



(86) Date de dépôt PCT/PCT Filing Date: 2005/09/05
(87) Date publication PCT/PCT Publication Date: 2006/03/09
(45) Date de délivrance/Issue Date: 2014/12/02
(85) Entrée phase nationale/National Entry: 2007/03/01
(86) N° demande PCT/PCT Application No.: EP 2005/009528
(87) N° publication PCT/PCT Publication No.: 2006/024547
(30) Priorité/Priority: 2004/09/03 (EP04 425 665.9)

(51) Cl.Int./Int.Cl. *C07K 14/47* (2006.01),
C12N 15/00 (2006.01)
(72) Inventeurs/Inventors:
FUMERO, SILVANO, IT;
BERTARIONE RAVA ROSSA, LUISA, IT;
BARONE, DOMENICO G., IT;
DRITTANTI, LILA, FR;
GUYON, THIERRY, FR;
BORRELLY, GILLES, FR;
CANEPA, BARBARA, IT;
LORENZETTO, CHIARA, IT
(73) Propriétaire/Owner:
CREABILIS THERAPEUTICS S.P.A., IT
(74) Agent: NORTON ROSE FULBRIGHT CANADA

(54) Titre : UTILISATION DE POLYPEPTIDES OBTENUS PAR LE BIAIS DE MUTATIONS SYSTEMATIQUES D'ACIDES AMINES UNIQUES D'UNE BOITE A HUMAINS ET NON HUMAINS DE HMGB1 POUR PREVENIR OU ANTAGONISER LES MALADIES INDUITES PAR LE HMGB
(54) Title: USE OF POLYPEPTIDES OBTAINED THROUGH SYSTEMATIC MUTATIONS OF SINGLE AMINO ACIDS OF HUMAN AND NON-HUMAN BOX-A OF HMGB1 TO PREVENT AND/OR ANTAGONIZE PATHOLOGIES INDUCED BY HMGB1

(57) **Abrégé/Abstract:**

The present invention relates to polypeptide variants of the HMGB-1 highaffinity binding domain Box-A (HMGB1 Box-A) or to a biologically active fragment of HMGB1 Box-A, which are obtained through systematic mutations of single amino acids of the wild-type HMGB1 Box-A protein and which show an increased resistance to proteases and which are therefore characterized by more favourable pharmacokinetic and pharmacodynamic profiles. Moreover, the present invention concerns the use of said polypeptide molecules of HMGB1 Box-A to diagnose, prevent, alleviate and/or treat pathologies associated with extracellular HMGB1.



(74) **Agent(suite/continued):** LLP/S.E.N.C.R.L., S.R.L.

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
9 March 2006 (09.03.2006)

PCT

(10) International Publication Number
WO 2006/024547 A3

(51) International Patent Classification:
C07K 14/47 (2006.01) **C12N 15/00** (2006.01)

(74) Agent: Weickmann & Weickmann; Postfach 860 820,
81635 München (DE).

(21) International Application Number:
PCT/EP2005/009528

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

(22) International Filing Date:
5 September 2005 (05.09.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
04 425 665.9 3 September 2004 (03.09.2004) EP

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): CRE-
ABILIS THERAPEUTICS S.P.A. [IT/IT]; BioIndustry
Park, Via Ribes, 5, I-10010 Colletterto Giacosa (IT).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): FUMERO, Sil-
vano [IT/IT]; Via delle Germane, 11, I-10015 Ivrea (IT).
BERTARIONE RAVA ROSSA, Luisa [IT/IT]; Via
Circonvallazione, 17/2, I-10018 Pavone Canavese(TO)
(IT). BARONE, Domenico, G. [IT/IT]; Corso Rosselli,
66, I-10129 Torino (IT). DRITTANTI, Lila [FR/FR];
49, rue Felix Faure, F-91270 Vigneux sur Siene (FR).
GUYON, Thierry [FR/FR]; 8, Résidence les Torres
Rouges, F-91120 Palaiseau (FR). BORRELLY, Gilles
[FR/FR]; 23b, rue Sainte Geneviève, 91860 Epinay sous
Senart (FR). CANEPA, Barbara [IT/IT]; Via Fassino, 39,
I-10020 San Sebastiano da Po (TO) (IT).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:
1 June 2006

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PROTEASE RESISTANT HUMAN AND NON-HUMAN HMGB1 BOX-A MUTANTS AND THEIR THERAPEUTIC/DIAGNOSTIC USE

(57) Abstract: The present invention relates to polypeptide variants of the HMGB-1 high affinity binding domain Box-A (HMGB1 Box-A) or to a biologically active fragment of HMGB1 Box-A, which are obtained through systematic mutations of single amino acids of the wild-type HMGB1 Box-A protein and which show an increased resistance to proteases and which are therefore characterized by more favourable pharmacokinetic and pharmacodynamic profiles. Moreover, the present invention concerns the use of said polypeptide molecules of HMGB1 Box-A to diagnose, prevent, alleviate and/or treat pathologies associated with extracellular HMGB1.

WO 2006/024547 A3

DEMANDES OU BREVETS VOLUMINEUX

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

CECI EST LE TOME __1__ DE __2__

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__

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

Use of polypeptides obtained through systematic mutations of single amino acids of human and non-human Box-A of HMGB1 to prevent and/or antagonize pathologies induced by HMGB1

Description

The present invention relates to polypeptide variants of the HMGB-1 high affinity binding domain Box-A (HMGB1 Box-A) or to a biologically active
10 fragment of HMGB1 Box-A, which are obtained through systematic mutations of single amino acids of the wild-type HMGB1 Box-A protein and which show an increased resistance to proteases and which are therefore characterized by more favourable pharmacokinetic and pharmacodynamic profiles. Moreover, the present invention concerns the use of said
15 polypeptide molecules of HMGB1 Box-A to diagnose, prevent, alleviate and/or treat pathologies associated with extracellular HMGB1.

Recent research in the field of sepsis and inflammation has led to an improved understanding of the pathogenic mechanisms and events
20 underlying their clinical onset and development. In the early stages of sepsis, for instance, bacterial endotoxins stimulate cells of the innate immune system which release pro-inflammatory cytokines (TNF, IL-1 α and IL-6). These early cytokines in turn induce the release of a later-acting downstream mediator (identified as the known protein HMGB1) that triggers
25 the pathological sequelae mediated by the subsequent release of cytokines such as TNF, IL-1 α , IL-1 β , IL-1Ra, IL-6, IL-8, IL-18, IFN- γ , PAF, etc., leading to a multisystem pathogenesis or to a lethal systemic inflammation (Andersson et al., 2002).

30 The HMGB1 protein belongs to the family of high mobility group (HMG) proteins. HMG proteins, so-called due to their high electrophoretic mobility in polyacrylamide gels, are the most ubiquitous non-histone proteins

- 2 -

associated with isolated chromatin in eukaryotic cells. These proteins play a generalized „architectural“ role in DNA bending, looping, folding and wrapping, since they either distort, bend or modify DNA structures and complexes with transcription factors or histones (Andersson et al., 2002; Agresti et al., 2003; Degryse et al., 2003). The high mobility group 1 (HMGB1) protein is usually a nuclear factor, in particular a transcriptional regulatory molecule causing DNA bending and facilitating the binding of several transcriptional complexes.

Structurally, the HMGB1 protein is a protein of approximately 25 kDa with a highly conserved sequence among mammals, whereby 2 out of 214 amino acids have conservative substitutions in all mammalian species. HMGB1 is ubiquitously present in all vertebrate nuclei and in particular can be found in fibroblasts, neurons, hepatocytes, glia and in cells derived from hematopoietic stem cells, including monocytes/macrophages, neutrophils and platelets. The HMGB1 molecule has a tripartite structure composed of three distinct domains: two DNA binding domains called HMG Box-A and Box-B, and an acid carboxyl terminus, making it bipolarly charged.

The two HMGB1 boxes are involved in the protein's function as non-sequence-specific architectural DNA-binding elements, conferring the ability to bind DNA into recognized distorted DNA structures and stabilizing nucleosome assembly, remodelling and sliding. Both the A- and B-HMG boxes are made up of highly conserved 84 amino acid residues, are strongly positively charged and are arranged in three α -helices having a similar L-shaped fold. The long arm of the “L” contains the N-terminal extended strand and helix III (Andersson et al. 2002; Agresti et al., 2003; Thomas, J. O. 2001), while the short arm comprises helices I and II. Structure-function analysis reveals that the pro-inflammatory cytokine domain of HMGB1 is the B-Box and in particular the sequence of its first 20 amino acids. The A-Box is an extremely weak agonist of the inflammatory cytokine release triggered by HMGB1 and competitively inhibits the pro-inflammatory activities of the B-Box and of the whole protein. Therefore, from a pharmacological point of

- 3 -

view, the A-Box acts as an antagonist of the pathological conditions induced and/or sustained by the B-Box and HMGB1.

5 The third domain, the carboxyl terminus or acidic tail, is extremely negatively charged since it contains 30 repetitive aspartic and glutamic acid residues, and is linked to the boxes by a basic region of about 20 residues. Mouse and rat HMGB1 differ from the human form by only two substitutions that are located in this continuous C-terminal stretch.

10 HMGB1 binds rather weakly to the B-form variety of linear double-stranded DNA with no sequence specificity, while it binds in the interior of the nucleus with high affinity to supercoiled DNA, to unusual DNA structures like 4-way junctions (cruciform DNA), bulged DNA and bent DNA (Ferrari et al., 1992; Pontiggia et al., 1993 and WO2006/002971 in the name of Creabilis
15 Therapeutics).

Besides its nuclear location and role as a transcription factor regulator, HMGB1 has also been found in the extracellular medium, actively released by activated cells of the immune systems (monocytes and macrophages) or
20 passively released by damaged or necrotic cells (Andersson et al., 2002; Scaffidi et al., 2002; Bonaldi et al., 2002; Taniguchi et al., 2003; Friedman et al., 2003; Palumbo et al., 2004).

25 Extracellularly released HMGB1 acts as a potent cytokine and as an extremely potent macrophage-stimulating factor. HMGB1 acts directly by binding to the cell membrane, inducing signaling and chemotaxis, having a chemokine-like function (Yang et al., 2001) and further acting indirectly by up-regulating the expression and secretion of pro-inflammatory cytokines. This makes extracellular HMGB1 protein a potent chemotactic and
30 immunoregulatory protein which promotes an effective inflammatory immune response. Furthermore, other proteins belonging to the family of HMG proteins, and which are able to bend DNA, are released together with HMGB1 in the extracellular medium. These proteins are inter alia HMGB2,

- 4 -

HMGB3, HMG-1L10, HMG-4L and SP100-HMG. They share with HMGB1 highly homologous amino acid sequences. Like HMGB1, they trigger/sustain inflammatory pathologies interacting with the same receptors, leading to the same downstream pathways of interaction.

5

In healthy cells, HMGB1 migrates to the cytoplasm both by passive and active transport. However, all cultured cells and resting monocytes contain the vast majority of HMGB1 in the nucleus, indicating that in baseline conditions import is much more effective than export. Cells might transport
10 HMGB1 from the nucleus by acetylating lysine residues which are abundant in HMGB1, thereby neutralizing their basic charge and rendering them unable to function as nuclear localization signals. Nuclear HMGB1 hyperacetylation determines the relocation of this protein from the nucleus to the cytoplasm (in the fibroblasts, for example) or its accumulation into
15 secretory endolysosomes (in activated monocytes and macrophages, for example) and subsequent redirection towards release through a non-classical vesicle-mediated secretory pathway. HMGB1 secretion by already activated monocytes is then triggered by bioactive lysophosphatidylcholine (LPC), which is generated later in the inflammation site from
20 phosphatidylcholine through the action of the secretory phospholipase sPLA2 produced by monocytes several hours after activation. Therefore, secretion of HMGB1 seems to be induced by two signals (Bonaldi et al., 2003) and to take place in three steps: 1) at first, an inflammatory signal promotes HMGB1 acetylation and its relocation from the nucleus to the
25 cytoplasm (step 1) and storage in cytoplasmic secretory vesicles (step 2); then, a secretion signal (extracellular ATP or lysophosphatidylcholine) promotes exocytosis (third step) (Andersson et al., 2002; Scaffidi et al. 2002; Gardella et al., 2002; Bonaldi et al., 2003; Friedman et al., 2003).

30 Released HMGB1 has been identified as one of the ligands binding to the RAGE receptor. This receptor is expressed in most cell types, and at a high level mainly in endothelial cells, in vascular smooth muscle cells, in

- 5 -

monocytes and macrophages and in mononuclear phagocytes. Recognition involves the C-terminal of HMGB1. The interaction of HMGB1 and RAGE triggers a sustained period of cellular activation mediated by RAGE up-regulation and receptor-dependent signaling. In particular, the interaction of HMGB1 and RAGE activates several intracellular signal transduction pathways, including mitogen-activated protein kinases (MAPKs), Cdc-42, p21ras, Rac and the nuclear translocation factor κ B (NF- κ B), the transcription factor classically linked to inflammatory processes (Schmidt et al., 2001).

According to several experimental evidences, released HMGB1 may also interact with receptors belonging to one or more subclasse(s) of the family of the Toll-like receptors. Further, HMGB1 may also interact with the functional N-terminal lectin-like domain (D1) of thrombomodulin. Due to the ability of the functional D1 domain of thrombomodulin to intercept and bind circulating HMGB1, the interaction with the RAGE receptors and the Toll-like receptors is prevented.

In the context of the present invention, "HMGB1" includes the non-acetylated form or/and the acetylated form of HMGB1. Likewise, "HMGB1 homologous proteins" include the non-acetylated form or/and the acetylated form of HMGB1 homologous proteins. Preferred HMGB1 homologous proteins are HMGB2, HMGB3, HMG-1L10, HMG-4L or/and SP100-HMG.

