormone, Figure 8), antibodies and antibody fragments (Figures 7 and 8), cytokines, growth factors, structural proteins, and the like. Insulin products, such as Lantus™, Novorapid™, Humalog™, Humulin™, Levemir™, Apidra™, and Tresiba™ may be useful with such modifications, as shown in SEQ ID NO: 1142-1155. The A chains of these insulin products can be modified as shown in SEQ ID NO: 1142, 1144, 1146, 1148, 1150, 1152, and 1154. The B chains of these insulin products can be modified as shown in SEQ ID NO: 1143, 1145, 1147, 1149, 1151, 1153, and 1155. The diselenide bonds may be intramolecular or intermolecular. Additional proteins that are suitable for such modifications include Programmed death-ligand 1 (PD-L1), PD-L1 fusion protein to an antibody domain(s), portions of PD-L1, such as the binding domain whether fused or unfused to an antibody domain(s), antibody domain(s), human growth hormone whether fused or unfused to an antibody domain(s), and Herceptin (Figures 4-8). This is not intended to be an exhaustive list of proteins that can be so modified.

[42] Although particular exemplary proteins are shown with a particular sequence in the sequence listing, which forms part of this disclosure, these sequences need not be used used and reproduced with total fidelity. As is known in the art, slight modifications to proteins including allelic variations and polymorphisms may occur in parts of proteins that are not detrimental to their use and function. Thus useful proteins according to

the present invention may have the precise sequences which are shown in the sequence listing or they may be slightly different. For example, useful proteins may be at least 99 percent, at least 98 percent, at least 97 percent, at least 96 percent, at least 95 percent, at least 94 percent, at least 93 percent, at least 92 percent, at least 91 percent, at least 90 percent, at least 89 percent, at least 88 percent, at least 87 percent, at least 86 percent, at least 85 percent, at least 84 percent, at least 83 percent, at least 82 percent, 81 percent, or at least 80 percent identical.

[43] The proteins that can be made according to the invention are unlimited in purpose. They can be therapeutic proteins, prophylactic proteins (such as vaccines), enzymatic proteins, endocrine proteins, signaling proteins, scaffolding proteins, diagnostic proteins, analytic proteins, etc. The diselenide bonds will generally make the proteins more stable in the presence of reducing agents.

[44] The proteins can be made in a host cell, or *in vitro*, in cell-free synthetic systems. Host cells may be any that can be robustly recoded. These can be bacterial cells that have well developed genetic systems, of which *E. coli* is exemplary. Other bacterial species can also be used. Cell-free systems for producing the proteins may be coupled transcription/translation systems or only translation systems. A notable aspect of the methods of the invention is the use of biological syntheses rather than chemical synthesis means.

[45] Culturing of recoded cells with the constructed nucleic acid sequences may be by any means known in the art. The culturing may be batch or continuous, in shaker flasks or in fermenters or immobilized on solid surfaces, such as small particles contained in larger vessels. Typically the culture medium will be supplemented with a source of selenium, such as Na_2SeO_3 . As is known in the art, production of the desired protein variant may be under the control of an inducer or a repressor. Any such systems that are known in the art may be selected for convenience of construction and protein production.

[46] Administration of a therapeutic, diagnostic, or prophylactic protein to a subject in need of such treatment may be by any means known in the art and suitable for the protein. These include without limitation intravenous, intramuscular, subcutaneous,

intrathecal, oral, intracoronary, and intracranial deliveries. Certain proteins are appropriate for certain types of delivery, due to stability and target. Administration of a therapeutic protein can include one or more pharmaceutically acceptable carriers, such as, for example, a liquid or solid filler, diluent, excipient, buffer, stabilizer, or encapsulating material.

[47] The above disclosure generally describes the present invention. All references disclosed herein are expressly incorporated by reference. A more complete understanding can be obtained by reference to the following specific examples which are provided herein for purposes of illustration only, and are not intended to limit the scope of the invention.

EXAMPLE 1

scFv Expression and Purification

[48] RTΔA 2X310K cells transformed with plasmid pACYC-PcaU-aMS2 bearing a gene encoding an anti-alpha MS2 scFv antibody with UAG codons directing selenocysteine incorporation substituted in place of all cysteine codons were diluted 1/250 in 1 L of terrific broth supplemented with 1000 µg.mL carbenicillin, 12.5 µg.mL tetracycline and 25 µM Na₂SeO₃ and induced with (1 mM) 3,4-dihydroxybenzoic acid during mid log phase followed by growth O/N at 37 °C. Cells were harvested by centrifugation at 8000 x g for 10 min and resuspended in 25 mL of wash buffer (50 mM K₂HPO₄, 300 mM NaCl, 20 mM imidazole and 10% glycerol at pH 8.0) with protease inhibitor cocktail (cOmplete, mini EDTA free, Roche) and lysozyme (0.5 mg.mL⁻¹). Following a 20 min incubation at 4 °C with gentle agitation cells were lysed by sonication (Model 500, Fisher Scientific) and clarified three times by centrifugation at 35000 x g for 30 min. Lysate was filtered through a 0.2 µm membrane and protein was recovered by IMAC using Ni-NTA resin and gravity flow columns. Eluate was concentrated to 3 mL and dialyzed against TBS pH 7.5 followed by purification to apparent homogeneity by size exclusion FPLC. A significant proportion of the sample precipitated during dialysis and an additional proportion of the sample was present in the less stable/soluble dimer form. Final yield of soluble aMS2 seleno-scFv monomer was 3 mg/L.

[49] RTΔA 2X310K-T7 cells transformed with plasmid pACYC-TetO-aTNF bearing a gene encoding an anti-TNF (tumor necrosis factor) scFv antibody with UAG codons directing selenocysteine incorporation substituted in place of all cysteine codons were diluted 1/250 in 1 L of terrific broth supplemented with 1000 µg.mL carbenicillin, 33 µg.mL chloramphenicol and 25 µM Na₂SeO₃ and induced with 200 ng.mL anhydrotetracycline during mid log phase followed by growth O/N at 30 °C. Cell were harvested and purified by IMAC as previously described. To reduce precipitation, samples were dialysed in larger volumes (> 6 mL). Protein samples were purified to homogeneity by either size exclusion FPLC or anion exchange chromatography (HiTrap Q HP column). Anti-RCA, anti-GM-CSF, and anti-interferon gamma seleno-scFvs were purified by the same procedure. Final yield of soluble seleno-scFv monomers ranged between 0.5-1 mg/L.

EXAMPLE 2

15 Mass Spectrometry

[50] Intact protein samples were analyzed using methods described previously. Selenoprotein samples were buffer exchanged into LC-MS grade water using 10 kDa molecular weight cut-off filters. Once the buffer exchange was complete the samples were diluted to 20 µM in a methanol/water/formic acid (50/49/1) solution. After dilution, protein solutions were infused into an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific Instruments, Bremen, Germany) at a rate of 3 µL.min⁻¹ via electrospray ionization. In order to confirm the incorporation of selenocysteine, intact mass analysis was carried out at 240k resolution and averaging 20 scans. Characterization of the protein sequences was undertaken by ultraviolet photodissociation (UVPD) using a 193 nm excimer laser (Coherent, Inc.) which was interfaced to the Orbitrap mass spectrometer as described previously. For each UVPD spectrum, two laser pulses of 2.5 mJ were used and 250 scans were averaged. MS1 spectra were deconvoluted using the Xtract deconvolution algorithm (Thermo Fisher Scientific). UVPD mass spectra were also deconvoluted using Xtract and then analyzed using ProsightPC 3.0. Proteins containing selenocysteine were searched by adding a modification of 61.9146 Da to serine residues at the four incorporation sites (including the subtraction of one hydrogen atom due to formation of diselenide

bonds). Incorporation efficiencies were calculated by dividing the area of the modified protein peak by the summed areas of the unmodified protein peak and the modified protein peak. The peak area used for each protein was the sum of the integrated areas of the five most abundant peaks from each isotope cluster.

5

EXAMPLE 3

**UAG Genomically Recoded Organism Suitable for
Expressing Polypeptides with Non-Standard Amino Acids**

- 10 [51] A Genomically Recoded Organism (GRO) in which the UAG codon translational function is completely removed is used to unambiguously incorporate non-standard amino acids (NSAAs) at UAG. See Lajoie, M. J. et al. Genomically recoded organisms expand biological functions. *Science* 342, 357-360, doi:10.1126/science.1241459 (2013), incorporated by reference in its entirety. A
- 15 genomically recoded organism may include one or more reassigned triplet codons to facilitate the incorporation of non-standard amino acids (NSAAa), such as selenocysteine. Triplet codons can be reassigned to incorporate non-standard amino acids, such as selenocysteine, using methods known to those of skill in the art. See Lajoie, M. J. et al. Probing the limits of genetic recoding in essential genes. *Science*
- 20 342, 361-363, doi:10.1126/science.1241460 (2013), incorporated by reference in its entirety. See also Thyer, R. et al. Evolving tRNA^{Sec} for Efficient Canonical Incorporation of Selenocysteine. *JACS* 137(1), 46-49, (2015), incorporated by reference in its entirety. See also Ostrov, et al. Design, synthesis, and testing toward a 57-codon genome, *Science* 353, 819-822, 2016. Alternatively, quadruplet codons can
- 25 be used to incorporate non-standard amino acids, using methods known to those of skill in the art. See Anderson, J. C. et al. An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7566-7571, doi:10.1073/pnas.0401517101 (2004), Neumann, H., Wang, K., Davis, L., Garcia-Alai, M. & Chin, J. W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464, 441-444, (2010) and Chatterjee, A., Lajoie, M. J., Xiao, H., Church, G. M. & Schultz, P. G. A Bacterial Strain with a Unique Quadruplet Codon Specifying Non-native Amino Acids. *Chembiochem*, n/a-n/a, doi:10.1002/cbic.201402104 (2014) each of which are incorporated by reference
- 30

in their entirety. An orthogonal aminoacyl-tRNA synthetase (aaRS)/tRNA pair is developed that specifically and efficiently decodes the quadruplet UAGA codon based based on the non-functional UAG triplet resulting in unambiguous incorporation of non-standard amino acids, such as selenocysteine, at UAGA codons producing high protein yields. Such quadruplet codons may be used in the present methods.

[52] Over 100 NSAAs with diverse chemistries have been synthesized and co-translationally incorporated into proteins using evolved orthogonal aminoacyl-tRNA synthetase (aaRSs)/tRNA pairs. See Liu, C. C. & Schultz, P. G. Adding new chemistries to the genetic code. *An. Rev. Biochem.* 79, 413-444, doi:10.1146/annurev.biochem.052308.105824 (2010), incorporated by reference in its entirety. Non-standard amino acids have been designed based on tyrosine or pyrrolysine. An aaRS/tRNA may be provided on a plasmid or into the genome of the genomically recoded organism. An orthogonal aaRS/tRNA pair is used to bioorthogonally incorporate NSAAs into proteins. Vector-based over-expression systems may be used to outcompete natural codon function with its reassigned function. See Wang, L., Brock, A., Herberich, B. & Schultz, P. G. Expanding the genetic code of *Escherichia coli*. *Science* 292, 498-500 (2001), Young, T. S., Ahmad, I., Yin, J. A. & Schultz, P. G. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *Journal of Molecular Biology* 395, 361-374 (2009), and Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H. & Schultz, P. G. A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in *Escherichia coli*. *Biochemistry* 52, 1828-1837, doi:10.1021/bi4000244 (2013) each of which are incorporated by reference in their entirety. If one completely abolishes natural UAG translation function, far lower aaRS/tRNA function may be sufficient to achieve efficient NSAA incorporation. GRO-based NSAA incorporation can use either vector- and/or genome-based aaRS/tRNA pairs. Genome-based aaRS/tRNA pairs have been used to reduce the mis-incorporation of canonical amino acids in the absence of available NSAAs (Mandell and Lajoie et al., *Nature*. 2015 Feb 5;518(7537):55-60. doi: 10.1038/nature14121, which is incorporated by reference).

[53] Since the UAG codon function has been completely reassigned in the genomically recoded organism, NSAAs, such as selenocysteine, can be incorporated in the

genomically recoded organism without any phenotypic consequences. NSAA incorporation in the genomically recoded organism may involve supplementing the growth media with the non-standard amino acid, such as selenocysteine, and an inducer for the aaRS. Alternatively, the aaRS may be expressed constitutively. Alternatively, as in the present disclosure, the endogenous seryl-tRNA synthetase may be used to seryl-ate selenocysteine tRNA, which tRNA is acted upon by enzymes comprising selA and selD to produce tRNA^{sec} (selenocysteine charged tRNA). Media may be supplemented with a selenium source like sodium selenite to improve production of tRNA^{sec}. The desired protein can be overexpressed using any desired protein overexpression system (*e.g.*, T7-RNAP, constitutive incorporation, or inducible expression based on IPTG/allolactose, anhydrotetracycline, arabinose, rhamnose, or other inducible systems). The protein cross-link (diselenide bond) may form spontaneously based on proximity and orientation during protein folding, and the the protein can be handled as any other over-expressed product.

15 EXAMPLE 4

Methods

[54] Sequences coding for the genes of interest were cloned into a chloramphenicol resistant vector with a T7 promoter controlled by a Tet-repressor. Vectors were transformed into 2x310k cells with an integrated T7 RNA Polymerase controlled by a tet operon. Cells were grown to an OD ~0.6 and induced with aTc at 100 ng/mL at room temperature overnight. Induced cells were pelleted and resuspended in 20 mM Tris pH 7.4, 100 mM NaCl, 20 mM imidazole, and 0.01% Triton X-100 followed by lysis by sonication. The lysates were centrifuged to separate and pellet. Supernate was run over Ni-affinity beads in the same buffer as lysis. Proteins of interest were eluted in PBS pH7.4 with addition of 300mM imidazole. Purified fractions were run on a SDS page gel. The results for 10 different selenobiologics are shown in Figs. 4-6.

REFERENCES

The disclosure of each reference cited is expressly incorporated herein.

1. Anderson, J. C. *et al.* An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7566-7571, doi:10.1073/pnas.0401517101 (2004),
2. Armishaw, C.J., Daly, N.L., Nevin, S.T., Adams, D.J., Craik, D.J. & Alewood, P.F. Alpha-selenoconotoxins, a new class of potent alpha7 neuronal nicotinic receptor antagonists. *J Biol Chem* **281**, 14136-43 (2006).
3. Chatterjee, A., Lajoie, M. J., Xiao, H., Church, G. M. & Schultz, P. G. A Bacterial Strain with a Unique Quadruplet Codon Specifying Non-native Amino Acids. *Chembiochem*, n/a-n/a, doi:10.1002/cbic.201402104 (2014)
4. Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H. & Schultz, P. G. A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in *Escherichia coli*. *Biochemistry* 52, 1828-1837, doi:10.1021/bi4000244 (2013)
5. Lajoie, M. J. *et al.* Genomically recoded organisms expand biological functions. *Science* 342, 357-360, doi:10.1126/science.1241459 (2013)
6. Lajoie, M. J. *et al.* Probing the limits of genetic recoding in essential genes. *Science* 342, 361-363, doi:10.1126/science.1241460 (2013)
7. Liu, C. C. & Schultz, P. G. Adding new chemistries to the genetic code. *An. Rev. Biochem.* 79, 413-444, doi:10.1146/annurev.biochem.052308.105824 (2010)
8. Mandell and Lajoie *et al.*, *Nature*. 2015 Feb 5;518(7537):55-60. doi: 10.1038/nature14121,
9. Muttenthaler, M., Andersson, A., de Araujo, A.D., Dekan, Z., Lewis, R.J. & Alewood, P.F. Modulating oxytocin activity and plasma stability by disulfide bond engineering. *J Med Chem* **53**, 8585-96 (2010).
10. Neumann, H., Wang, K., Davis, L., Garcia-Alai, M. & Chin, J. W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464, 441-444, (2010)
11. Ostrov, *et al.* Design, synthesis, and testing toward a 57-codon genome, *Science* 353, 819-822, 2016
12. Safavi-Hemami, H., Gajewiak, J., Karanth, S., Robinson, S.D., Ueberheide, B., Douglass, A.D., Schlegel, A., Imperial, J.S., Watkins, M., Bandyopadhyay, P.K.,

- Yandell, M., Li, Q., Purcell, A.W., Norton, R.S., Ellgaard, L. & Olivera, B.M. Specialized insulin is used for chemical warfare by fish-hunting cone snails. *Proc Natl Acad Sci U S A* **112**, 1743-8 (2015).
13. Thyer, R. *et al.* Evolving tRNA^{Sec} for Efficient Canonical Incorporation of Selenocysteine. *JACS* 137(1), 46-49, (2015)
14. Wang, L., Brock, A., Herberich, B. & Schultz, P. G. Expanding the genetic code of *Escherichia coli*. *Science* 292, 498-500 (2001)
15. Young, T. S., Ahmad, I., Yin, J. A. & Schultz, P. G. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *Journal of Molecular Biology* 395, 361-374 (2009),

CLAIMS

1. A protein comprising at least two selenocysteine residues with at least one diselenide bond between the at least two selenocysteine residues, wherein the protein comprises a sequence selected from the group consisting of: SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence.
2. A protein according to claim 1 comprising at least two selenocysteine residues with at least one diselenide bond between the at least two selenocysteine residues, wherein the protein comprises a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence.
3. An antibody molecule comprising:

a heavy chain sequence selected from the group consisting of SEQ ID NO: 472-801 and sequences that are at least 80% identical to said sequence; and

a light chain sequence selected from the group consisting of SEQ ID NO: 812-1131 and sequences that are at least 80% identical to said sequence,

wherein from 0-15 pairs of selenocysteine residues in the antibody molecule are replaced with from 0-15 pairs of cysteine residues, and wherein the antibody contains at least one diselenide bond between the heavy chain and the light chain.
4. A modified insulin protein comprising:

an A chain having a sequence selected from SEQ ID NO: 1142, 1144, 1146, 1148, 1150, 1152, and 1154 and sequences that are at least 80% identical to said sequence;

a B chain having a sequence selected from SEQ ID NO: 1143, 1145, 1147, 1149, 1151, 1153, and 1155 and sequences that are at least 80% identical to said sequence;

wherein the protein comprises at least one diselenide bond between the A chain and the B chain.
5. A bacterial host that has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon, wherein the bacterial host contains a constructed DNA sequence encoding a variant form of a protein in which a pair of cysteine codons has been replaced with the unassigned codon wherein the variant form comprises a sequence selected from the group

consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said said sequence.

6. A bacterial host according to claim 5 that has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon, wherein the bacterial host contains a constructed DNA sequence encoding a variant form of a protein in which a pair of cysteine codons has been replaced with the unassigned codon wherein the variant form comprises a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence.
7. The bacterial host of claim 5 or 6 wherein the variant form is a monoclonal antibody.
8. The bacterial host of claim 7 wherein the monoclonal antibody comprises a heavy chain sequence selected from the group consisting of SEQ ID NO: 472-801 and sequences that are at least 80% identical to said sequence.
9. The bacterial host of claim 7 wherein the monoclonal antibody comprises a light chain sequence selected from the group consisting of SEQ ID NO: 812-1131 and sequences that are at least 80% identical to said sequence.
10. The bacterial host of claim 7 wherein the monoclonal antibody comprises a heavy chain variable region sequence selected from the group consisting of SEQ ID NO: 16-243 and sequences that are at least 80% identical to said sequence.
11. The bacterial host of claim 7 wherein the monoclonal antibody comprises a light chain variable region sequence selected from the group consisting of SEQ ID NO: 244-471 and sequences that are at least 80% identical to said sequence.
12. The bacterial host of claim 10 or 11 wherein the monoclonal antibody further comprises a constant chain sequence selected from the group consisting of SEQ ID NO: 1-15 and sequences that are at least 80% identical to said sequence.
13. The bacterial host of claim 5 or 6 wherein the variant form is a therapeutic protein.
14. The bacterial host of claim 5 or 6 wherein the variant form is a prophylactic protein.
15. The bacterial host of claim 5 or 6 wherein the unassigned codon is a UAG.
16. A bacterial host that has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon, wherein the bacterial host contains a constructed DNA sequence encoding a variant form of a

monoclonal antibody in which a pair of cysteine codons has been replaced with the unassigned codon, wherein the monoclonal antibody comprises a VL chain and positions 23 and 88 of the VL chain are substituted with selenocysteine residues, and the monoclonal antibody comprises a VH chain and positions 22 and 92 of the VH chain are substituted with selenocysteine residues.