When released *in vivo*, HMGB1 is an extremely potent cytokine and a potent macrophage-stimulating factor. In fact, like other cytokine mediators of endotoxemia, HMGB1 activates *in vitro* a cascade of multiple pro-inflammatory cytokines (TNF, IL-1 α , IL-1 β , IL-1Ra, IL-6, IL-8, MIP-1 α and MIP-1 β) from human macrophages. Therefore, HMGB1 acts as a late mediator during acute inflammation and participates in an important way in the pathogenesis of systemic inflammation after the early mediator response has been resolved.

- 6 -

The observed pro-inflammatory effects of HMGB1 *in vitro* and the correlation between circulating HMGB1 levels and the development of the pathogenic sequence of systemic inflammation *in vivo* indicate that therapeutically targeting of this cytokine-like molecule should be of relevant clinical value, suggesting novel therapeutic approaches by a "late" administration of (selective) antagonists/ inhibitors of the extracellular activities of HMGB1.

Therefore, several attempts were performed in order to block this extracellular HMGB1 chemo-cytokine protein. Several important approaches were addressed to the administration of antibodies against HMGB1, of HMGB1 fragments (for example HMGB1 A-Box), of antibodies to RAGE, of soluble RAGE (sRAGE) and of ethyl pyruvate (Czura et al., 2003; Lotze et al., 2003).

15

The passive immunization of mice with HMGB1-neutralizing antibodies conferred a highly significant, dose-dependent and lasting protection against lethal doses of endotoxin, even when the first doses of antibodies were given after the TNF peak had passed, suggesting that antagonizing HMGB1 activity late in the clinical course may be an effective treatment approach to potentially lethal sepsis (Yang et al., 2004).

Another possibility is to administer mono- or oligoclonal antibodies against the HMGB1 B-Box, or its 20 amino acid relevant core which signals through RAGE. Furthermore, HMGB1 A-Box, one of the two DNA-binding domains in HMGB1, has been identified as a specific antagonist of HMGB1: highly purified recombinant A-Box has protected mice from lethal experimental sepsis even when initial treatment has been delayed for 24 hours after pathology induction, further suggesting that HMGB1 antagonists may be administered successfully in a clinically relevant window wider than the one used for other known cytokines (Yang et al., 2004).

- 7 -

Structural function analysis of HMGB1-truncated mutants has revealed that the A-Box domain of HMGB1 competitively displaces the saturable binding of HMGB1 to macrophages, specifically antagonizing HMGB1 activities. As
5 has been already seen with the protective activity of anti-HMGB1 antibodies, the administration of the A-Box rescues mice from sepsis even when treatment has been initiated as late as 24 hours after surgical induction of sepsis (Yang H. et al., 2004). HMGB1 antagonists or inhibitors selected from the group of antibodies or antibody fragments that bind to an HMGB1
10 protein, HMGB1 gene antisense sequences and HMGB1 receptor antagonists are known from US 6,468,533, WO 02/074337 and US 2003/0144201.

Moreover, saturation of circulating HMGB1 by the administration of sRAGE
15 leads to the block of its activities mediated by cellular RAGE, a result which can also be obtained by inhibiting RAGE itself with the administration of anti-RAGE antibodies.

Furthermore, a similar protective response late in the course of sepsis has
20 been observed by administering ethyl-pyruvate, a stable lipophilic derivative and relatively non-toxic food additive also used as an experimental anti-inflammatory agent, which attenuates the systemic inflammation of ischemia/reperfusion tissue injury and lethal hemorrhagic shock. Ethyl-pyruvate inhibited HMGB1 and TNF release *in vitro* from endotoxin-
25 stimulated murine macrophages, while *in vivo* protected mice from peritonitis-induced lethal sepsis, again when dosing was begun 24 hours after this pathology was experimentally induced.

Finally, it has been shown that the N-terminal lectin-like domain (D1) of
30 thrombomodulin is an inhibitor of HMGB1, since it binds to and sequesters this chemokine, preventing the binding of HMGB1 to RAGE and Toll-like

- 8 -

receptors such that the downstream cascade of events leading to inflammatory pathologies is inhibited.

As described above, several attempts were performed with the aim of
5 inhibiting and/or antagonising the extracellular HMGB1 chemo-cytokine
protein. The present invention is based on the experimental evidence that
the two high affinity binding domains for DNA, i.e. HMGB1 Box-A and
HMGB1 Box-B, which are present in the HMGB1 molecule, have two
opposing roles in the protein released in the extracellular space. The main
10 activity of HMGB1 Box-A is to mediate the pro-inflammatory activities
attributed to the HMGB1 protein. On the other hand, HMGB1 Box-A acts as
an antagonist competing with the pro-inflammatory activity of the Box-B
domain.

15 The problem underlying the present invention was therefore the provision of
novel agents for the prevention, alleviation and/or treatment of HMGB1-
associated pathologies. In particular, the problem of the present invention
was to develop novel agents as selective extracellular HMGB1 antagonist
and/or inhibitors, in order to prevent, alleviate and/or treat the broad
20 spectrum of pathological effects induced by the HMGB1 chemokine itself
and/or by the cascade of multiple inflammatory cytokines caused by the
extracellular release of the HMGB1 protein.

The solution to the above problem is therefore the provision of a polypeptide
25 variant of the human and/or non-human HMGB1 high affinity binding domain
Box-A (HMGB1 Box-A) or of a biologically active fragment of human and/or
non-human HMGB1 Box-A, characterized in that the amino acid sequence of
said polypeptide variant differs from the amino acid sequence of the wild
type HMGB1 Box-A protein by the mutation of one or more single amino
30 acids. Surprisingly, it was found by the inventors of the present invention
that said polypeptide variant exhibits an increased resistance to proteolysis

- 9 -

compared to wild type HMGB1 Box-A or to the biological active fragment of the wild type HMGB1 Box-A.

By increasing the resistance to the proteolytic activity of the proteases, a
5 more favourable pharmacokinetic and pharmacodynamic profile can be
achieved, since an increased stability in body fluids is obtained for the
inventive polypeptide variants. As a result thereof, an increase in the half-life
in body fluids of the protein's variants of the present invention is observed as
well. It is known that the estimated half-life of proteins *in vivo* can be as short
10 as a few minutes. The variants of the present invention preferably have an
increased half-life, e.g. because they are more resistant to proteases.

In a most preferred embodiment of the present invention, polypeptide
variants are obtained by using a directed evolution process, which
15 technology is extensively described in WO 2004/7022593 and in several
further patent applications (PCT/FR00/03503, PCT/FR01/01366, US
10/022,249, US 10/022,390, US 10/375,192, US 60/409,898, US
60/457,135, US 60/410,258 and US 60/410,263), all in the name of Nautilus
Biotech S.A. (Paris, France),

20

In general, the term "directed evolution" refers to biotechnological processes
devoted to the improvement of target protein features by means of specific
changes introduced into their amino acid sequences. The directed evolution
process includes the generation of a library of mutant versions of the gene of
25 interest, followed by the selection of those variants that display the desired
features. These processes can be iterative when gene products having an
improvement in a desired property are subjected to further cycles of
mutation and screening.

30 In order to optimise the Box-A of HMGB1 protein and to obtain the
polypeptide variants of the present invention with higher stability against

- 10 -

proteases, a particular Nautilus proprietary technology for directed evolution has been applied. In particular, a so-called two-dimensional rational mutagenesis scanning approach ("2-D scanning") has been applied, which is described in the Nautilus patent application WO 2004/022593,

5

Nautilus 2-D scanning approach for protein rational evolution is based on a process, in which two dimensions of the target protein are scanned by serial mutagenesis in order to find the right amino acid change that is needed at the right amino acid position. The first dimension that is scanned is the amino acid position along the target protein sequence, in order to identify those specific amino acid residues to be replaced with different amino acids. These amino acid positions are referred to as is-HIT target positions. The second dimension is the specific amino acid type selected for replacing a particular is-HIT target position. According to a particular approach of the 2-D scanning method, a number of target positions along the protein sequence are selected, *in silico*. As used herein, *in silico* refers to research and experiments performed using a computer. In this context, *in silico* methods include, but are not limited to, molecular modeling studies and biomolecular docking experiments. Therefore, the amino acid target positions on the protein sequence are identified without use of experimental biological methods. Once a protein feature to be optimised is selected, diverse sources of information or previous knowledge are exploited in order to determine those amino acid positions that may be amenable to improve the protein's fitness by replacement with a different amino acid. In particular the "is-HIT target positions" are identified based on three factors, being (i) the protein feature to be evolved and optimised, (ii) the protein's amino acid sequence and/or (iii) the known properties of the individual amino acids.

30 In the specific context of the present invention, the "*in silico* HITs" ("is-HITs") are all possible candidate amino acid positions along the target protein's primary sequence that might be involved as target for the proteolytic activity of proteases. Based on the specific list of proteases considered in the

- 11 -

context of the present invention (Fig. 1), the complete list of all amino acid sequences that could potentially be targeted within the wild type HMGB1 Box-A amino acid sequence is determined.

5 Once the is-HIT target positions have been selected, mutagenesis then is performed by the replacement of single amino acid residues at the specific acid target positions on the protein backbone. The mutagenesis is performed by residue replacement "one-by-one" in addressable arrays and molecules containing the preselected amino acid changes at each of the
10 target amino acid positions are produced.

The choice of the replacing amino acid takes into account the need to preserve the physicochemical properties such as hydrophobicity, charge and/or polarity of essential residues (such as catalytic and binding residues).
15 Numerous methods of selecting replacing amino acids are well known in the art, in particular, amino acid substitution matrixes are used for this purpose. A very preferred technology according to the present invention makes use of the so-called "Percent Accepted Mutation" (PAM) (Dayhoff et al., Atlas of protein sequence and structure, 5(3):345-352, 1978), as shown in Fig. 2.
20 PAM values are used in order to select an appropriate group of replacement amino acids. PAM values, originally developed to produce alignments between protein sequences, are available in the form of probability matrixes, which reflect an evolutionary distance. "Conservative substitutions" of a residue in a reference sequence are those substitutions that are physically
25 and functionally similar to the corresponding reference residues, e.g. those that have a similar size, shape, electric charge, chemical properties, including the ability to form covalent or hydrogen bonds, or the like. Preferred conservative substitutions show the highest scores fitting with the PAM matrix criteria in the form of "accepted point mutations". The PAM250
30 matrix is used in 2-D scanning to identify the replacing amino acids for the is-HITs in order to generate conservative mutations without affecting the protein function. At least, the two amino acids with the highest values in PAM250 matrix, corresponding to "conservative substitutions" or "accepted

- 12 -

point mutations", are chosen. The replacement of amino acids by cysteine residues is explicitly avoided, since this change would potentially lead to the formation of intermolecular disulfide bonds.

5 Using the above-resumed Nautilus Biotech directed evolution technology, the inventors of the present application were able to obtain polypeptide variants of the HMGB1 Box-A which differ from the amino acid sequence of the native target polypeptide by one or more mutation.

10 In the context of the present invention, where reference is made to the term "HMGB1 Box-A or amino acid sequence of HMGB1 Box-A", it is referred to both human and non-human HMGB1 Box-A. In a preferred embodiment of the present invention, the systematic mutation of single amino acid on the critical is-HITs positions for proteases has been obtained on the wild type of
15 human HMGB1 Box-A protein and on the wild type of *Anopheles gambia* HMGB1 Box-A protein. The choice of the species *Anopheles gambia* was made by the inventors of the present application after a proper structural and phylogenetic analysis showing a 68% identity and a 88% homology of the human and *Anopheles* HMGB1 Box-A.

20

"Biologically active fragments of HMGB1 Box-A" as used herein are meant to encompass parts of the known wild type HMGB1 Box-A protein, for which at least one of the biological activities of the corresponding mature protein is still observable when known tests are being used. Preferably, a fragment of
25 the mature protein is considered as biologically active if an antagonist activity with respect to the pro-inflammatory activity of the HMGB1 B-Box and the HMGB1 protein as a whole can be determined. Biologically active fragments of native HMGB1 Box-A are fragments of at least 20, 25, 30, 35, 45, 50, 55, 60, 65, 70, 75 or 80 amino acids. Preferred biologically active
30 fragments of native HMGB1 Box-A used in the context of the present invention comprises fragments of at least 77 or of at least 54 amino acids, respectively.

- 13 -

The term "mutation" as used in the context of the present invention can be understood as substitution, deletion and/ or addition of single amino acid in the target sequence. Preferably, the mutation of the target sequence in the present invention is a substitution. The substitution can occur with different
5 genetically encoded amino acid or by non-genetically encoded amino acids. Examples for non-genetically encoded amino acids are homocystein, hydroxyproline, ornithin, hydroxylysine, citrulline, carnitine, etc.

The polypeptide variants of the present invention obtained by using directed
10 evolution technology are mutant proteins which differ from the amino acid sequence of the wild type HMGB1 Box-A by the mutation of one or more single amino acid. In a very preferred embodiment of the present invention, only one amino acid replacement occurs on the sequence of the native protein. In this case, the polypeptide variant of the invention is obtained by
15 the modification of the native protein, such that the amino acid sequence of the variant differs from that of the native protein by a single amino acid change at only one of the is-HIT target positions. It is, however, encompassed by the subject of the present invention that the native protein can be further optimised by replacement of a plurality, e.g two or more, of is-
20 HIT target positions on the same protein molecule. According to this variant of the invention, polypeptide variants are obtained by combining the single mutation into a single protein molecule. The modified polypeptide variants having more single amino acid replacement can differ from the wild type protein sequence by amino acid replacements on 1-10, preferably 2, 3, 4, 5
25 and 6 different amino acid target positions.

The selection of the candidate lead of the series of polypeptide variants produced with the technology used in the present invention is based both on the more favourable pharmacokinetic profile, obtained thanks to the longer
30 resistance to proteases and on a better pharmacodynamic profile thanks to an increased intrinsic activity and binding affinity which gives a greater antagonistic activity than the native HMGB1 Box-A protein.