17. The bacterial host of claim 16 wherein the unassigned codon is a UAG.
18. A method of producing a variant form of a protein, wherein the protein has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues, comprising:

culturing a bacterial host that has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon, wherein the bacterial host contains a constructed DNA sequence encoding the variant form of the protein in which a pair of cysteine codons has been replaced with the unassigned codon, wherein the culturing is under conditions in which the variant form of the protein is expressed, wherein the variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence.
19. A method of producing a variant form of a protein according to claim 18, wherein the protein has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues, comprising:

culturing a bacterial host that has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon, wherein the bacterial host contains a constructed DNA sequence encoding the variant form of the protein in which a pair of cysteine codons has been replaced with the unassigned codon, wherein the culturing is under conditions in which the variant form of the protein is expressed, wherein the variant form comprises a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence.
20. The method of claim 18 or 19 further comprising the step of harvesting the variant form of the protein from culture medium or from cells of the bacterial host.
21. The method of claim 18 or 19 wherein the variant form is a monoclonal antibody.

22. The method of claim 21 wherein the monoclonal antibody comprises a heavy chain sequence selected from the group consisting of SEQ ID NO: 472-801 and sequences that are at least 80% identical to said sequence.
23. The method of claim 21 wherein the monoclonal antibody comprises a light chain sequence selected from the group consisting of SEQ ID NO: 812-1131 and sequences that are at least 80% identical to said sequence.
24. The method of claim 21 wherein the monoclonal antibody comprises a heavy chain variable region sequence selected from the group consisting of SEQ ID NO: 16-243 and sequences that are at least 80% identical to said sequence.
25. The method of claim 21 wherein the monoclonal antibody comprises a light chain variable region sequence selected from the group consisting of SEQ ID NO: 244-471 and sequences that are at least 80% identical to said sequence.
26. The method of claim 24 or 25 wherein the monoclonal antibody further comprises a constant chain sequence selected from the group consisting of SEQ ID NO: 1-15 and sequences that are at least 80% identical to said sequence.
27. The method of claim 18 or 19 wherein the variant form is a therapeutic protein.
28. The method of claim 18 or 19 wherein the variant form is a prophylactic protein.
29. The method of claim 18 or 19 wherein the unassigned codon is UAG.
30. A method of producing a variant form of a monoclonal antibody, wherein the monoclonal antibody has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues, comprising:
culturing a bacterial host that has been genomically recoded so that an unassigned codon is recognized by a selenocysteine tRNA with a corresponding anticodon, wherein the bacterial host contains a constructed DNA sequence encoding the variant form of the monoclonal antibody in which a pair of cysteine codons has been replaced with the unassigned codon, wherein the culturing is under conditions in which the variant form of the monoclonal antibody is expressed, wherein the monoclonal antibody comprises a VL chain and positions 23 and 88 of the VL chain are substituted with selenocysteine residues, and the monoclonal antibody comprises a

VH chain and positions 22 and 92 of the VH chain are substituted with selenocysteine residues.

31. The method of claim 30 wherein the unassigned codon is a UAG.
32. A method of treating a mammal with a variant form of a protein, comprising the step of:

administering to the mammal the variant form of the protein, wherein the protein has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues, wherein the variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence.
33. A method of treating a mammal according to claim 32 with a variant form of a protein, comprising the step of:

administering to the mammal the variant form of the protein, wherein the protein has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues, wherein the variant form comprises a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence.
34. The method of claim 32 or 33 wherein the administering is intravenous.
35. The method of claim 32 or 33 wherein the protein is a monoclonal antibody.
36. The method of claim 32 or 33 wherein the protein is a therapeutic protein.
37. The method of claim 32 or 33 wherein the protein is a prophylactic protein.
38. A method of treating a mammal with a variant form of a monoclonal antibody, comprising the step of:

administering to the mammal the variant form of the monoclonal antibody, wherein the monoclonal antibody has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues, wherein the monoclonal antibody comprises a VL chain and positions 23 and 88 of the VL chain are substituted with

selenocysteine residues, and the monoclonal antibody comprises a VH chain and positions 22 and 92 of the VH chain are substituted with selenocysteine residues.

39. The antibody of claim 3 wherein at least one pair of selenocysteine residues in the antibody molecule is replaced with one pair of cysteine residues.
40. The antibody molecule of claim 3 wherein one pair of selenocysteine residues in the antibody molecule is replaced with one pair of cysteine residues.
41. The antibody molecule of claim 3 wherein two pairs of selenocysteine residues in the antibody molecule are replaced with two pairs of cysteine residues.
42. The antibody molecule of claim 3 wherein three pairs of selenocysteine residues in the antibody molecule are replaced with three pairs of cysteine residues.
43. The antibody molecule of claim 3 wherein four pairs of selenocysteine residues in the antibody molecule are replaced with four pairs of cysteine residues.
44. The antibody molecule of claim 3 wherein at five pairs of selenocysteine residues in the antibody molecule are replaced with five pairs of cysteine residues.
45. The antibody molecule of claim 3 wherein six pairs of selenocysteine residues in the antibody molecule are replaced with six pairs of cysteine residues.
46. The antibody molecule of claim 3 wherein seven pairs of selenocysteine residues in the antibody molecule are replaced with seven pairs of cysteine residues.
47. The antibody molecule of claim 3 wherein eight pairs of selenocysteine residues in the antibody molecule are replaced with eight pairs of cysteine residues.
48. The antibody molecule of claim 3 wherein nine pairs of selenocysteine residues in the antibody molecule are replaced with nine pairs of cysteine residues.
49. The antibody molecule of claim 3 wherein ten pairs of selenocysteine residues in the antibody molecule are replaced with ten pairs of cysteine residues.
50. The antibody molecule of claim 3 wherein eleven pairs of selenocysteine residues in the antibody molecule are replaced with eleven pairs of cysteine residues.
51. The antibody molecule of claim 3 wherein twelve pairs of selenocysteine residues in the antibody molecule are replaced with twelve pairs of cysteine residues.

52. The antibody molecule of claim 3 wherein thirteen pairs of selenocysteine residues in the antibody molecule are replaced with thirteen pairs of cysteine residues.
53. The antibody molecule of claim 3 wherein fourteen pairs of selenocysteine residues in the antibody molecule are replaced with fourteen pairs of cysteine residues.
54. The antibody molecule of claim 3 wherein fifteen pairs of selenocysteine residues in the antibody molecule are replaced with fifteen pairs of cysteine residues.
55. An *in vitro* translation system or *in vitro* transcription and translation system comprising a constructed nucleic acid molecule and a tRNA molecule, wherein the constructed nucleic acid molecule encodes a variant form of a protein in which a pair of cysteine residues has been replaced with a pair of selenocysteine residues, wherein the variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1631 and sequences that are at least 80% identical to said sequence, wherein the nucleic acid molecule comprises an unassigned codon or leverages cognate codon suppression to incorporate selenocysteine residues, wherein the tRNA molecule is a selenocysteine tRNA with an anticodon that corresponds to the unassigned codon.
56. The *in vitro* translation system or *in vitro* transcription and translation system of claim 55 wherein the variant form is a monoclonal antibody.
57. The *in vitro* translation system or *in vitro* transcription and translation system of claim 55 wherein the monoclonal antibody comprises a heavy chain sequence selected from the group consisting of SEQ ID NO: 472-801 and sequences that are at least 80% identical to said sequence.
58. The *in vitro* translation system or *in vitro* transcription and translation system of claim 55 wherein the monoclonal antibody comprises a light chain sequence selected from the group consisting of SEQ ID NO: 812-1131 and sequences that are at least 80% identical to said sequence.
59. The *in vitro* translation system or *in vitro* transcription and translation system of claim 55 wherein the monoclonal antibody comprises a heavy chain variable region sequence selected from the group consisting of SEQ ID NO: 16-243 and sequences that are at least 80% identical to said sequence.
60. The *in vitro* translation system or *in vitro* transcription and translation system of claim 55 wherein the monoclonal antibody comprises a light chain variable region sequence

selected from the group consisting of SEQ ID NO: 244-471 and sequences that are at least 80% identical to said sequence.

61. The *in vitro* translation system or *in vitro* transcription and translation system of claim 55 wherein the monoclonal antibody further comprises a constant chain sequence selected from the group consisting of SEQ ID NO: 1-15 and sequences that are at least 80% identical to said sequence.
62. The *in vitro* translation system or *in vitro* transcription and translation system of claim 55 wherein the variant form is a therapeutic protein.
63. The *in vitro* translation system or *in vitro* transcription and translation system of claim 55 wherein the variant form is a prophylactic protein.
64. The *in vitro* translation system or *in vitro* transcription and translation system of claim 55 wherein the unassigned codon is a UAG.
65. An *in vitro* translation system or *in vitro* transcription and translation system comprising a constructed nucleic acid molecule and a tRNA molecule, wherein the constructed nucleic acid molecule encodes a variant form of a monoclonal antibody which comprises a VL chain and positions 23 and 88 of the VL chain are substituted with selenocysteine residues and the monoclonal antibody comprises a VH chain and positions 22 and 92 of the VH chain are substituted with selenocysteine residues, wherein the constructed nucleic acid molecule comprises at least a pair of cysteine codons that has been replaced with an unassigned codon in the nucleic acid, wherein the tRNA molecule is a selenocysteine tRNA with an anticodon that is complementary to the unassigned codon.
66. The *in vitro* translation system or *in vitro* transcription and translation system of claim 65 wherein the unassigned codon is a UAG.
67. A method of producing a variant form of a protein, wherein the protein has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues, comprising:

translating *in vitro* a constructed nucleic acid sequence encoding the variant form of the protein, wherein in said constructed nucleic acid sequence a pair of cysteine codons has been replaced with an unassigned codon, wherein the variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1630 and

sequences that are at least 80% identical to said sequence, wherein said step of translating employs a selenocysteine tRNA with a corresponding anticodon to the unassigned codon.

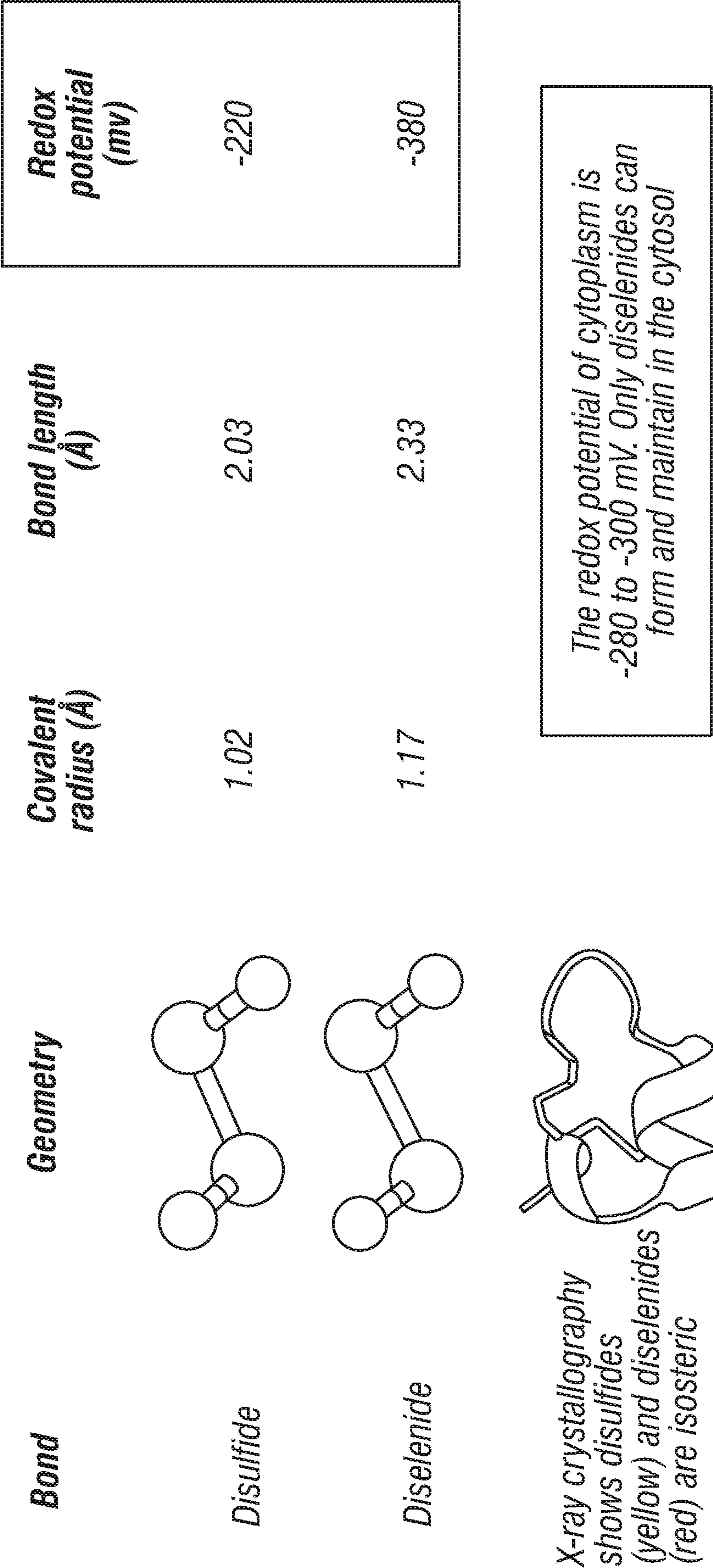
68. A method according to claim 67 of producing a variant form of a protein, wherein the protein has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues, comprising:

translating *in vitro* a constructed nucleic acid sequence encoding the variant form of the protein, wherein in said constructed nucleic acid sequence a pair of cysteine codons has been replaced with an unassigned codon, wherein the variant form comprises a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence, wherein said step of translating employs a selenocysteine tRNA with a corresponding anticodon to the unassigned codon.

69. The method of claim 67 or 68 wherein the variant form is a monoclonal antibody.
70. The method of claim 69 wherein the monoclonal antibody comprises a heavy chain sequence selected from the group consisting of SEQ ID NO: 472-801 and sequences that are at least 80% identical to said sequence.
71. The method of claim 69 wherein the monoclonal antibody comprises a light chain sequence selected from the group consisting of SEQ ID NO: 812-1131 and sequences that are at least 80% identical to said sequence.
72. The method of claim 69 wherein the monoclonal antibody comprises a heavy chain variable region sequence selected from the group consisting of SEQ ID NO: 16-243 and sequences that are at least 80% identical to said sequence.
73. The method of claim 69 wherein the monoclonal antibody comprises a light chain variable region sequence selected from the group consisting of SEQ ID NO: 244-471 and sequences that are at least 80% identical to said sequence.
74. The method of claim 72 or 73 wherein the monoclonal antibody further comprises a constant chain sequence selected from the group consisting of SEQ ID NO: 1-15 and sequences that are at least 80% identical to said sequence.
75. The method of claim 67 or 68 wherein the variant form is a therapeutic protein.
76. The method of claim 67 or 68 wherein the variant form is a prophylactic protein.

77. The method of claim 67 or 68 wherein the unassigned codon is UAG.
78. A method of testing a sample, comprising the step of:
contacting a sample with a variant form of a protein to form a reaction mixture, wherein the protein has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues, wherein the variant form comprises a sequence selected from the group consisting of SEQ ID NO: 1-1630 and sequences that are at least 80% identical to said sequence; and
testing the reaction mixture to detect a specific binding partner bound to the variant form of the protein.
79. A method according to claim 78 of testing a sample, comprising the step of:
contacting a sample with a variant form of a protein to form a reaction mixture, wherein the protein has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues, wherein the variant form comprises a sequence of SEQ ID NO: 1631 and sequences that are at least 80% identical to said sequence; and
testing the reaction mixture to detect a specific binding partner bound to the variant form of the protein.
80. A method testing a sample, comprising the step of:
contacting a sample with a variant form of a monoclonal antibody to form a reaction mixture, wherein the monoclonal antibody has a disulfide bond between cysteine residues, wherein the variant form has a diselenide bond between selenocysteine residues in place of the disulfide bond between cysteine residues, wherein the monoclonal antibody comprises a VL chain and positions 23 and 88 of the VL chain are substituted with selenocysteine residues, and the monoclonal antibody comprises a VH chain and positions 22 and 92 of the VH chain are substituted with selenocysteine residues; and
testing the reaction mixture to detect a specific binding partner bound to the variant form of the protein.