- 14 -

In a particular embodiment of the invention, starting from Human HMGB1 Box-A as starting native protein, three groups of polypeptide variants are obtained. In particular, one group of polypeptide variants is derived from single mutations introduced into the full-length amino acid sequence (84 amino acids) from Human HMGB1 Box-A. The other two groups of inventive polypeptide variants are generated starting from biologically active fragments of Human HMGB1 Box-A of 77 amino acids and 54 amino acids, respectively.

In a further particular embodiment of the invention, starting from *Anopheles gambia* HMGB1 Box-A as starting native protein, three groups of polypeptide variants are obtained. In particular, one group of polypeptide variants is derived from single mutations introduced into the full-length amino acid sequence (84 amino acids) from *Anopheles gambia* HMGB1 Box-A. The other two groups of inventive polypeptide variants are generated starting from biologically active fragments of *Anopheles gambia* HMGB1 Box-A of 77 amino acids and 54 amino acids, respectively.

Hence, the above-mentioned very preferred polypeptide variants of this invention are defined as below.

1) On the human HMGB1 Box-A full-length fragment of 84 amino acids defined by the sequence SEQ ID NO:1 (Fig. 3a), 53 amino acid positions, recognized as substrate for different proteases (cf. Fig. 1), are identified. The numbering corresponds to that in the wild type protein:

K2, D4, P5, K6, K7, P8, R9, K11, M12, Y15, F17, F18, R23, E24, E25, K27, K28, K29, P31, D32, F37, E39, F40, K42, K43, E46, R47, W48, K49, M51, K54, E55, K56, K58, F59, E60, D61, M62, K64, D66, K67, R69, Y70, E71, R72, E73, M74, K75, Y77, P79, P80, K81, E83.

The native amino acid at each of these positions is replaced by residues defined by the substitution matrix PAM250 (cf. Fig. 2). In particular, the

- 15 -

performed residue substitutions are as listed below.

R to H, Q
E to H, Q, N
5 K to Q, T
D to N, Q
M to I, V
P to A, S
Y to I, H
10 F to I, V
W to Y, S

A total of 115 polypeptide variants of Box-A of human HMGB1 are generated (Fig. 3a). These polypeptide variants are defined in sequences SEQ ID
15 NOs:2 to 116.

2) On the Human HMGB1 Box-A biologically active fragment of 77 amino acids, defined in sequence SEQ ID NO:117 (Fig. 4a), 48 amino acid positions, recognized as substrate for different proteases (cf. Fig. 1), are
20 identified. The numbering is in accordance to their position in SEQ ID NO:117:

P1, R2, K4, M5, Y8, F10, F11, R16, E17, E18, K20, K21, K22, P24, D25,
F30, E32, F33, K35, K36, E39, R40, W41, K42, M44, K47, E48, K49, K51,
25 F52, E53, D54, M55, K56, D59, K60, R62, Y63, E64, R65, E66, M67, K68,
Y70, P72, P73, K74, E76.

The native amino acid in each of these positions is replaced by residues defined by the substitution matrix PAM250 (cf. Fig. 2). In particular, the
30 performed residue substitutions are as listed below.

R to H, Q
E to H, Q, N

- 16 -

K to Q, T

D to N, Q

M to I, V

P to A, S

5 Y to I, H

F to I, V

W to Y, S

10 A total of 105 polypeptide variants of Box-A of human HMGB1 fragment of 77 amino acids are generated (Fig. 4b) and defined as in sequences SEQ ID NOs:118 to 222.

15 3) On the Human HMGB1 Box-A biologically active fragment of 54 amino acids defined in sequence SEQ ID NO:223 (Fig. 5a), 35 amino acid positions, recognized as substrate for different proteases (Fig. 1), are identified. The numbering is in accordance to their position in SEQ ID NO:223:

20 P1, D2, F7, E9, F10, K12, K13, E16, R17, W18, K19, M21, K24, E25, K26, K28, F29, E30, D31, M32, K34, D36, K37, R39, Y40, E41, R42, E43, M44, K45, Y47, P49, P50, K51, E53.

25 The native amino acid at each of these positions is replaced by residues defined by the substitution matrix PAM250 (cf. Fig. 2). In particular, the performed residue substitutions are as listed below.

R to H, Q

E to H, Q, N

K to Q, T

30 D to N, Q

M to I, V

P to A, S

Y to I, H

- 17 -

F to I, V

W to Y, S

A total of 77 polypeptide variants of Box-A of human HMGB1 fragment of 54
5 amino acids are generated (Fig. 5b) and defined as in sequences SEQ ID
NOs:224 to 300.

4) On the *Anopheles gambia* (XP_311154) HMGB1 Box-A full-length
fragment of 84 amino acids, defined by the sequence SEQ ID NO:301 (Fig.
10 6a), 53 amino acid positions, recognized as substrate for different proteases
(Fig. 1), were identified. The numbering is in accordance with the position in
the native protein.

K2, K4, D5, K7, P8, R9, R11, M12, Y15, F17, F18, R23, E24, E25, K27, K28,
15 K29, P31, E32, E33, F37, E39, F40, R42, K43, E46, R47, W48, K49, M51,
L52, D53, K54, E55, K56, R58, F59, E61, M62, E64, K65, D66, K67, R69,
Y70, E71, L72, E73, M74, Y77, P79, P80, K81.

The native amino acid at each of these positions was replaced by residues
20 defined by the substitution matrix PAM250 (cf. Fig. 2).

The performed actual residue substitutions are as listed below.

R to H, Q

E to H, Q, N

25 K to Q, T

D to N, Q

M to I, V

P to A, S

Y to I, H

30 F to I, V

W to Y, S

A total of 117 variants of Box A of HMGB1 *Anopheles gambia* (XP_311154)

- 18 -

were generated (Fig. 6b) and identified in the sequences as defined in SEQ ID NOs:302 to 418.

5) On the *Anopheles gambia* (XP_311154) HMGB1 Box-A biologically active
 5 fragment of 77 amino acids, defined in sequence SEQ ID NO:419 (Fig. 7a),
 49 amino acid positions, recognized as substrate for different proteases (cf.
 Fig. 1), were identified. The numbering is in accordance with the position in
 the sequence as defined in SEQ ID NO:419.

10 P1, R2, R4, M5, Y8, F10, F11, R16, E17, E18, K20, K21, K22, P24, E25,
 E26, F30, E32, F33, R35, K36, E39, R40, W41, K42, M44, L45, D46, K47,
 E48, K49, R51, F52, E54, M55, E57, K58, D59, K60, R62, Y63, E64, L65,
 E66, M67, Y70, P72, P73, K74.

15 The native amino acid at each of these positions was replaced by residues
 defined by the substitution matrix PAM250 (cf. Fig. 2).

The performed actual residue substitutions are as listed below.

R to H, Q
 20 E to H, Q, N
 K to Q, T
 D to N, Q
 M to I, V
 P to A, S
 25 Y to I, H
 F to I, V
 W to Y, S

A total of 109 polypeptide variants of Box-A of HMGB1 fragment of 77 amino
 30 acids were generated (Fig. 7b) and identified as defined in sequences SEQ
 ID NOs:420 to 529.

6) On the *Anopheles gambia* (XP_311154) HMGB1 Box-A biologically active

- 19 -

fragment of 54 amino acids defined in sequence SEQ ID NO:530 (Fig. 8a), 36 amino acid positions, recognized as substrate for different proteases (cf. Fig. 1), were identified. The numbering is in accordance with the position on the sequence as defined in SEQ ID NO:530.

5

P1, E2, E3, F7, E9, F10, R12, K13, E16, R17, W18, K19, M21, L22, D23, K24, E25, K26, R28, F29, E31, M32, E34, K35, D36, K37, R39, Y40, E41, L42, E43, M44, Y47, P49, P50, K51.

10 The native amino acid in each of these positions was replaced by residues defined by the substitution matrix PAM250 (cf. Fig. 2).

The performed actual residue substitutions are as listed below.

R to H, Q

15 E to H, Q, N

K to Q, T

D to N, Q

M to I, V

P to A, S

20 Y to I, H

F to I, V

W to Y, S

25 A total of 81 polypeptide variants of Box-A of HMGB1 *Anopheles gambia* (XP_311154) fragment of 54 amino acids were generated (Fig. 8b) and identified in the sequences as defined in SEQ ID NOs:531 to 612.

30 It is noted that the amino acids which occur in the various amino acid sequences appearing herein are identified according to their known one-letter code abbreviations. It should be further noted that all amino acid residue sequences represented herein by their one-letter abbreviation code have a left-to-right orientation in the conventional direction of amino-terminus to carboxyl-terminus.

It is provided a polypeptide variant of the human high affinity binding domain Box-A (HMGB1 Box-A) or of a biologically active fragment of the human HMGB1 Box-A, characterised in that the amino acid sequence of the polypeptide variant differs from the amino acid sequence of the wild type HMGB1 Box-A by the mutation of one or more
5 single amino acid, wherein the polypeptide variant differs from the wild type HMGB1 Box-A sequence by the mutation of 1 to 10 single amino acid, preferably by only one single amino acid, wherein the polypeptide variant of the human HMGB1 Box-A is selected from from the group consisting of the amino acid sequences as defined in any of SEQ ID NOs:33, 35, 37-39, 42-45, 47-49, 52, 55, 57, 59, 62, 64, 67, 69 and 104, and
10 wherein a biologically active fragment of HMGB1 Box-A encompasses parts of the wild type HMGB1 Box-A protein, for which at least one of the biological activities of the corresponding mature protein is still observable when known tests are used.

- 20 -

Accordingly, the present invention provides modified polypeptide variants that exhibit increased resistance to the proteolytic activity of proteases and/or peptidases compared to the wild type HMGB1 Box-A protein. The polypeptide variants of the invention in particular exhibit an increase in the resistance to the proteolytic activity of the human proteases and/or peptidases, in particular of the human serum proteases and/or human gastro-intestinal proteases or peptidases. Preferred proteases are listed in Fig. 1. In a more preferred embodiment of the invention, polypeptide variants exhibit an increase in the resistance to the proteolytic activity of at least a protease selected from the group comprising chymotrypsin, trypsin, endoprotease, endopeptidases or a combination thereof.

In particular, the resistance to proteolysis is at least 10%, 20%, 30%, 40%, 50%, 70%, 80%, 90%, 95% or higher compared to the unmodified wild type HMGB1 Box-A. Protease resistance was measured at different timepoints (between 5 minutes and 8 hours) at 25°C after incubation of 20 µg of Box-A wild type or variants with a mixture of proteases at 1% w/w of total proteins. The mixture of the proteases was prepared freshly at each assay from stock solutions of endoproteinase Glu-C (SIGMA) 200 µg/ml; trypsin (SIGMA) 400µg/ml and α-chymotrypsin (SIGMA) 400 µg/ml. After protease incubation the reaction was stopped adding 10 µl of anti-proteases solution (Roche) and the samples were stored at -20°C for the biological activity assay.

As a consequence of the increased stability due to the increased resistance to proteases activity, the polypeptide variants of the present invention also exhibit a longer half-life in body fluids compared to the wild type HMGB1 Box-A. In particular, the half-life in serum and/or in blood is increased, whereby an increase of at least 10 minutes, 20 minutes, 30 minutes, 60 minutes or even longer, compared to the wild type HMGB1 Box-A is observed.

A further aspect of the present invention is a nucleic acid molecule encoding

- 21 -

a polypeptide variant of the present invention. In particular, the present invention refers to nucleic acid molecules encoding polypeptide variants as defined in SEQ ID NO:2 to 116, 118 to 222, 224 to 300, 302 to 418, 420 to 526 and 531 to 612.

5

A still further aspect of the present invention is a vector comprising a nucleic acid molecule as defined above.

10

Furthermore, the present invention refers to a method for producing a polypeptide variant as described above comprising (i) introducing a nucleic acid molecule as defined above into a host cell and (ii) culturing the cell, under conditions in which the encoded polypeptide variant is expressed. Preferably the host cell is a mammalian, insect or bacterial cell, in particular E. Coli, preferably the M15 strain.

15

A further method for producing a polypeptide variant as described above is the use of chemical peptide synthesis, e.g. a solid phase peptide synthesis according to standard methods.

20

The polypeptide variants of the present invention exhibit an increased resistance to proteolysis and thus a higher stability compared to the unmodified wild type protein. Consequently, the peptides of the invention also exhibit improved therapeutic and biological properties and activity. In fact, they show a more favorable pharmacokinetic and pharmacodynamic profile than native HMGB1 Box-A.

25

The invention is therefore directed to the use of the above-mentioned polypeptide variants of HMGB1 Box-A, obtained through systematic mutations of single amino acids in the sequence of HMGB1 Box-A or of its biologically active fragments as active agent in a medicament.

30

A still further aspect of the invention is hence the use of the inventive polypeptide variants for the manufacture of a medicament for the prevention

- 22 -

and/or treatment of extracellular HMGB1-associated pathologies or pathologies associated with the HMGB1 homologous proteins. In particular, the HMGB1 associated pathologies are pathologies which are mediated by a multiple inflammatory cytokine cascade.

5

The broad spectrum of pathological conditions induced by the HMGB1-chemokine and by the HMGB1-induced cascade of inflammatory cytokines are grouped in the following categories: inflammatory disease, autoimmune disease, systemic inflammatory response syndrome, reperfusion injury after organ transplantation, cardiovascular affections, obstetric and gynecologic disease, infectious (viral and bacterial) disease, allergic and atopic disease, solid and liquid tumor pathologies, transplant rejection diseases, congenital diseases, dermatological diseases, neurological diseases, cachexia, renal diseases, iatrogenic intoxication conditions, metabolic and idiopathic diseases.