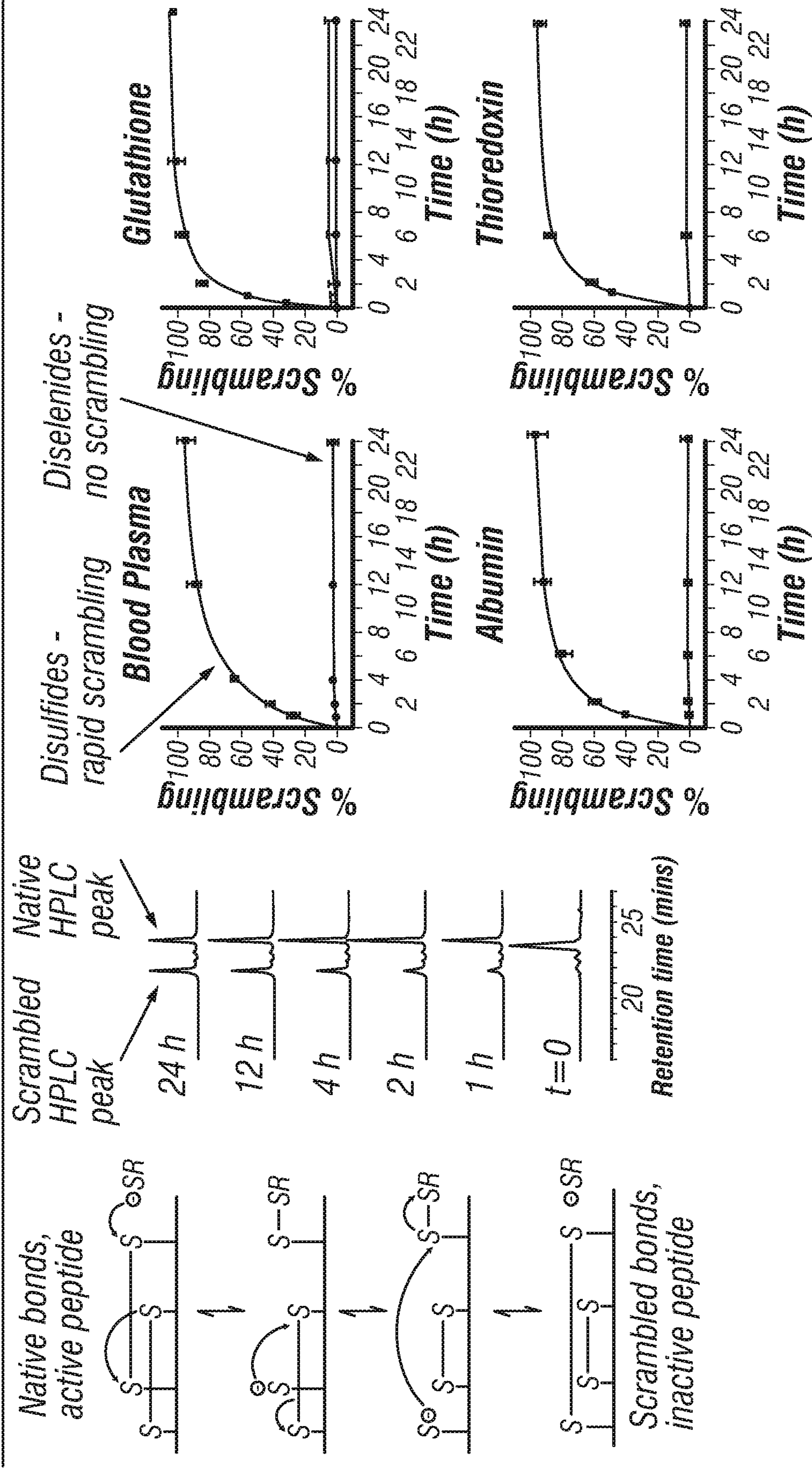
81. A modified antibody comprising at least one diselenide bond
82. The modified antibody of claim 81 wherein the diselenide bond is in the location of a disulfide bond in an occurring antibody of the same specificity.
83. The modified antibody of claim 81 wherein the diselenide bond is intermolecular.
84. The modified antibody of claim 81 which is selected from an isotype selected from the group consisting of IgA, IgD, IgE, IgG, and IgM.
85. A modified insulin protein comprising: an A chain having residues 90-110 of SEQ ID NO: 1630 with at least one cysteine to selenocysteine substitution; a B chain having residues 25-54 of SEQ ID NO: 1630 with at least one cysteine to selenocysteine substitution; and at least one diselenide bridge between the A chain and the B chain.
86. A modified human insulin protein comprising at least two selenocysteine residues replacing two cysteine residues in a native human insulin protein, with at least one diselenide bond between the at least two selenocysteine residues.
87. The modified human insulin protein of claim 86 which is produced in a recombinant bacterium.
88. The modified human insulin protein of claim 87 wherein the recombinant bacterium is recoded so that a selenocysteine tRNA recognizes an amber codon in the bacterium.



Muttenthaler et al, JACS, 2010

FIG. 1

Diselenide biologics show dramatic improvements to serum half-life over disulfides



- Reducing agents in blood may disrupt and scramble stabilizing disulfides
- Replacement of disulfides with diselenides substantially reduces scrambling

FIG. 2

Armishaw et al., JBC, 2006

3/9

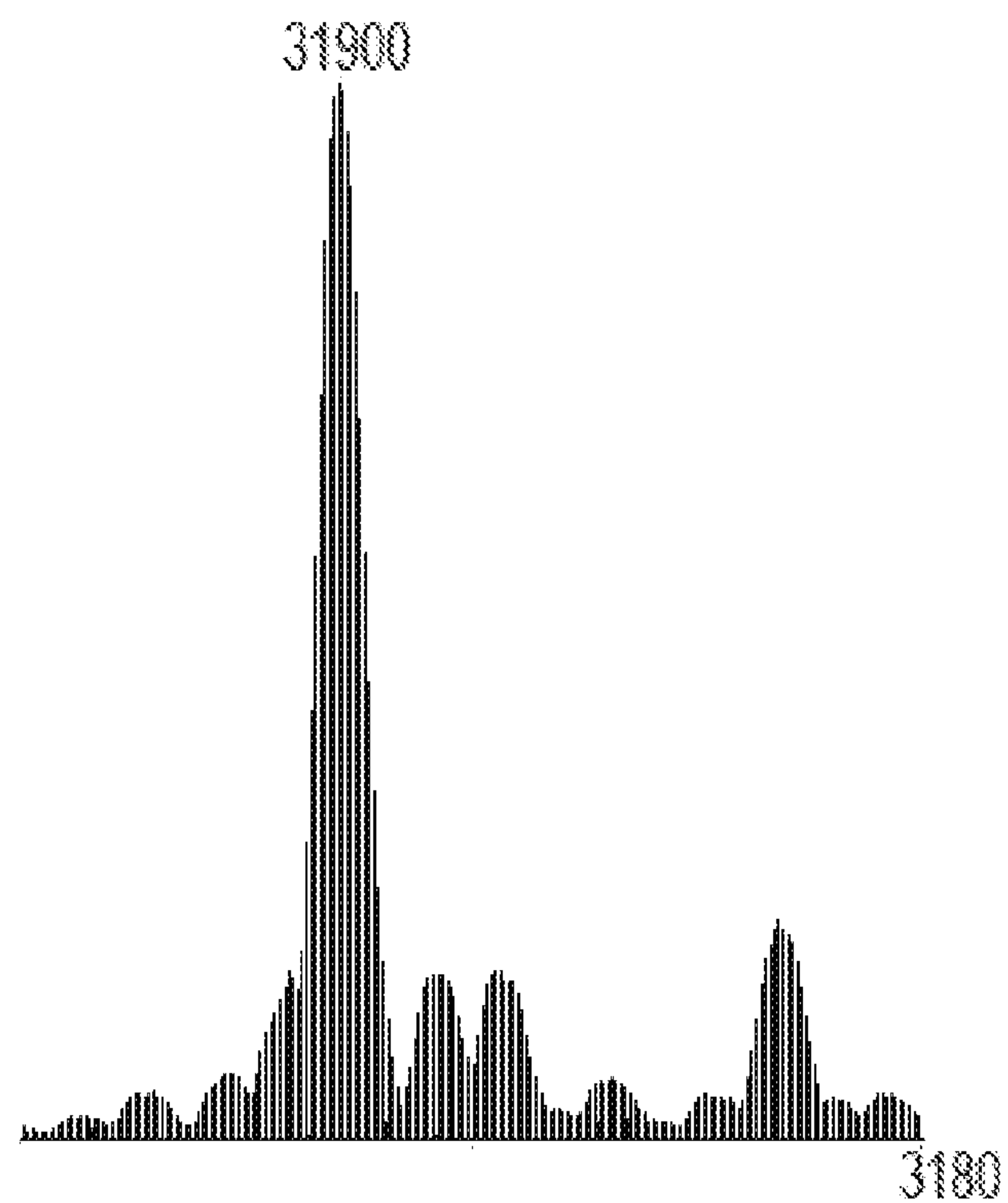


FIG. 3A

N M H H H H H L V P R G S G K S A M A A E V K L V 25
26 E S G G L V K P G G S L K L S U A A S G F T F S 50
51 S Y A M S W V R Q T P E K R L E W V A T I S T G G 75
76 G Y T Y F P D S V K G R F T I S R D N A K N A L Y 100
101 L Q M K S L R S E D T A M Y Y U A R Q G D F G D W 125
126 Y F D V W G A G T T V T V S S G G G S G G G G S 150
151 G G G S T D V L M T Q T P L S L P V S L G D Q A 175
176 S I S U R S S Q S L V H S N G N T Y L H W Y L Q K 200
201 P G Q S P K L L I Y K V S N R F S G V P D R F S G 225
226 S G S G T D F T L K I S R V E A E D L G V Y F U S 250
251 Q S T H V P W T F G G G T K L E I K R A A A D Y K 275
276 D H D G D Y K D H D I D Y K D D D K C