15

HMGB1-associated pathologies according to the present invention are preferably pathological conditions mediated by activation of the inflammatory cytokine cascade. Non limiting examples of conditions which can be usefully treated using the present invention include the broad spectrum of pathological conditions induced by the HMGB1-chemokine and by the HMGB1-induced cascade of inflammatory cytokines grouped in the following categories: restenosis and other cardiovascular diseases, reperfusion injury, inflammation diseases such as inflammatory bowel disease, systemic inflammation response syndrome, e.g. sepsis, adult respiratory distress syndrome, etc, autoimmune diseases such as rheumatoid arthritis and osteoarthritis, obstetric and gynaecological diseases, infectious diseases, atopic diseases, such as asthma, eczema, etc, tumor pathologies, e.g. solid or non-solid tumor diseases associated with organ or tissue transplants, such as reperfusion injuries after organ transplantation, organ rejection and graft-versus-host disease, congenital diseases, dermatological diseases such as psoriasis or alopecia, neurological diseases, ophthalmological diseases, renal, metabolic or idiopathic diseases and intoxication conditions,

20

25

30

- 23 -

e.g. iatrogenic toxicity, wherein the above diseases are caused by, associated with and/or accompanied by HMGB1 protein release.

In particular, the pathologies belonging to inflammatory and autoimmune
5 diseases include rheumatoid arthritis/seronegative arthropathies, osteoarthritis, inflammatory bowel disease, Crohn's disease, intestinal infarction, systemic lupus erythematosus, iridocyclitis/uveitis, optic neuritis, idiopathic pulmonary fibrosis, systemic vasculitis/Wegener's granulomatosis, sarcoidosis, orchitis/vasectomy reversal procedures. Systemic inflammatory
10 response includes sepsis syndrome (including gram positive sepsis, gram negative sepsis, culture negative sepsis, fungal sepsis, neutropenic fever, urosepsis, septic conjunctivitis), meningococemia, trauma hemorrhage, burns, ionizing radiation exposure, acute and chronic prostatitis, acute and chronic pancreatitis, appendicitis, peptic, gastric and duodenal ulcers, peritonitis, ulcerative, pseudomembranous, acute and ischemic colitis,
15 diverticulitis, achalasia, cholangitis, cholecystitis, enteritis, adult respiratory distress syndrome (ARDS). Reperfusion injury includes post-pump syndrome and ischemia-reperfusion injury. Cardiovascular disease includes cardiac stun syndrome, myocardial infarction and ischemia, atherosclerosis, thrombophlebitis, endocarditis, pericarditis, congestive heart failure and
20 restenosis. Obstetric and gynecologic diseases include premature labour, endometriosis, miscarriage, vaginitis and infertility. Infectious diseases include HIV infection/HIV neuropathy, meningitis, B- and C-hepatitis, herpes simplex infection, septic arthritis, peritonitis, E. coli 0157:H7, pneumonia epiglottitis, haemolytic uremic syndrome/thrombolytic thrombocytopenic
25 purpura, candidiasis, filariasis, amebiasis, malaria, Dengue hemorrhagic fever, leishmaniasis, leprosy, toxic shock syndrome, streptococcal myositis, gas gangrene, mycobacterium tuberculosis, mycobacterium avium intracellulare, pneumocystis carinii pneumonia, pelvic inflammatory disease, orchitis/epididymitis, legionella, Lyme disease, influenza A, Epstein-Barr
30 Virus, Cytomegalovirus, viral associated hemophagocytic syndrome, viral encephalitis/aseptic meningitis. Allergic and atopic disease include asthma, allergy, anaphylactic shock, immune complex disease, hay fever, allergic

- 24 -

rhinitis, eczema, allergic contact dermatitis, allergic conjunctivitis, hypersensitivity pneumonitis. Malignancies (liquid and solid tumor pathologies) include ALL, AML, CML, CLL, Hodgkin's disease, non Hodgkin's lymphoma, Kaposi's sarcoma, colorectal carcinoma, nasopharyngeal carcinoma, malignant histiocytosis and paraneoplastic syndrome/hypercalcemia of malignancy. Transplant diseases include organ transplant rejection and graft-versus-host disease. Congenital disease includes cystic fibrosis, familial hemophagocytic lymphohistiocytosis and sickle cell anemia. Dermatologic disease includes psoriasis, psoriatic arthritis and alopecia. Neurologic disease includes neurodegenerative diseases (multiple sclerosis, migraine, headache, amyloid-associated pathologies, prion diseases/Creutzfeld-Jacob disease, Alzheimer and Parkinson's diseases, multiple sclerosis, amyotrophic emilateral sclerosis) and peripheral neuropathies, migraine, headache. Renal disease includes nephrotic syndrome, hemodialysis and uremia. Iatrogenic intoxication condition includes OKT3 therapy, Anti-CD3 therapy, Cytokine therapy, Chemotherapy, Radiation therapy and chronic salicylate intoxication. Metabolic and idiopathic disease includes Wilson's disease, hemochromatosis, alpha-1 antitrypsin deficiency, diabetes, weight loss, anorexia, cachexia, obesity, Hashimoto's thyroiditis, osteoporosis, hypothalamic-pituitary-adrenal axis evaluation and primary biliary cirrhosis. Ophthalmological disease include glaucoma, retinopathies and dry-eye. A miscellanea of other pathologies comprehends: multiple organ dysfunction syndrome, muscular dystrophy, septic meningitis, atherosclerosis, epiglottitis, Whipple's disease, asthma, allergy, allergic rhinitis, organ necrosis, fever, septicaemia, endotoxic shock, hyperpyrexia, eosinophilic granuloma, granulomatosis, sarcoidosis, septic abortion, urethritis, emphysema, rhinitis, alveolitis, bronchiolitis, pharyngitis, epithelial barrier dysfunctions, pneumoultramicroscopic silicovolcanoconiosis, pleurisy, sinusitis, influenza, respiratory syncytial virus infection, disseminated bacteremia, hydatid cyst, dermatomyositis, burns, sunburn, urticaria, warst, wheal, vasulitis, angiitis, myocarditis, arteritis, periarteritis nodosa, rheumatic fever, celiac disease, encephalitis, cerebral embolism, Guillame-Barre syndrome,

- 25 -

neuritis, neuralgia, iatrogenic complications/peripheral nerve lesions, spinal
cord injury, paralysis, uveitis, arthriditis, arthralgias, osteomyelitis, fasciitis,
Paget's disease, gout, periodontal disease, synovitis, myasthenia gravis,
Goodpasture's syndrome, Babcets's syndrome, ankylosing spondylitis,
5 Barger's disease, Retier's syndrome, bullous dermatitis (bullous
pemphigoid), pemphigous and pemphigous vulgaris and alopecia.

In a further aspect of the invention, the use of the polypeptide variants
obtained through systematic mutations of amino acid sequences of human
10 and non-human Box-A of HMGB1, or of its biologically relevant fragments
described above, is in combination with a further agent.

The further agent is preferably an agent capable of inhibiting an early
mediator of the inflammatory cytokine cascade. Preferably, this further agent
15 is an antagonist or inhibitor of a cytokine selected from the group consisting
of TNF, IL-1 α , IL-1 β , IL-Ra, IL-6, IL-8, IL-10, IL 13, IL-18, IFN- γ MIP-1 α , MIF-
1 β , MIP-2, MIF and PAF.

The further agent used in combination with the polypeptide variants of
20 HMGB1 Box-A, or of its biologically relevant fragments, may also be an
inhibitor of RAGE, e.g. an antibody directed to RAGE, a nucleic acid or
nucleic acid analogue capable of inhibiting RAGE expression, e.g. an
antisense molecule, a ribozyme or a RNA interference molecule, or a small
synthetic molecule antagonist of the interaction of HMGB1 with RAGE,
25 preferably of the interaction of the non-acetylated or/and acetylated form of
HMGB1 with RAGE, or soluble RAGE (sRAGE). The antibody to RAGE is
preferably a monoclonal antibody, more preferably a chimeric or humanised
antibody or a recombinant antibody, such as a single chain antibody or an
antigen-binding fragment of such an antibody. The soluble RAGE analog
30 may be optionally present as a fusion protein, e.g. with the Fc domain of a
human antibody. The small synthetic molecular antagonist of the HMGB1
interaction with RAGE preferably has a molecular weight of less than 1000

- 26 -

Dalton. The small synthetic molecular antagonist preferably inhibits the interaction of RAGE with the non-acetylated form or/and with the acetylated form of HMGB1 and with the non-acetylated form or/and with the acetylated form of HMGB1 homologous proteins, particularly HMGB2, HMGB3, HMG-1L10, HMG-4L or/and SP100-HMG.

The further agent used in combination with the polypeptide variants of HMGB1 Box-A, or of its biologically relevant fragments, may also be an inhibitor of the interaction of a Toll-like receptor (TLR), e.g. of TLR2, TLR4, TLR7, TLR8 or/and TLR9, with HMGB1, which inhibitor is preferably a monoclonal or polyclonal antibody, a nucleic acid or nucleic acid analogue capable of inhibiting TLR expression, e.g. an antisense molecule, a ribozyme or a RNA interference molecule, or a synthetic molecule preferably having a size of less than 1000 Dalton. The inhibitor may be a known inhibitor of a Toll-like receptor, in particular of TLR2, TLR4, TLR7, TLR8 or/and TLR9. The inhibitor preferably inhibits the interaction of the Toll-like receptor with the non-acetylated form or/and the acetylated form of HMGB1 and with the non-acetylated form or/and with the acetylated form of HMGB1 homologous proteins, in particular HMGB2, HMGB3, HMG-1L10, HMG-4L or/and SP100-HMG.

In still another embodiment, the further agent used in combination with the polypeptide variants of HMGB1 Box-A, or of its biologically relevant fragments, is the functional N-terminal lectin-like domain (D1) of thrombomodulin. The D1 domain of thrombomodulin is able to intercept the non-acetylated form and/or the acetylated form of released HMGB1 and of released HMGB1 homologous proteins, in particular HMGB2, HMGB3, HMG-1L10, HMG-4L or/and SP100-HMG, preventing thus their interaction with RAGE and Toll-like receptors. The D1 domain of thrombomodulin may be native or mutated in order to make it resistant to proteases.

The further agent may also be a synthetic double-stranded nucleic acid or

- 27 -

nucleic acid analogue molecule with a bent shape structure, particularly a double-stranded bent DNA, PNA or DNA/PNA chimera or hybrid or a double-stranded cruciform DNA, PNA or DNA/PNA chimera or hybrid structure, capable of binding to the HMGB1 protein. Preferred nucleic acids and
5 nucleic analogue molecules are disclosed in a co-owned and co-pending international patent application publication no. WO2006/002971 filed on 4 July 2005.

The synthetic
double-stranded nucleic acid or nucleic acid analogue molecule with a bent
10 shape structure is preferably capable of binding to the non-acetylated or/and to the acetylated form of HMGB1 and the non-acetylated or/and the acetylated form of HMGB1 homologous proteins, in particular HMGB2, HMGB3, HMG-1L10, HMG4L or/and SP100-HMG.

15 In a still further embodiment, the further agent used in combination with the inventive polypeptide variants is K-252a or/and a salt or derivative thereof or a polymer conjugate of K-252a or/and of a derivative thereof. The use of K-252a or polymer conjugate of K-252a and derivatives thereof is disclosed in a co-owned and co-pending international patent application
20 publication no. WO2006/002971.

Therefore, a further aspect of the present invention is a pharmaceutical composition comprising an effective amount of at least one of the
25 polypeptide variants of HMGB1 Box-A or a biologically active fragment thereof as an active ingredient for the treatment of HMGB1-associated pathologies and pharmaceutically acceptable carriers, diluents and/or adjuvants. The pharmaceutical composition of the present invention is preferably suitable for the treatment of pathologies associated with the non-
30 acetylated or/and the acetylated form of HMGB1 and/or of HMGB1 homologous proteins. In a further preferred embodiment, the pharmaceutical composition of the present invention comprising the at least one polypeptide variant also comprises a further agent as defined above. The

- 28 -

pharmaceutical composition of the present invention may be used for diagnostic or for therapeutic applications.

5 The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's conditions. Administration may be achieved in a single dose or repeated doses at intervals. Dosage amount and interval may be adjusted individually in order to provide the therapeutical effect which results in amelioration of symptoms or a prolongation of the survival in a patient. The actual amount of composition
10 administered will, of course, be dependent on the subject being treated, on the subject's weight, the severity of the affliction, the manner of administration and the judgement of the prescribing physician. A suitable daily dosage will be between 0,001 to 10 mg/kg, particularly 0,1 to 5 mg/kg.

15 The administration may be carried out by known methods, e.g. by injection, in particular by intravenous, intramuscular, transmucosal, subcutaneous or intraperitoneal injection and/or by oral, topical, nasal, inhalation, aerosol and/or rectal application, etc. The administration may be local or systemic.

20 In addition, the variants of Box-A of HMGB1, or of its pharmacologically active fragments, object of this invention can be reversibly immobilized and/or adsorbed on the surface and/or inside medical devices or drug release/vehicling systems (microspheres). Medical devices and microspheres can be reversibly loaded with the polypeptide variants of Box-
25 A object of this invention, through their binding, impregnation and/or adsorption on the surface of the medical device or of the microsphere or on a layer that coats its surface. When the medical device or the microsphere come into contact with biological fluids, the reversibly immobilized variant of Box-A is released. Therefore, the medical device and the microsphere act as
30 drug-releasing tools that elute the molecule object of this invention in such a way that their release kinetics can be controlled, ensuring controlled or

- 29 -

sustained release, as required by the treatment. The methods for coating/impregnating the medical devices and loading microspheres are well known by experts in these technologies.

5 Thus, a further aspect of this invention is the way of using the variants of Box-A of HMGB1 or its pharmacologically relevant fragments, wherein the mutated polypeptide molecules are reversibly immobilized on the surface of medical devices or of microspheres or are adsorbed within them. These medical instruments are preferably surgical tools, implants, catheters or
10 stents, for example stents for angioplasty and, in particular, medicated drug-eluting stents.

Another aspect of the invention concerns a medical device reversibly coated with at least one polypeptide variant of the invention. Such a device can be
15 selected from surgical instruments, implants, catheters or stents. Such a device may be useful for angioplasty.

The invention is further illustrated by the following Figures and Examples. The examples are intended to exemplify generic processes and are included
20 for illustrative purpose only, without intention of limiting the scope of the present invention.

Fig. 1 shows the proteases used for the *in silico* identification of the amino acid positions (is-HITs) on the HMGB1 Box-A amino acid sequence which
25 are targets for the proteolytic activity.

Fig. 2 depicts the "Percent Accepted Mutation" (PAM 250) matrix. Values given to identical residues are shown in grey square. Highest values in the matrix are shown in black square and correspond to the highest occurrence
30 of substitution between two residues.

Fig. 3a displays the amino acid sequence of the native Human HMGB1 Box-A made of 84 amino acid residues. In bold, the amino acids sensitive to

- 30 -

proteases proteolysis are identified, showing the is-HIT residue positions.

Fig. 3b shows the type of replacing amino acids on the respective is-HITs target positions selected to generate the polypeptide variant of the full-length human HMGB1 Box-A. Further, the specific amino acid sequences of the generated polypeptide variant are displayed in SEQ ID NOs:2 to 116.

Fig. 4a displays the amino acid sequence of the biologically active fragment of Human HMGB1 Box-A made of 77 amino acid residues. In bold, the amino acids sensitive to proteases proteolysis are identified, showing the is-HIT residue positions.

Fig. 4b shows the type of replacing amino acids on the respective is-HITs target positions selected to generate the polypeptide variant of the biologically active fragment of Human HMGB1 Box-A made of 77 amino acid residues. Further the specific amino acid sequences of the generated polypeptide variant are displayed in SEQ ID NOs: 118 to 222.

Fig. 5a displays the amino acid sequence of the biologically active fragment of Human HMGB1 Box-A made of 54 amino acid residues. In bold, the amino acids sensitive to proteases proteolysis are identified, showing the is-HIT residue positions.

Fig. 5b shows the type of replacing amino acids on the respective is-HITs target positions selected to generate the polypeptide variant of the biologically active fragment of Human HMGB1 Box-A made of 54 amino acid residues. Further, the specific amino acid sequences of the generated polypeptide variant are displayed in SEQ ID NOs: 224 to 300.

Fig. 6a displays the amino acid sequence of the native *Anopheles gambia* HMGB1 Box-A made of 84 amino acid residues. In bold, the amino acids sensitive to proteases proteolysis are identified, showing the is-HIT residue positions.

- 31 -

Fig. 6b shows the type of replacing amino acids on the respective is-HITs target positions selected to generate the polypeptide variant of the full-length *Anopheles gambia* HMGB1 Box-A. Further, the specific amino acid sequences of the generated polypeptide variant are displayed in SEQ ID NOs: 302 to 419.

Fig. 7a displays the amino acid sequence of the biologically active fragment of *Anopheles gambia* HMGB1 Box-A made of 77 amino acid residues. In bold, the amino acids sensitive to proteases proteolysis are identified, showing the is-HIT residue positions.

Fig. 7b shows the type of replacing amino acids on the respective is-HITs target positions selected to generate the polypeptide variant of the biologically active fragment of *Anopheles gambia* HMGB1 Box-A made of 77 amino acid residues. Further the specific amino acid sequences of the generated polypeptide variant are displayed in SEQ ID NOs: 420 to 529.

Fig. 8a displays the amino acid sequence of the biologically active fragment of *Anopheles gambia* HMGB1 Box-A made of 54 amino acid residues. In bold, the amino acids sensitive to proteases proteolysis are identified, showing the is-HIT residue positions.

Fig. 8b shows the type of replacing amino acids on the respective is-HITs target positions selected to generate the polypeptide variant of the biologically active fragment of *Anopheles gambia* HMGB1 Box-A made of 54 amino acid residues. Further, the specific amino acid sequences of the generated polypeptide variant are displayed in SEQ ID NOs: 531 to 612.

Fig. 9 shows the plasmid vector containing the nucleic acid sequence encoding for the polypeptide variant of the present invention. The plasmid contains the gene encoding for the polypeptide variant of the present invention, which is under control of the IPTG inducible T5 promoter. The

- 32 -

plasmid further contains an ampicillin resistant gene, a 6x His-tag and several restriction sites.

Fig. 10 shows a graph displaying the correlation between the TNF-alpha release induced by the stimulation of HMGB1 in RAW 264.7 cells.

Fig. 11 displays a dose-dependent inhibition of HMGB1-induced TNF-alpha release by a Box-A His-tagged protein.

10

EXAMPLES

1. PRODUCTION OF HMGB1 BOX-A NATIVE AND VARIANTS IN BACTERIA

The *in silico* generated variants of HMGB1 Box-A were cloned from HMGB1 protein into an inducible plasmid vector (Fig. 9) used to transform E. coli M15 strain competent cells. M15 cells were grown overnight in 1 mL of LB medium containing Kanamycin and Ampicillin in 96 deep-well plates under agitation (750 rpm). At OD_{600 nm} of 0.2-0.3 the cultures were diluted in 5 mL of LB medium in 24-well plates to reach an OD_{600 nm} of 0.07.

The M15 cells were incubated at 37°C under constant agitation (200 rpm). The production of Box-A (native or variants) was induced by the addition of IPTG (1mM final concentration) at OD_{600 nm} of 0.6. The culture was continued for three hours at 37°C under agitation (200 rpm). M15 cells were then harvested by centrifugation at 1000 g for 15 minutes, the supernatant was discarded and the pellet stored at -80°C at least for 1 hour before cells lysis and Box-A purification.

2. PURIFICATION OF HMGB1 BOX A NATIVE AND VARIANTS

- 33 -

M15 cells pellet was thawed on ice for 15 min. The cells were resuspended in 1 mL NPI-10 buffer containing 1 mg/mL Lysozyme and incubated for 30 min at RT under agitation at 750 rpm on a plate shaker. After the equilibration of Ni-NTA QIAfilter with 200 μ L of Superflow resin (QIAGEN catalog#969261) and 600 μ L of NPI-10 buffer the bacterial lysate was loaded and 200 μ L of absolute EtOH added. Four wash steps with 1 mL of NPI-20 were performed. The second and third washes were done with 1mL NPI-20 added with 100 μ g/mL Polymyxin (Fluka catalog#81271) in order to deplete LPS contaminants. After wash steps Box-A native and variants were eluted with 450 μ L NPI-250. The samples were stored at 4°C.

Box-A native and variants were re-purified with a DetoxiGel polymyxin column (PIERCE) at 4°C according to the supplier instructions. Finally the eluted proteins were filtered (0.22 μ m) in PBS and stored at 4°C to be tested.

3. BOX-A BIOLOGICAL ACTIVITY ASSAY

HMGB1 stimulates the secretion of TNF-alpha and of other cytokines as well as the proliferation of macrophages and monocytes. Box-A acts as an antagonist by inhibiting the activity of HMGB1.

The activity of Box-A native and variants produced were measured by the level of inhibition on the stimulation produced by HMGB1 on RAW 264.7 cells (murine macrophages, ATCC).

HMGB1 Box-A native and variants produced as described above were tested in a two-step process of screening directed to test i) their inhibition of HMGB1 induced TNF-alpha release and ii) their resistance to proteolysis.

30

In order to determine the proper HMGB1 concentration to be used in inhibition assay RAW 264.7 cells were seeded in 96 well plates (4×10^5 cells/well) and grown overnight in RPMI 1640 medium supplemented with

- 34 -

0.1% BSA. After overnight culture, cells were stimulated with HMGB1 (two times serial dilution concentrations between 100 µg/mL and 0.05 µg/mL) for 24 hours. The level of TNF-alpha produced was measured from cell media using ELISA (R&D systems), according to the manufacturer instructions. As presented in Fig. 10, HMGB1 significantly stimulated TNF-alpha release in macrophage cultures.

4. BOX-A INHIBITION OF HMGB1 TNF-ALPHA RELEASE AS SCREENING TEST

Murine macrophage-like RAW 264.7 cells were seeded in 96 well plates (4×10^5 cells/well) and grown overnight in RPMI 1640 medium supplemented with 0.1% BSA. After overnight culture, cells were stimulated with an adequate concentration of HMGB1 and Box-A native or variants or His-tagged (two times serial dilution between 20 µg/mL and 0.5 µg/mL) for 24 hours. The level of TNF-alpha was measured from cell media using ELISA (R&D systems), according to the manufacturer instructions.

Fig. 11 shows an example of dose-dependent inhibition of HMGB1 induced TNF release by Box-A, with an EC50 of 7.5 µg/ml (solid line). 100% inhibition of TNF-alpha release is obtained with a concentration of 20 µg/ml of Box-A. In parallel, TNF-alpha levels are measured in Box-A stimulated cells without HMGB1 in order to determine the presence or absence of contaminating endotoxin in Box-A preparation and quantify any non-HMGB1 dependent release of TNF-alpha. No release of TNF-alpha is observed at all concentrations of Box-A used in the assay (dashed line).

5. RESISTANCE TO PROTEOLYSIS OF BOX-A VARIANTS

Resistance of Box-A variants to proteolysis is determined as the residual biological activity (in the HMGB1/RAW cells system) following exposure to a mixture of selected proteases at increasing times of incubation.

- 35 -

20 µg of Box-A native or variants were treated with a mixture of proteases at 1% w/w of total proteins. The mixture of proteases was freshly prepared for each assay from stock solutions of endoproteinase Glu-C (SIGMA; 200 µg/ml), trypsin (SIGMA; 400µg/ml) and α-chymotrypsin (SIGMA; 400 µg/ml).

5

Samples were collected at different time points between 5 minutes and 8 hours of incubation with proteases after stopping the reaction with the addition of 10 µl of anti-proteases solution (Roche). Biological activity of each sample was then evaluated by the screening test described above in order to assess the residual activity at each time point.

10

- 36 -

REFERENCES

- Andersson, U., Erlandsson-Harris, H., Yang, H. and Tracey, K.J. (2002) *HMGB1 as a DNA-binding cytokine* J. Leucocyte Biol., 72: 1084-1091
- 5 Agresti, A. and Bianchi, M.E. (2003) *HMGB-proteins and gene expression* Current Opin. In Genetics and Develop., 13: 170-178
- Degryse, B., de Virgilio, M. (2003) *The nuclear protein HMGB1, a new kind of chemokine ?* FEBS Letters, 553: 11-17
- 10 Thomas, J.O. (2001) *HMGB1 and 2: architectural DNA-binding proteins* Biochemical Society Transactions, 29: 395-401
- Ferrari, S., Harley, V. H., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R. and Bianchi, M. E. (1992) *SRY, like HMGB1, recognizes sharp angles in DNA* The EMBO J., 11: 4497-4506
- 15 Pontiggia, A., Negri, A., Beltrame, M. and Bianchi, M.E. (1993) *Protein HU binds specifically to kinked DNA* Mol. Biol., 7: 343-350
- 20 Scaffidi, P., Misteli, T. and Bianchi, M.E. (2002) *Release of chromatin protein HMGB1 by necrotic cells triggers inflammation* Nature, 418: 191-195
- Bonaldi, T., Talamo, F., Scaffidi, P., Ferrera, D., Porto, A., Bachi, A., Rubartelli, A., Agresti, A. and Bianchi M.E. (2003) *Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards Secretion* The EMBO Journal, 22: 5551-5560
- 25 Taniguchi, N., Kawahara, K., Yone, K., Hashiguchi, T., Yamakuchi, M., Inoue, K., Yamada, S., Ijiri, K., Matsunaga, S., Nakajima, T., Komiya S. and Maruyama, I. (2003) *High mobility group box chromosomal protein 1 plays a role in the pathogenesis of arthritis as a novel cytokine* Arthritis and Rheumatism, 48:971-981
- 30

- 37 -

- Palumbo, R., Sanpaolesi, M., De Marchis, F., Tonlorenzi, R., Colombetti, S., Mondino, A., Cossu, G. and Bianchi, M.E. (2004) *Extracellular HMGB1, a signal of tissue damage, induces neoangioblast migration and proliferation* The J. of Cell Biology, 164: 441-449
- Friedman, S.G., Czura, C., J. and Tracey, K.J. (2003) *The gesture life of high mobility group box 1* Current Opinion in Clinical Nutrition and Metabolica Care, 6: 283-287
- Yang, H., Wang, H., and Tracey, K. J. (2001) *HMGB1 rediscovered as a cytokine* Shock, 15: 247-253
- Gardella, S., Andrei, C., Ferrera, D., Lotti, L.V., Torrisi, M.R., Bianchi, M.E. And Rubartelli, A. (2002) *The nuclear protein HMGB1 is secreted by monocytes via a non-classical, vesicle- mediated secretory pathway.* EMBO. Rep., 3: 995-1001
- Schmidt, A. M., Yan, S.D., Yan, S. F. and Stern, D. M. (2001) *The multiligand receptor RAGE as a progression factor amplifying immune and inflammatory responses* J. Clin. Invest., 108: 949-955
- Czura, C.J., Tracey, K.J. (2003) *Targeting high mobility group box 1 as a late acting mediator of inflammation* Crit. Care Med., 31: S46-S50
- Lotze, M. T. and De Marco, R. A. (2003) *Editorial overview – Dealing with death: HMGB1 as a novel target for cancer therapy* Current Opinion in Investigational Drugs, 4: 1405-1409
- Pullerits, R., Jonsson, I. M., Verdreng, M., Bokarewa, M., Andersson, U., Erlandsson-Harris, H. and Tarkowski, A. (2003) *High mobility group box chromosomal protein 1, a DNA binding cytokine, induces arthritis* Arthritis and Rheumatism, 48: 1693-1700