FIG. 3B

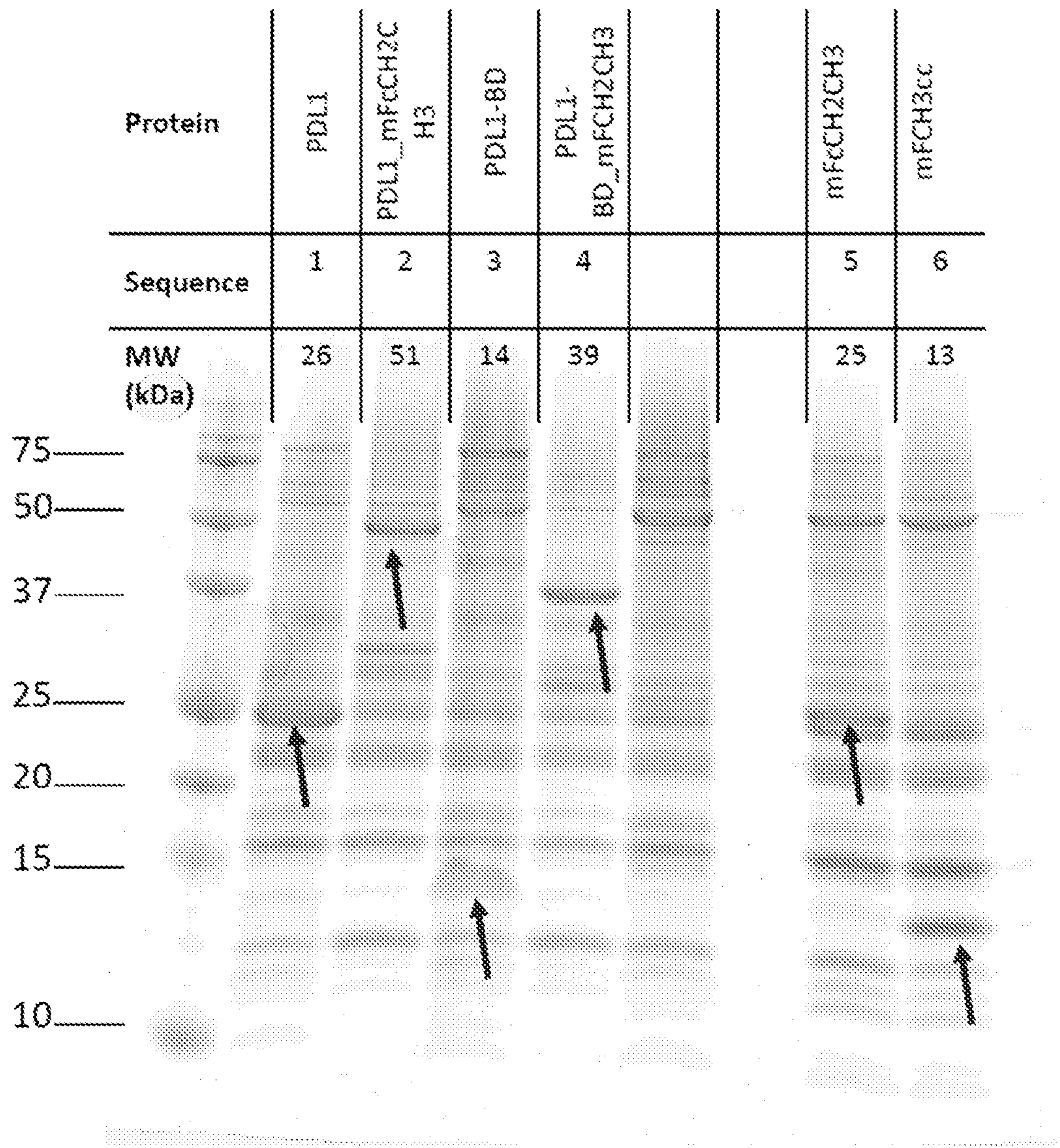


FIG. 4

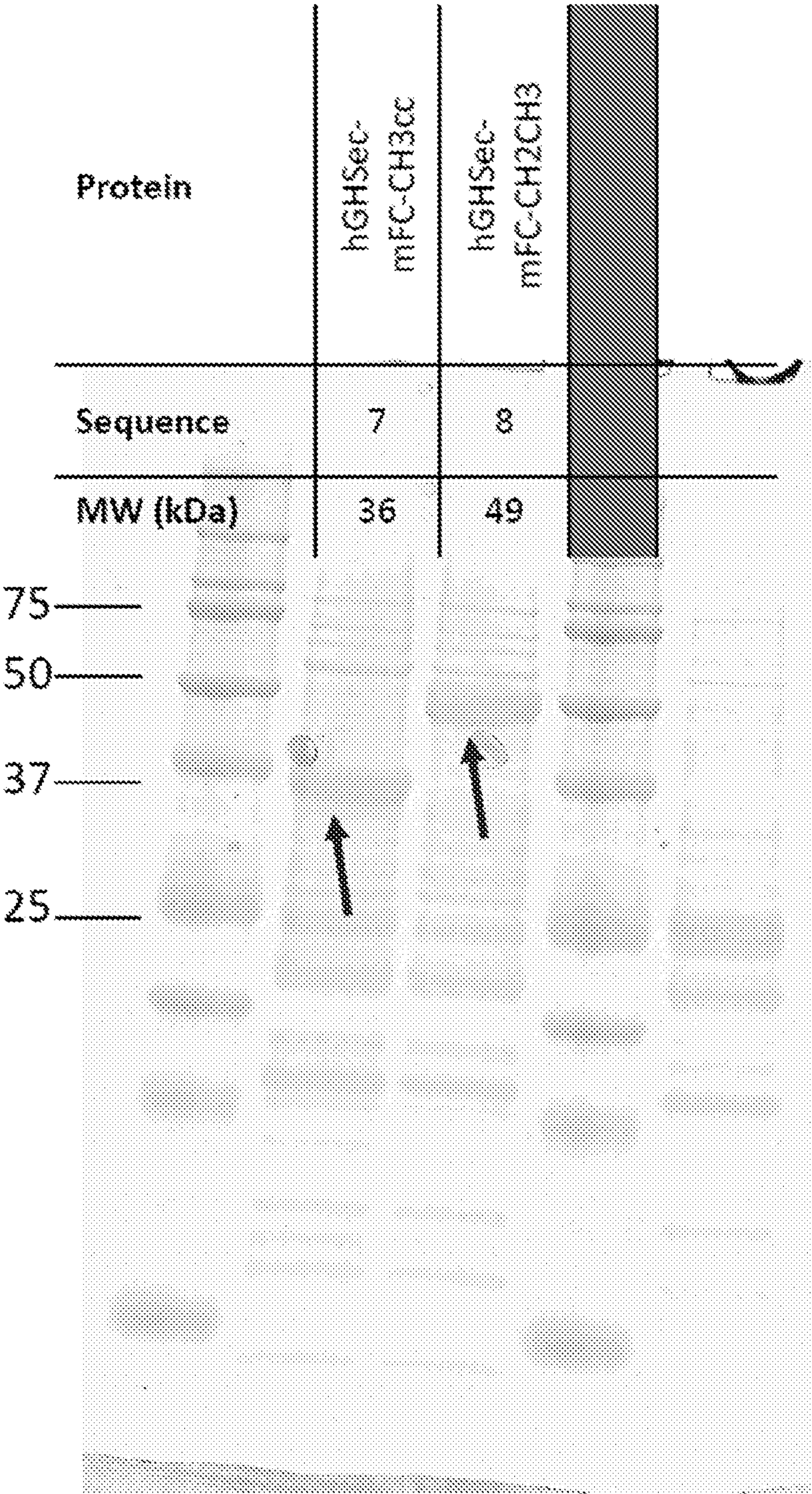


FIG. 5

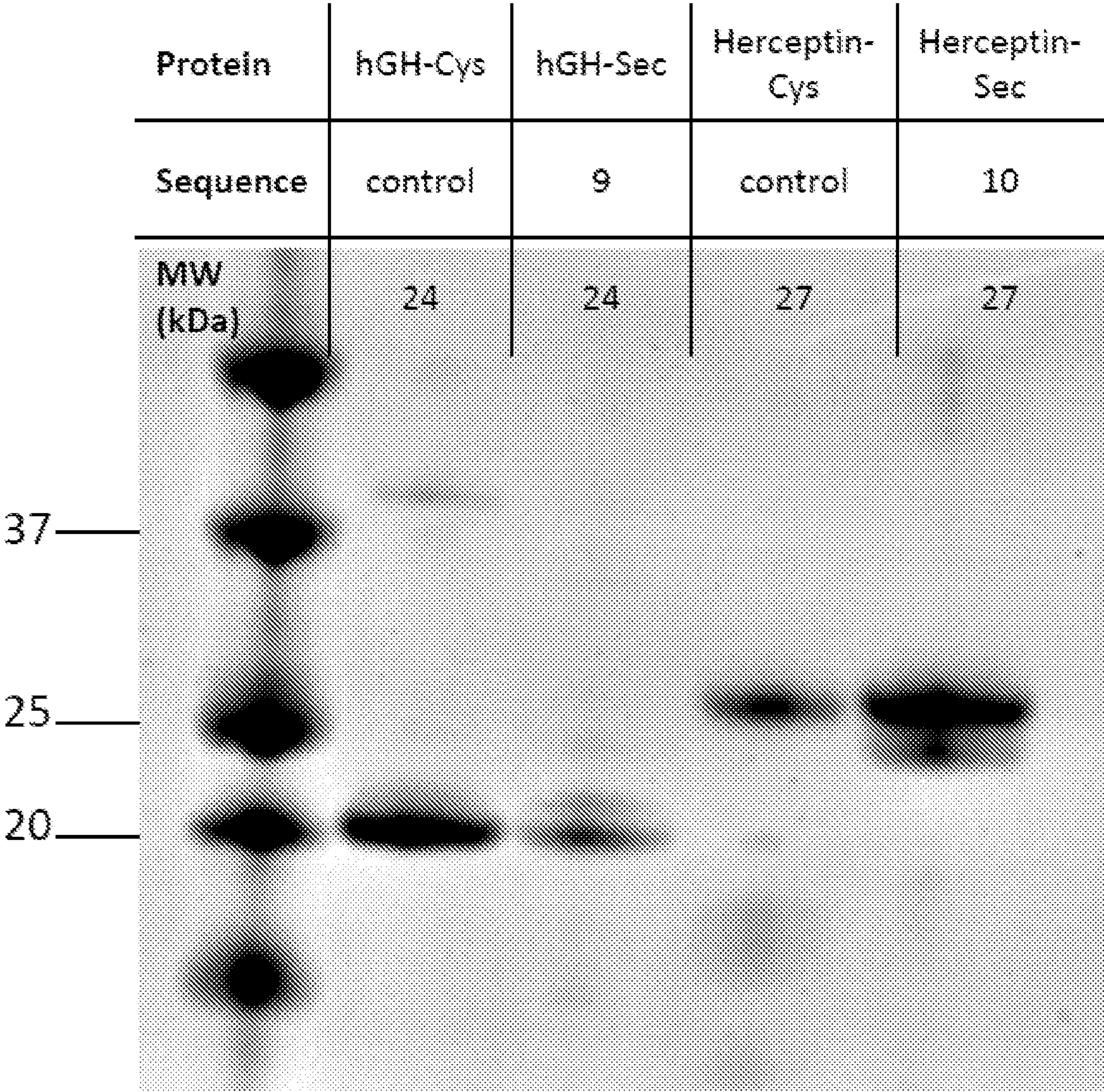


FIG. 6

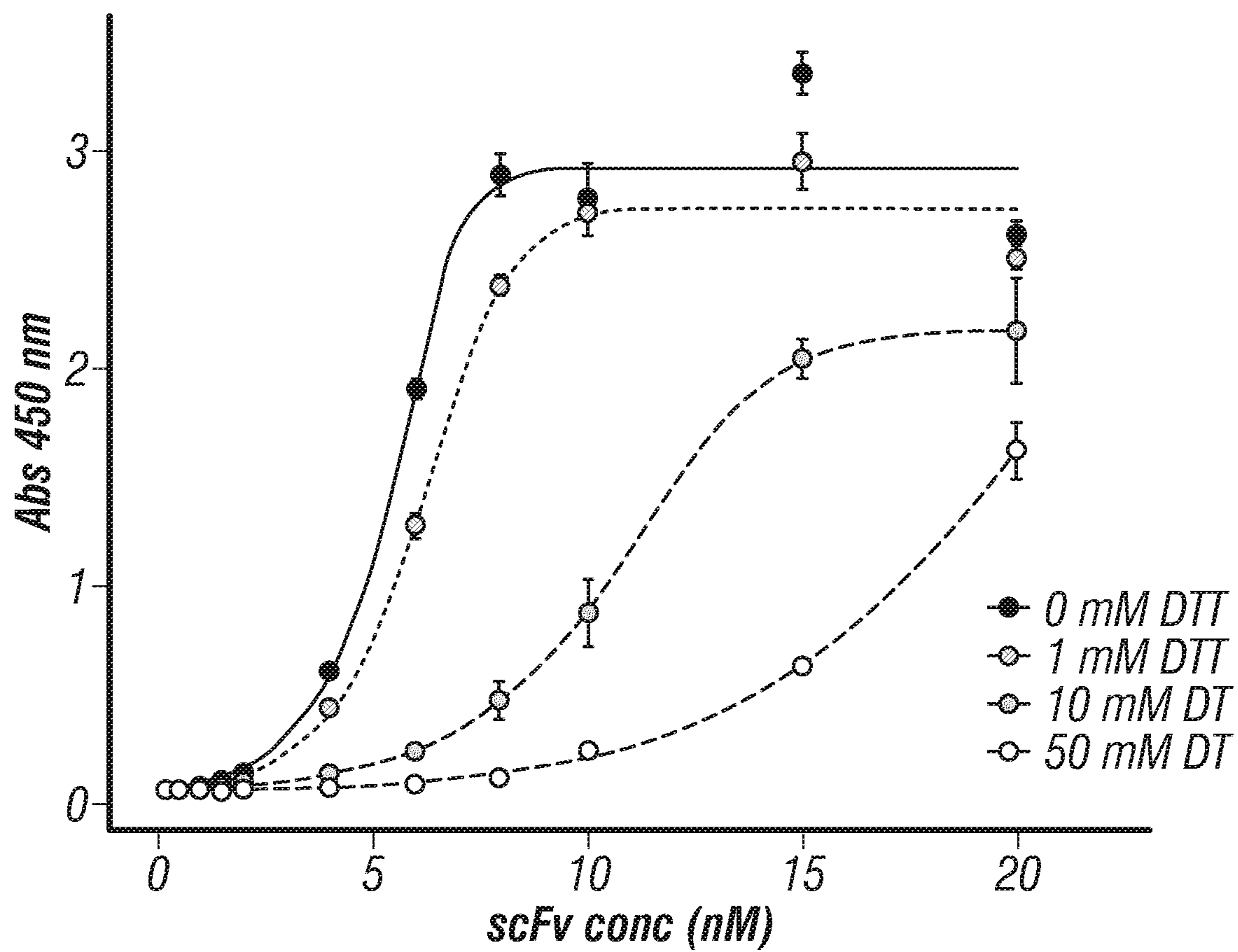


FIG. 7A

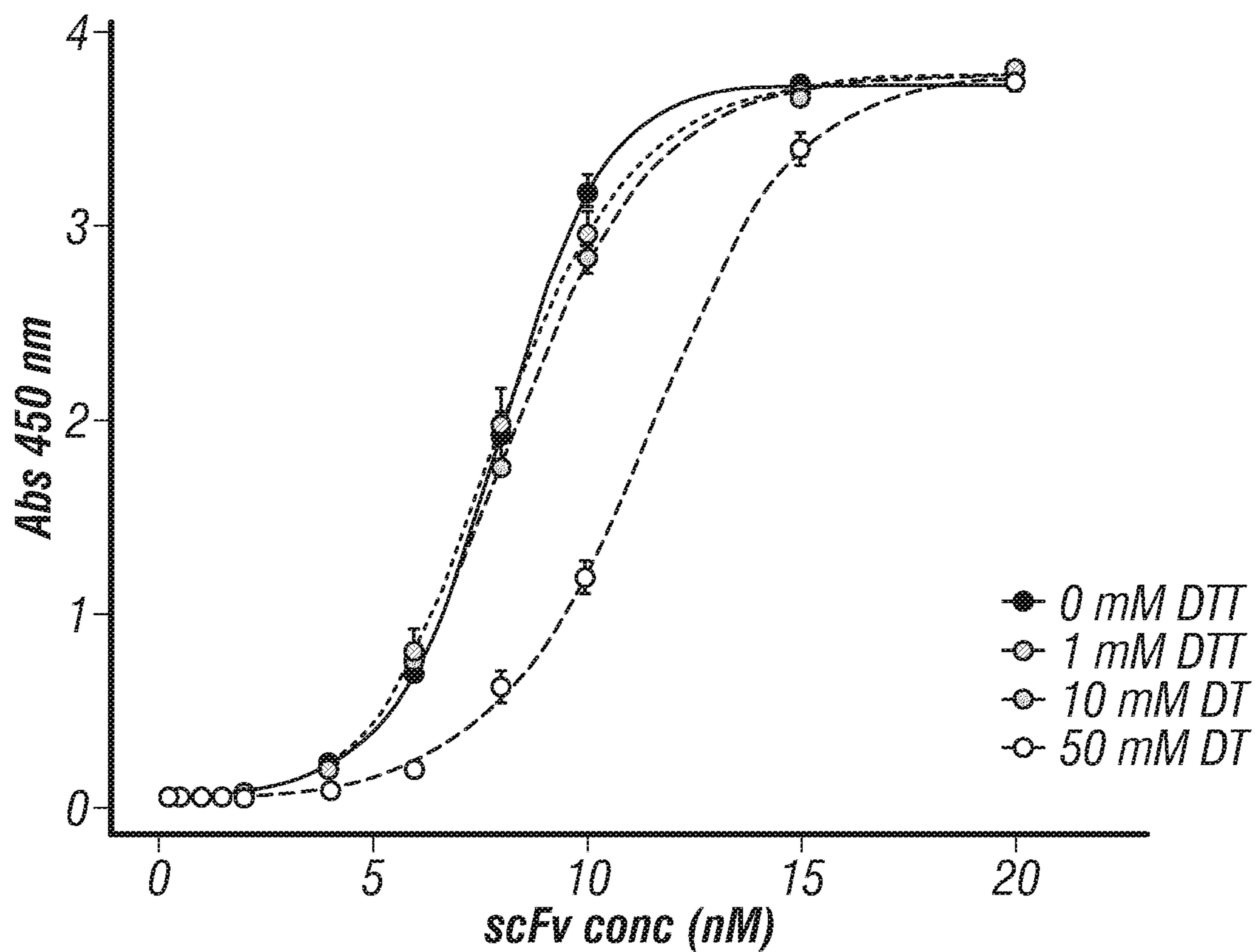


FIG. 7B

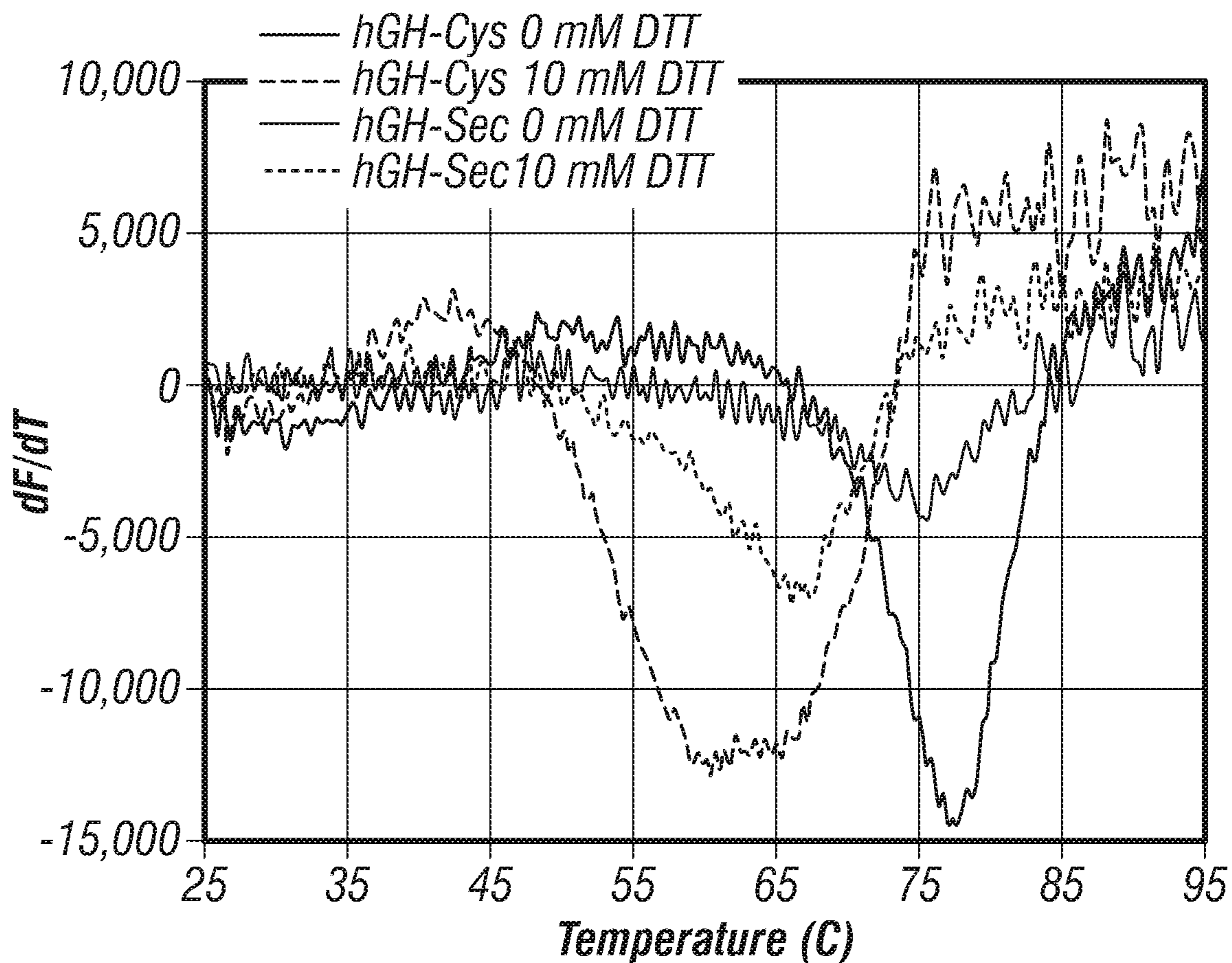