- 38 -

Kokkola, R., Li, J., Sundenberg, E., Aveberger, A. C., Palmblad, K., Yang, H., Tracey, K. J., Andersson, U. and Erlandsson-Harris, H. (2003) *Successful treatment of collagen-induced arthritis in mice and rats by targeting extracellular high mobility group box chromosomal protein 1 activity* Arthritis and Rheumatism, 48: 2052-2058

Yan, S. D., Chen, X., Fu, J., Chen, M., Zhu, H., Roher, A., Slattery, T., Zhao, L., Nagashima, M., Morser, J., Migheli, A., Nawroth, P., Stern, D., Schmidt, A. M. (1996) *RAGE and amyloid-beta peptide neurotoxicity in Alzheimer's disease* Nature, 382: 685-691

Yang, H., Ochani, M., Li, J., Qiang, X., Tanovic, M., Harris, H.E., Susarla, S.M., Ulloa, L., Wang, H., DiRaimo, R., Czura, C.J., Wang, H., Roth, J., Warren, H.S., Fink, M.P., Fenton, M.J., Andersson, U. and Tracey, K.J. (2004) *Reversing established sepsis with antagonists of endogenous high-mobility group box-1* PNAS, 101:296-301

DEMANDES OU BREVETS VOLUMINEUX

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

CECI EST LE TOME __1__ DE __2__

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__

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

CLAIMS

1. Polypeptide variant of the human high affinity binding domain Box-A (HMGB1 Box-A) comprising an amino acid sequence selected from the group consisting of SEQ ID NOs:33, 35, 37-39, 42-45, 47-49, 52, 55, 57, 59, 62, 64, 67, 69 and 104.
2. The polypeptide variant of claim 1, wherein the polypeptide variant differs from the wild type HMGB1 Box-A sequence by the mutation of only one single amino acid.
3. The polypeptide variant of claim 1 or 2, wherein said polypeptide variant exhibits an increased resistance to the proteolytic activity of proteases compared to the wild type HMGB1 Box-A.
4. The polypeptide variant of any one of claims 1-3, wherein the increase in resistance to proteolysis is in respect to at least one protease selected from the group consisting of chymotrypsin, trypsin, endoprotease, endopeptidase and a combination thereof.
5. The polypeptide variant of any one claims 1-4, wherein the increase in resistance to proteolysis is of 10%, 20%, 30%, 40%, 50%, 70%, 80%, 90% or 95% compared to the wild type HMGB1 Box-A.
6. The polypeptide variant of any one of claims 1-5, wherein the polypeptide variant exhibits a longer half life in body fluids compared to the wild type HMGB1 Box-A.
7. The polypeptide variant of claim 6, wherein the half life is 10 minutes, 20 minutes, 30 minutes or 60 minutes compared to the wild type HMGB1 Box-A.
8. A nucleic acid molecule encoding the polypeptide variant of any one of claims 1 to 7.
9. A vector comprising the nucleic acid molecule of claim 8.
10. A method for producing the polypeptide variant of any one of claims 1 to 7, comprising:
 - (i) introducing the nucleic acid molecule of claim 8 into a host cell; and
 - (ii) culturing the cell, under conditions in which the encoded polypeptide variant is expressed.

11. A method for producing the polypeptide variant of any one of claims 1 to 7 comprising the step of synthesizing said polypeptide using chemical peptide synthesis.
12. The polypeptide variant of any one of claims 1 to 7 for the treatment of inflammatory disease, autoimmune disease, systemic inflammatory response syndrome, reperfusion injury after organ transplantation, cardiovascular affections, obstetric and gynecologic disease, infectious viral or bacterial disease, allergic and atopic disease, solid and liquid tumor pathologies, transplant rejection diseases, congenital diseases, dermatological diseases, neurological diseases, cachexia, renal diseases, iatrogenic intoxication conditions, metabolic and idiopathic diseases, or ophthalmological diseases.
13. Use of the polypeptide variant of any one of claims 1 to 7 for the manufacture of a medicament for the prevention or treatment of inflammatory disease, autoimmune disease, systemic inflammatory response syndrome, reperfusion injury after organ transplantation, cardiovascular affections, obstetric and gynecologic disease, infectious viral or bacterial disease, allergic and atopic disease, solid and liquid tumor pathologies, transplant rejection diseases, congenital diseases, dermatological diseases, neurological diseases, cachexia, renal diseases, iatrogenic intoxication conditions, metabolic and idiopathic diseases, or ophthalmological diseases.
14. The use of claim 13 in combination with a further agent inhibiting an early mediator of the inflammatory cytokine cascade.
15. The use of claim 14, wherein the further agent is an antagonist or inhibitor of a cytokine selected from the group consisting of TNF, IL-1 α , IL-1 β , IL-R α , IL-6, IL-8, IL-10, IL-13, IL-18, IFN- γ , MIP-1 β , MIF-1 β , MIP-2, MIF and PAF.
16. The use of claim 14, wherein the further agent is an antibody to RAGE, a nucleic acid or nucleic acid analogue inhibiting RAGE expression, or a small synthetic molecule antagonist of the HMGB1 interaction with RAGE or soluble RAGE (sRAGE).
17. The use of claim 16, wherein the nucleic acid or nucleic acid analogue is an antisense molecule, a ribozyme or an RNA interference molecule.

18. The use of claim 14, wherein the further agent is an inhibitor of the interaction of a Toll-like receptor (TLR), with HMGB1, or a synthetic molecule having a size of less than 1000 Dalton.
19. The use of claim 18, wherein the TLR is at least one of TLR2, TLR4, TLR7, TLR8 and TLR9.
20. The use of claim 18 or 19, wherein the further agent is a monoclonal antibody, a polyclonal antibody, a nucleic acid or nucleic acid analogue inhibiting TLR expression.
21. The use of claim 20, wherein the nucleic acid or nucleic acid analogue is an antisense molecule, a ribozyme or an RNA interference molecule.
22. The use of claim 14, wherein the further agent is the N-terminal lectin-like domain (D1) of native or mutated thrombomodulin.
23. The use of claim 14, wherein the further agent is a synthetic double-stranded nucleic acid or nucleic acid analogue molecule with a bent shape structure, selected from bent or cruciform DNA, PNA, DNA/PNA chimeria and hybrid.
24. The use of claim 14, wherein the further agent is K-252a , a salt thereof or a polymer conjugate of K-252a.
25. A pharmaceutical composition comprising an effective amount of at least one polypeptide variant of any one of claims 1 to 7 and a pharmaceutically acceptable carrier.
26. The composition of claim 25, wherein the at least one polypeptide variant is in combination with at least one further agent as defined in any one of claims 15 to 24.
27. The composition of claims 25 or 26 for use in diagnosing of inflammatory disease, autoimmune disease, systemic inflammatory response syndrome, reperfusion injury after organ transplantation, cardiovascular affections, obstetric and gynecologic disease, infectious viral or bacterial disease, allergic and atopic disease, solid and liquid tumor pathologies, transplant rejection diseases, congenital diseases, dermatological diseases, neurological diseases, cachexia, renal diseases, iatrogenic

intoxication conditions, metabolic and idiopathic diseases, or ophthalmological diseases.

28. The composition of claims 25 or 26 for treating inflammatory disease, autoimmune disease, systemic inflammatory response syndrome, reperfusion injury after organ transplantation, cardiovascular affections, obstetric and gynecologic disease, infectious viral or bacterial disease, allergic and atopic disease, solid and liquid tumor pathologies, transplant rejection diseases, congenital diseases, dermatological diseases, neurological diseases, cachexia, renal diseases, iatrogenic intoxication conditions, metabolic and idiopathic diseases, or ophthalmological diseases.
29. Use of at least one polypeptide variant of any one of claims 1 to 7, for coating or impregnating medical devices, wherein said molecules are reversibly immobilised on the surface of medical devices.
30. The use of claim 29, wherein said medical devices are surgical instruments, implants, catheters or stents.
31. Medical device reversibly coated with at least one polypeptide variant of any one of claims 1 to 7.
32. The medical device of claim 31, wherein the medical device is selected from surgical instruments, implants, catheters and stents.

Figure 1

In silico identification of all amino acid positions that are targets for proteolysis using a large number of selected proteases and chemical treatments.

AspN	'D	Endoproteinase Asp-N
Chymo	(F,W,Y,M,L)~P	Chymotrypsin
Clos	R'	Clostripain
CnBr	M'	Cyanogen Bromide
IbzO	W'	IodosoBenzoate
Myxo	K'	Myxobacter
NH2OH	N'G	Hydroxylamine
pH2.5	D'P	pH 2.5
ProEn	P'	Proline Endopeptidase
Staph	E'	Staphylococcal Protease
Tryp	(K,R)~P	Trypsin
TrypK	K'~P	Trypsin(Arg blocked)
TrypR	R'~P	Trypsin(Lys blocked)

Figure 2 – Percent Accepted Mutation (PAM 250)

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	2	-2	0	0	-2	0	0	1	-1	-1	-2	-1	-1	-3	1	1	1	-6	-3	0
R	-2	6	0	-1	-4	1	-1	-3	2	-2	-3	3	0	-4	0	0	-1	2	-4	-2
N	0	0	2	2	-4	1	1	0	2	-2	-3	1	-2	-3	0	1	0	-4	-2	-2
D	0	-1	2	4	-5	2	3	1	1	-2	-4	0	-3	-6	-1	0	0	-7	-4	-2
C	-2	-4	-4	-5	12	-5	-5	-3	-3	-2	-6	-5	-5	-4	-3	0	-2	-8	0	-2
Q	0	1	1	2	-5	4	2	-1	3	-2	-2	1	-1	-5	0	-1	-1	-5	-4	-2
E	0	-1	1	3	-5	2	4	0	1	-2	-3	0	-2	-5	-1	0	0	-7	-4	-2
G	1	-3	0	1	-3	-1	0	5	-2	-3	-4	-2	-3	-5	0	1	0	-7	-5	-1
H	-1	2	2	1	-3	3	1	-2	6	-2	-2	0	-2	-2	0	-1	-1	-3	0	-2
I	-1	-2	-2	-2	-2	-2	-2	-3	-2	5	2	-2	2	1	-2	-1	0	-5	-1	4
L	-2	-3	-3	-4	-6	-2	-3	-4	-2	2	6	-3	4	2	-3	-3	-2	-2	-1	2
K	-1	3	1	0	-5	1	0	-2	0	-2	-3	5	0	-5	-1	0	0	-3	-4	-2
M	-1	0	-2	-3	-5	-1	-2	-3	-2	2	4	0	6	0	-2	-2	-1	-4	-2	2
F	-3	-4	-3	-6	-4	-5	-5	-5	-2	1	2	-5	0	9	-5	-3	-3	0	7	-1
P	1	0	0	-1	-3	0	-1	0	0	-2	-3	-1	-2	-5	6	1	0	-6	-5	-1
S	1	0	1	0	0	-1	0	1	-1	-1	-3	0	-2	-3	1	2	1	-2	-3	-1
T	1	-1	0	0	-2	-1	0	0	-1	0	-2	0	-1	-3	0	1	3	-5	-3	0
W	-6	2	-4	-7	-8	-5	-7	-7	-3	-5	-2	-3	-4	0	-6	-2	-5	17	0	-6
Y	-3	-4	-2	-4	0	-4	-4	-5	0	-1	-1	-4	-2	7	-5	-3	-3	0	10	-2
V	0	-2	-2	-2	-2	-2	-2	-1	-2	4	2	-2	2	-1	-1	-1	0	-6	-2	4

■ Value given for identical residues.

■ Positive value of substitution between two residues.

3/56

Figure 3a**Box A 84 amino acids**

Protection against proteolysis
If sequence:

GKGDPPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKE
KGKFEDMAKADKARYEREMKTYIPPKGET

In bold amino acids sensitive to proteases proteolysis

Figure 3b**Box A 84 amino acids**

Mutant list:

K2N	K27Q	E55Q	E73N
K2Q	K28N	E55H	M74I
D4N	K28Q	E55N	M74V
D4Q	K29N	K56N	K75N
P5A	K29Q	K56Q	K75Q
P5S	P31A	K58N	Y77H
K6N	P31S	K58Q	Y77I
K6Q	D32N	F59I	P79A
K7N	D32Q	F59V	P79S
K7Q	F37I	E60Q	P80A
P8A	F37V	E60H	P80S
P8S	E39Q	E60N	K81N
R9H	E39H	D61N	K81Q
R9Q	E39N	D61Q	E83Q
K11N	F40I	M62I	E83H
K11Q	F40V	M62V	E83N
M12I	K42N	K64N	
M12V	K42Q	K64Q	
Y15H	K43N	D66N	
Y15I	K43Q	D66Q	
F17I	E46Q	K67N	
F17V	E46H	K67Q	
F18I	E46N	R69H	
F18V	R47H	R69Q	
R23H	R47Q	Y70H	
R23Q	W48Y	Y70I	
E24Q	W48S	E71Q	
E24H	K49N	E71H	
E24N	K49Q	E71N	
E25Q	M51I	R72H	
E25H	M51V	R72Q	
E25N	K54N	E73Q	
K27N	K54Q	E73H	

Figure 3b continued**Box A 84 amino acid sequences:**

> sequence 1 Wild type

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 2 K2N

GNGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 3 K2Q

GQGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 4 D4N

GKGNPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 5 D4Q

GKGQPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 6 P5A

GKGDAKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 7 P5S

GKGDSKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 8 K6N

GKGDPNKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 9 K6Q

GKGDPQKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 10 K7N

GKGDPKNPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 11 K7Q

GKGDPKQPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 12 P8A

GKGDPKKARGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 13 P8S

GKGDPKKSARGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKE

Figure 3b continued

5/56

KGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 14 R9H

GKGDPKKPHGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 15 R9Q

GKGDPKKPQGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 16 K11N

GKGDPKKPRGNMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 17 K11Q

GKGDPKKPRGQMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 18 M12I

GKGDPKKPRGKISSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 19 M12V

GKGDPKKPRGKVSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 20 Y15H

GKGDPKKPRGKMSSHAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 21 Y15I

GKGDPKKPRGKMSSIAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 22 F17I

GKGDPKKPRGKMSSYAIFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 23 F17V

GKGDPKKPRGKMSSYAVFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 24 F18I

GKGDPKKPRGKMSSYAFIVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 25 F18V

GKGDPKKPRGKMSSYAFVVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 26 R23H

GKGDPKKPRGKMSSYAFFVQTCHEEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

Figure 3b continued

6/56

> sequence 27 R23Q

GKGDPKKPRGKMSSYAFFVQTCQEEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADK
AYEREMKTYIPPKGET

> sequence 28 E24Q

GKGDPKKPRGKMSSYAFFVQTCRQEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 29 E24H

GKGDPKKPRGKMSSYAFFVQTCRHEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 30 E24N

GKGDPKKPRGKMSSYAFFVQTCRNEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 31 E25Q

GKGDPKKPRGKMSSYAFFVQTCREQHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 32 E25H

GKGDPKKPRGKMSSYAFFVQTCREHHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 33 E25N

GKGDPKKPRGKMSSYAFFVQTCRENHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 34 K27N

GKGDPKKPRGKMSSYAFFVQTCREEHNKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 35 K27Q

GKGDPKKPRGKMSSYAFFVQTCREEHQKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 36 K28N

GKGDPKKPRGKMSSYAFFVQTCREEHKNKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 37 K28Q

GKGDPKKPRGKMSSYAFFVQTCREEHKQKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 38 K29N

GKGDPKKPRGKMSSYAFFVQTCREEHKNHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 39 K29Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKQHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

Figure 3b continued

7/56

> sequence 40 P31A

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHADASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 41 P31S

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHSDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 42 D32N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHHPNASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 43 D32Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHHPQASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 44 F37I

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNISEFSKKCSERWKTMSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 45 F37V

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNVSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 46 E39Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNFSEQFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 47 E39H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 48 E39N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 49 F40I

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNFSEISKCSERWKTMSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 50 F40V

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNFSEVSKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 51 K42N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNFSEFSNKC SERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 52 K42Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNFSEFSQKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 53 K43N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHDPASVNFSEFSKNC SERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

Figure 3b continued

> sequence 54 K43Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKQCSEWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 55 E46Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSQRWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 56 E46H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSHRWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 57 E46N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSNRWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 58 R47H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSEHWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 59 R47Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSEQWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 60 W48Y

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERYKTMSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 61 W48S

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERSKTMSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 62 K49N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWNTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 63 K49Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWQTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 64 M51I

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTISAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 65 M51V

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTVSAKEKGKFEDMAKADK
ARYEREMKTYIPPKGET

> sequence 66 K54N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSANEKGKFEDMAKAD
KARYEREMKTYIPPKGET

> sequence 67 K54Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAQEKGKFEDMAKAD

Figure 3b continued

9/56

KARYEREMKTYIPPKGET

> sequence 68 E55Q

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKQKGKFEDMAKAD
KARYEREMKTYIPPKGET**

> sequence 69 E55H

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKHKGKFEDMAKAD
KARYEREMKTYIPPKGET**

> sequence 70 E55N

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKNKGKFEDMAKAD
KARYEREMKTYIPPKGET**

> sequence 71 K56N

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKENGKFEDMAKAD
KARYEREMKTYIPPKGET**

> sequence 72 K56Q

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEQGKFEDMAKAD
KARYEREMKTYIPPKGET**

> sequence 73 K58N

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGNFEDMAKAD
KARYEREMKTYIPPKGET**

> sequence 74 K58Q

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGQFEDMAKAD
KARYEREMKTYIPPKGET**

> sequence 75 F59I

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKIEDMAKADK
ARYEREMKTYIPPKGET**

> sequence 76 F59V

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKVEDMAKAD
KARYEREMKTYIPPKGET**

> sequence 77 E60Q

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFQDMAKAD
KARYEREMKTYIPPKGET**

> sequence 78 E60H

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFHDMAKAD
KARYEREMKTYIPPKGET**

> sequence 79 E60N

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFNDMAKAD
KARYEREMKTYIPPKGET**

> sequence 80 D61N

**GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFENMAKAD
KARYEREMKTYIPPKGET**

> sequence 81 D61Q

Figure 3b continued

10/56

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEQMAKAD
KARYEREMKTYIPPKGET

> sequence 82 M62I

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDIAKADK
ARYEREMKTYIPPKGET

> sequence 83 M62V

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDVAKADK
ARYEREMKTYIPPKGET

> sequence 84 K64N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMANAD
KARYEREMKTYIPPKGET

> sequence 85 K64Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAQAD
KARYEREMKTYIPPKGET

> sequence 86 D66N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAN
KARYEREMKTYIPPKGET

> sequence 87 D66Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAQ
KARYEREMKTYIPPKGET

> sequence 88 K67N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
NARYEREMKTYIPPKGET

> sequence 89 K67Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
QARYEREMKTYIPPKGET

> sequence 90 R69H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KAHYEREMKTYIPPKGET

> sequence 91 R69Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KAQYEREMKTYIPPKGET

> sequence 92 Y70H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARHEREMKTYIPPKGET

> sequence 93 Y70I

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARIEREMKTYIPPKGET

> sequence 94 E71Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYQREMKTYIPPKGET

Figure 3b continued

11/56

> sequence 95 E71H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYHREMPTYIPPKGET

> sequence 96 E71N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYNREMPTYIPPKGET

> sequence 97 R72H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEHMPTYIPPKGET

> sequence 98 R72Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEQEMPTYIPPKGET

> sequence 99 E73Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYERQMPTYIPPKGET

> sequence 100 E73H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYERHMPTYIPPKGET

> sequence 101 E73N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYERNMPTYIPPKGET

> sequence 102 M74I

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREIKTYIPPKGET

> sequence 103 M74V

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREVKTYIPPKGET

> sequence 104 K75N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMNTYIPPKGET

> sequence 105 K75Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMQTYIPPKGET

> sequence 106 Y77H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTHIPPKGET

> sequence 107 Y77I

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTIIPPKGET

> sequence 108 P79A

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIAPKGET

12/56

Figure 3b continued

> sequence 109 P79S

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYISPKGET

> sequence 110 P80A

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPAKGET

> sequence 111 P80S

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPSKGET

> sequence 112K81N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPNGET

> sequence 113 K81Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPQGET

> sequence 114 E83Q

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGQT

> sequence 115 E83H

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGHT

> sequence 116 E83N

GKGDPKKPRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAD
KARYEREMKTYIPPKGNT

13/56

Figure 4a**Box A 77 amino acids**

Protection against proteolysis
If sequence:

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDM
AKADKARYEREMKTYIPPKGET**

In bold amino acids sensitive to proteases proteolysis

Figure 4b**Box A 77 amino acids**

Mutant list:

P1A	F30V	E53H	P73S
P1S	E32Q	E53N	K74N
R2H	E32H	D54N	K74Q
R2Q	E32N	D54Q	E76Q
K4N	F33I	M55I	E76H
K4Q	F33V	M55V	E76N
M5I	K35N	K57N	
M5V	K35Q	K57Q	
Y8H	K36N	D59N	
Y8I	K36Q	D59Q	
F10I	E39Q	K60N	
F10V	E39H	K60Q	
F11I	E39N	R62H	
F11V	R40H	R62Q	
R16H	R40Q	Y63H	
R16Q	W41Y	Y63I	
E17Q	W41S	E64Q	
E17H	K42N	E64H	
E17N	K42Q	E64N	
E18Q	M44I	R65H	
E18H	M44V	R65Q	
E18N	K47N	E66Q	
K20N	K47Q	E66H	
K20Q	E48Q	E66N	
K21N	E48H	M67I	
K21Q	E48N	M67V	
K22N	K49N	K68N	
K22Q	K49Q	K68Q	
P24A	K51N	Y70H	
P24S	K51Q	Y70I	
D25N	F52I	P72A	
D25Q	F52V	P72S	
F30I	E53Q	P73A	

Figure 4b continued**Box A 77 amino acid sequences**

> sequence 117 Wild type

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET

> sequence 118 P1A

ARGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET

> sequence 119 P1S

SRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET

> sequence 120 R2H

PHGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET

> sequence 121 R2Q

PQGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET

> sequence 122 K4N

PRGNMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET

> sequence 123 K4Q

PRGQMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET

> sequence 124 M5I

PRGKISSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET

> sequence 125 M5V

PRGKVSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET

> sequence 126 Y8H

PRGKMSSHAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET

> sequence 127 Y8I

PRGKMSSIAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET

> sequence 128 F10I

PRGKMSSYAIFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET

> sequence 129 F10V

PRGKMSSYAVFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE

Figure 4b continued**MKTYIPPKGET**

> sequence 130 F11I

**PRGKMSSYAFIVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 131 F11V

**PRGKMSSYAFVVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 132 R16H

**PRGKMSSYAFFVQTCHEEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 133 R16Q

**PRGKMSSYAFFVQTCQEEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 134 E17Q

**PRGKMSSYAFFVQTCRQEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 135 E17H

**PRGKMSSYAFFVQTCRHEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 136 E17N

**PRGKMSSYAFFVQTCRNEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 137 E18Q

**PRGKMSSYAFFVQTCREQHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 138 E18H

**PRGKMSSYAFFVQTCREHHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 139 E18N

**PRGKMSSYAFFVQTCRENHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 140 K20N

**PRGKMSSYAFFVQTCREEHNKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 141 K20Q

**PRGKMSSYAFFVQTCREEHQKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 142 K21N

**PRGKMSSYAFFVQTCREEHKNKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 143 K21Q

PRGKMSSYAFFVQTCREEHKQKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE

Figure 4b continued

16/56

MKTYIPPKGET

> sequence 144 K22N

**PRGKMSSYAFFVQTCREEHKKNHDPASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 145 K22Q

**PRGKMSSYAFFVQTCREEHKKQHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 146 P24A

**PRGKMSSYAFFVQTCREEHKKKHADASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 147 P24S

**PRGKMSSYAFFVQTCREEHKKKHSDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 148 D25N

**PRGKMSSYAFFVQTCREEHKKKHPNASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 149 D25Q

**PRGKMSSYAFFVQTCREEHKKKHPQASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 150 F30I

**PRGKMSSYAFFVQTCREEHKKKHPDASVNISEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 151 F30V

**PRGKMSSYAFFVQTCREEHKKKHPDASVNVSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 152 E32Q

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEQFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 153 E32H

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 154 E32N

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 155 F33I

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 156 F33V

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 157 K35N

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSNKKCSERWKTMSAKEKGKFEDMAKADKARYEREM

Figure 4b continued**KTYIPPKGET****> sequence 158 K35Q****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSQKCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET****> sequence 159 K36N****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKNCSEWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET****> sequence 160 K36Q****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKQCSERWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET****> sequence 161 E39Q****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSQRWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET****> sequence 162 E39H****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSHRWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET****> sequence 163 E39N****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSNRWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET****> sequence 164 R40H****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSEHWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET****> sequence 165 R40Q****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSEQWKTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET****> sequence 166 W41Y****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERYKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET****> sequence 167 W41S****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERSKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET****> sequence 168 K42N****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWNTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET****> sequence 169 K42Q****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWQTMSAKEKGKFEDMAKADKARYERE
MKTYIPPKGET****> sequence 170 M44I****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTISAKEKGKFEDMAKADKARYEREM
KTYIPPKGET****> sequence 171 M44V**

Figure 4b continued

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTVSAKEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 172 K47N

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSANEKGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 173 K47Q

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAQEKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 174 E48Q

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAQKKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 175 E48H

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKHKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 176 E48N

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKNKGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 177 K49N

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKENGKFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 178 K49Q

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEQGKFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 179 K51N

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGNFEDMAKADKARYEREM
KTYIPPKGET**

> sequence 180 K51Q

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGQFEDMAKADKARYERE
MKTYIPPKGET**

> sequence 181 F52I

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKIEDMAKADKARYEREM
KTYIPPKGET**

> sequence 182 F52V

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKVEDMAKADKARYERE
MKTYIPPKGET**

> sequence 183 E53Q

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFQDMAKADKARYERE
MKTYIPPKGET**