FIG. 8A

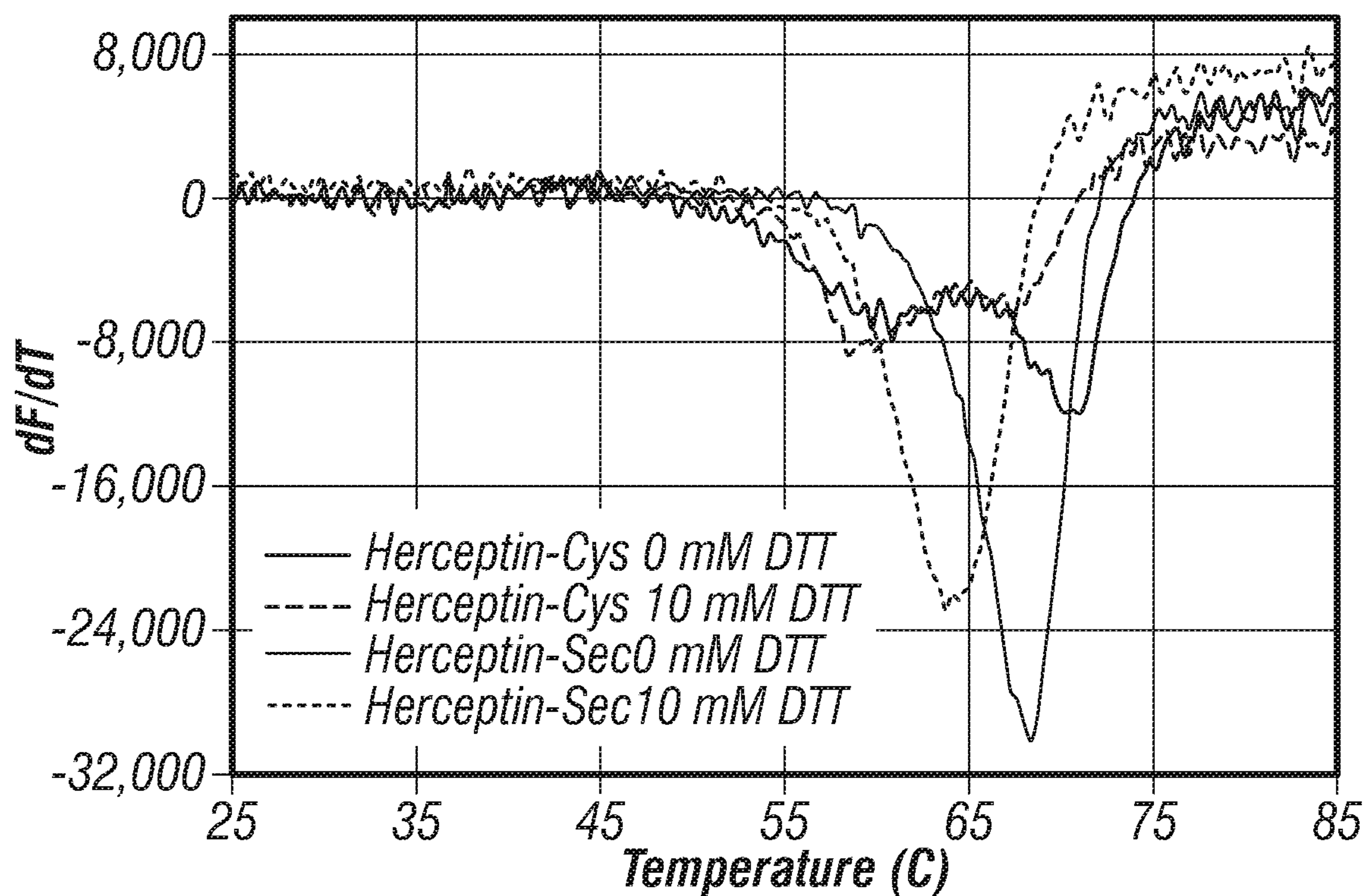
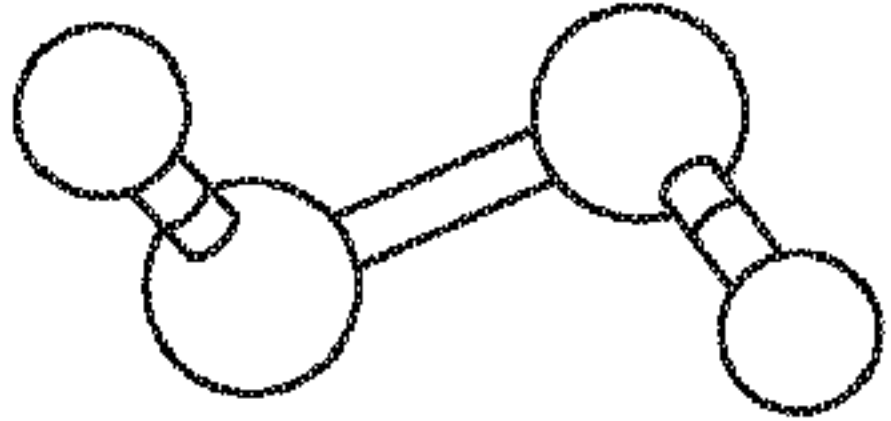
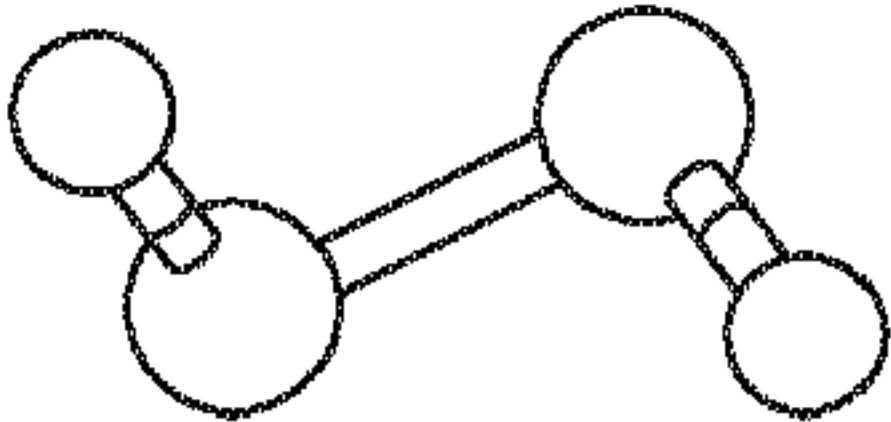
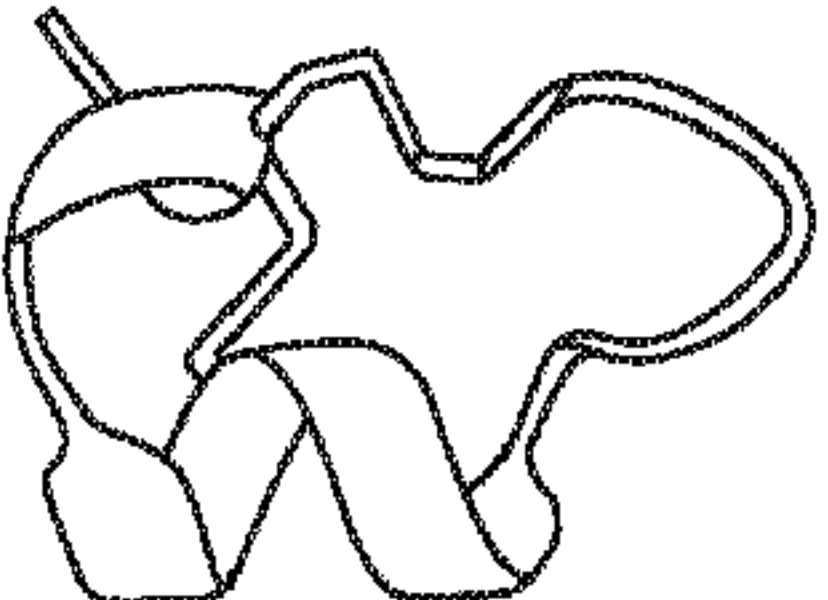


FIG. 8B

<i>Bond</i>	<i>Geometry</i>	<i>Covalent radius (Å)</i>	<i>Bond length (Å)</i>	<i>Redox potential (mv)</i>
<i>Disulfide</i>		1.02	2.03	-220
<i>Diselenide</i>		1.17	2.33	-380
<p><i>X-ray crystallography shows disulfides (yellow) and diselenides (red) are isosteric</i></p> 				
<p><i>Muttenthaler et al, JACS, 2010</i></p>				

The redox potential of cytoplasm is -280 to -300 mV. Only diselenides can form and maintain in the cytosol

FIG. 1