> sequence 184 E53H

**PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFHDMAKADKARYERE
MKTYIPPKGET**

Figure 4b continued

> sequence 185 E53N

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFNDMAKADKARYERE
MKTYIPPKGET

> sequence 186 D54N

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFENMAKADKARYEREM
KTYIPPKGET

> sequence 187 D54Q

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEQMAKADKARYERE
MKTYIPPKGET

> sequence 188 M55I

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDIAKADKARYEREM
KTYIPPKGET

> sequence 189 M55V

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDVAKADKARYEREM
KTYIPPKGET

> sequence 190 K57N

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMANADKARYEREM
KTYIPPKGET

> sequence 191 K57Q

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAQADKARYERE
MKTYIPPKGET

> sequence 192 D59N

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKANKARYEREM
KTYIPPKGET

> sequence 193 D59Q

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAQKARYERE
MKTYIPPKGET

> sequence 194 K60N

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADNARYEREM
KTYIPPKGET

> sequence 195 K60Q

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADQARYERE
MKTYIPPKGET

> sequence 196 R62H

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKAHYEREM
KTYIPPKGET

> sequence 197 R62Q

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKAQYERE
MKTYIPPKGET

> sequence 198 Y63H

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARHERE
MKTYIPPKGET

Figure 4b continued

> sequence 199 Y63I

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARIEREM
KTYIPPKGET

> sequence 200 E64Q

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYQRE
MKTYIPPKGET

> sequence 201 E64H

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYHRE
MKTYIPPKGET

> sequence 202 E64N

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYNRE
MKTYIPPKGET

> sequence 203 R65H

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEHM
KTYIPPKGET

> sequence 204 R65Q

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEQE
MKTYIPPKGET

> sequence 205 E66Q

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERQ
MKTYIPPKGET

> sequence 206 E66H

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERH
MKTYIPPKGET

> sequence 207 E66N

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERN
MKTYIPPKGET

> sequence 208 M67I

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREI
KTYIPPKGET

> sequence 209 M67V

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREV
KTYIPPKGET

> sequence 210 K68N

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
NTYIPPKGET

> sequence 211 K68Q

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
QTYIPPKGET

> sequence 212 Y70H

PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM

Figure 4b continued**KTHIPPKGET****> sequence 213 Y70I****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTIIPPKGET****> sequence 214 P72A****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIAPKGET****> sequence 215 P72S****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYISPKGET****> sequence 216 P73A****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPAKGET****> sequence 217 P73S****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPSKGET****> sequence 218 K74N****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPNGET****> sequence 219 K74Q****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPQGET****> sequence 220 E76Q****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGQT****> sequence 221 E76H****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGHT****> sequence 222 E76N****PRGKMSSYAFFVQTCREEHKKKHPDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREM
KTYIPPKGNT**

Figure 5a

Box A 54 amino acids

Protection against proteolysis
If sequence:

PDASVNFSEFSKKCSERWKTMSAKEKGFEDMAKADKARYEREMKTYIPPKGET

In bold amino acids sensitive to proteases proteolysis

Figure 5b

Box A 54 amino acids

Mutant list:

P1A	F29I	P50A
P1S	F29V	P50S
D2N	E30Q	K51N
D2Q	E30H	K51Q
F7I	E30N	E53Q
F7V	D31N	E53H
E9Q	D31Q	E53N
E9H	M32I	
E9N	M32V	
F10I	K34N	
F10V	K34Q	
K12N	D36N	
K12Q	D36Q	
K13N	K37N	
K13Q	K37Q	
E16Q	R39H	
E16H	R39Q	
E16N	Y40H	
R17H	Y40I	
R17Q	E41Q	
W18Y	E41H	
W18S	E41N	
K19N	R42H	
K19Q	R42Q	
M21I	E43Q	
M21V	E43H	
K24N	E43N	
K24Q	M44I	
E25Q	M44V	
E25H	K45N	
E25N	K45Q	
K26N	Y47H	
K26Q	Y47I	
K28N	P49A	
K28Q	P49S	

Figure 5b continued**Box A 54 amino acid sequences:**

> sequence 223 Wild type

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 224 P1A

ADASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 225 P1S

SDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 226 D2N

PNASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 227 D2Q

PQASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 228 F7I

PDASVNISEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 229 F7V

PDASVNVSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 230 E9Q

PDASVNFSEQFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 231 E9H

PDASVNFSEHFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 232 E9N

PDASVNFSEHFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 233 F10I

PDASVNFSEISKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 234 F10V

PDASVNFSEVSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 235 K12N

PDASVNFSEFSNKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 236 K12Q

PDASVNFSEFSQKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 237 K13N

PDASVNFSEFSKNCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 238 K13Q

PDASVNFSEFSKQCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 239 E16Q

PDASVNFSEFSKKCSQRWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

Figure 5b continued

> sequence 240 E16H

PDASVNFSEFSKKCSHRWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 241 E16N

PDASVNFSEFSKKCSNRWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 242 R17H

PDASVNFSEFSKKCSEHWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 243 R17Q

PDASVNFSEFSKKCSEQWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 244 W18Y

PDASVNFSEFSKKCSERYKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 245 W18S

PDASVNFSEFSKKCSERSKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 246 K19N

PDASVNFSEFSKKCSERWNTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 247 K19Q

PDASVNFSEFSKKCSERWQTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 248 M21I

PDASVNFSEFSKKCSERWKTISAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 249 M21V

PDASVNFSEFSKKCSERWKTVSAKEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 250 K24N

PDASVNFSEFSKKCSERWKTMSANEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 251 K24Q

PDASVNFSEFSKKCSERWKTMSAQEKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 252 E25Q

PDASVNFSEFSKKCSERWKTMSAQKKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 253 E25H

PDASVNFSEFSKKCSERWKTMSAKHKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 254 E25N

PDASVNFSEFSKKCSERWKTMSAKNKGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 255 K26N

PDASVNFSEFSKKCSERWKTMSAKENGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 256 K26Q

PDASVNFSEFSKKCSERWKTMSAKEQGKFEDMAKADKARYEREMKTYIPPKGET

> sequence 257 K28N

PDASVNFSEFSKKCSERWKTMSAKEKGNFEDMAKADKARYEREMKTYIPPKGET

> sequence 258 K28Q

Figure 5b continued

PDASVNFSEFSKKCSERWKTMSAKEKGQFEDMAKADKARYEREMKTYIPPKGET

> sequence 259 F29I

PDASVNFSEFSKKCSERWKTMSAKEKGKIEDMAKADKARYEREMKTYIPPKGET

> sequence 260 F29V

PDASVNFSEFSKKCSERWKTMSAKEKGKVEDMAKADKARYEREMKTYIPPKGET

> sequence 261 E30Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFQDMAKADKARYEREMKTYIPPKGET

> sequence 262 E30H

PDASVNFSEFSKKCSERWKTMSAKEKGKFHDMAKADKARYEREMKTYIPPKGET

> sequence 263 E30N

PDASVNFSEFSKKCSERWKTMSAKEKGKFNDMAKADKARYEREMKTYIPPKGET

> sequence 264 D31N

PDASVNFSEFSKKCSERWKTMSAKEKGKFENMAKADKARYEREMKTYIPPKGET

> sequence 265 D31Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEQMAKADKARYEREMKTYIPPKGET

> sequence 266 M32I

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDIAKADKARYEREMKTYIPPKGET

> sequence 267 M32V

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDVAKADKARYEREMKTYIPPKGET

> sequence 268 K34N

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMANADKARYEREMKTYIPPKGET

> sequence 269 K34Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAQADKARYEREMKTYIPPKGET

> sequence 270 D36N

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKANKARYEREMKTYIPPKGET

> sequence 271 D36Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKAQKARYEREMKTYIPPKGET

> sequence 272 K37N

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADNARYEREMKTYIPPKGET

> sequence 273 K37Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADQARYEREMKTYIPPKGET

> sequence 274 R39H

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKAHYEREMKTYIPPKGET

> sequence 275 R39Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKAQYEREMKTYIPPKGET

> sequence 276 Y40H

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARHEREMKTYIPPKGET

Figure 5b continued

> sequence 277 Y40I

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARIEREMKTYIPPKGET

> sequence 278 E41Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYQREMKTYPKGET

> sequence 279 E41H

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYHREMKTYPKGET

> sequence 280 E41N

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYNREMKTYPKGET

> sequence 281 R42H

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEHMKTYPKGET

> sequence 282 R42Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEQEMKTYPKGET

> sequence 283 E43Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERQMKTYPKGET

> sequence 284 E43H

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERHMKTYPKGET

> sequence 285 E43N

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYERNMKTYPKGET

> sequence 286 M44I

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREIKTYPKGET

> sequence 287 M44V

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREVKTYPKGET

> sequence 288 K45N

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMNTYPKGET

> sequence 289 K45Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMQTYIPKGET

> sequence 290 Y47H

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTHIPKGET

> sequence 291 Y47I

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTIIPKGET

> sequence 292 P49A

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIAPKGET

> sequence 293 P49S

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYISPKGET

> sequence 294 P50A

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPAKGET

27/56

Figure 5b continued

> sequence 295 P50S

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPSKGET

> sequence 296 K51N

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPNGET

> sequence 297 K51Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPQGET

> sequence 298 E53Q

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGQT

> sequence 299 E53H

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGHT

> sequence 300 E53N

PDASVNFSEFSKKCSERWKTMSAKEKGKFEDMAKADKARYEREMKTYIPPKGNT

28/56

Figure 6a**Box A 84 amino acid of HMGB1** *Anopheles gambia* (XP_311154)

Protection against proteolysis
If sequence:

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEK
QRFHEMAEKDKARYELEMQSYVPPKGAV

In bold amino acids sensitive to proteases proteolysis

Figure 6b**Box A 84 amino acid**

Mutant list:

K2N	E24H	F40V	K56N	E71H
K2Q	E24N	R42H	K56Q	E71N
K4N	E25Q	R42Q	R58H	L72I
K4Q	E25H	K43N	R58Q	L72V
D5N	E25N	K43Q	F59I	E73Q
D5Q	K27N	E46Q	F59V	E73H
K7N	K27Q	E46H	E61Q	E73N
K7Q	K28N	E46N	E61H	M74I
P8A	K28Q	R47H	E61N	M74V
P8S	K29N	R47Q	M62I	Y77H
R9H	K29Q	W48Y	M62V	Y77I
R9Q	P31A	W48S	E64Q	P79A
R11H	P31S	K49N	E64H	P79S
R11Q	E32Q	K49Q	E64N	P80A
M12I	E32H	M51I	K65N	P80S
M12V	E32N	M51V	K65Q	K81N
Y15H	E33Q	L52I	D66N	K81Q
Y15I	E33H	L52V	D66Q	
F17I	E33N	D53N	K67N	
F17V	F37I	D53Q	K67Q	
F18I	F37V	K54N	R69H	
F18V	E39Q	K54Q	R69Q	
R23H	E39H	E55Q	Y70H	
R23Q	E39N	E55H	Y70I	
E24Q	F40I	E55N	E71Q	

Figure 6b continued**> SEQUENCE 301 Wild type**

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 302 K2N

GNVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 303 K2Q

GQVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 304 K4N

GKVNNDKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 305 K4Q

GKVQDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 306 D5N

GKVKNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 307 D5Q

GKVQKNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 308 K7N

GKVKDNNPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 309 K7Q

GKVKDNQPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 310 P8A

GKVKDNKARGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 311 P8S

GKVKDNKSRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 312 R9H

GKVKDNKPHGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 313 R9Q

GKVKDNKPQGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 314 R11H

GKVKDNKPRGHMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK

Figure 6b continued**ARYELEMQSYVPPKGAV****> > SEQUENCE 315 R11Q****GKVKDNKPRGQMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 316 M12I****GKVKDNKPRGRITAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKA
RYELEMQSYVPPKGAV****> > SEQUENCE 317 M12V****GKVKDNKPRGRVTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 318 Y15H****GKVKDNKPRGRMTAHAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 319 Y15I****GKVKDNKPRGRMTAIAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKA
RYELEMQSYVPPKGAV****> > SEQUENCE 320 F17I****GKVKDNKPRGRMTAYAIFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKA
RYELEMQSYVPPKGAV****> > SEQUENCE 321 F17V****GKVKDNKPRGRMTAYAVFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****>> SEQUENCE 322 F18I****GKVKDNKPRGRMTAYAFIVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKA
RYELEMQSYVPPKGAV****> > SEQUENCE 323 F18V****GKVKDNKPRGRMTAYAFVVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 324 R23H****GKVKDNKPRGRMTAYAFFVQTCHEEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 325 R23Q****GKVKDNKPRGRMTAYAFFVQTCQEEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 326 E24Q****GKVKDNKPRGRMTAYAFFVQTCRQEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 327 E24H****GKVKDNKPRGRMTAYAFFVQTCRHEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 328 E24N**

Figure 6b continued

GKVKDNKPRGRMTAYAFFVQTCRNEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 329 E25Q

GKVKDNKPRGRMTAYAFFVQTCREQHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 330 E25H

GKVKDNKPRGRMTAYAFFVQTCREHHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 331 E25N

GKVKDNKPRGRMTAYAFFVQTCRENHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 332 K27N

GKVKDNKPRGRMTAYAFFVQTCREEHNKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 333 K27Q

GKVKDNKPRGRMTAYAFFVQTCREEHQKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 334 K28N

GKVKDNKPRGRMTAYAFFVQTCREEHKNKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 335 K28Q

GKVKDNKPRGRMTAYAFFVQTCREEHKQKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 336 K29N

GKVKDNKPRGRMTAYAFFVQTCREEHKKNHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 337 K29Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKQHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 338 P31A

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHAEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 339 P31S

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHSEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 340 E32Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHHPQEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 341 E32H

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHHPHEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK

Figure 6b continued**ARYELEMQSYVPPKGAV****>> SEQUENCE 342 E32N****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPNEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 343 E33Q****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEQQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 344 E33H****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEHQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 345 E33N****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPENQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 346 F37I****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIAEFSRKCAERWKTMLDKEKQRFHEMAEKDKA
RYELEMQSYVPPKGAV****> > SEQUENCE 347 F37V****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIVAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 348 E39Q****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAQFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 349 E39H****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAHFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 350 E39N****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFANFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 351 F40I****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEISRKCAERWKTMLDKEKQRFHEMAEKDKA
RYELEMQSYVPPKGAV****> > SEQUENCE 352 F40V****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEVSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 353 R42H****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSHKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 354 R42Q****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSQKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 355 K43N**

Figure 6b continued

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRNCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 356 K43Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRQCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 357 E46Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAQRWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 358 E46H

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAHRWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 359 E46N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCANRWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 360 R47H

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAEHWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 361 R47Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAEQWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 362 W48Y

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERYKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 363 W48S

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERSKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 364 K49N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWNTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 365 K49Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWQTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 366 M51I

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTILDKEKQRFHEMAEKDKA
RYELEMQSYVPPKGAV

> > SEQUENCE 367 M51V

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTVLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > > SEQUENCE 368 L52I

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKA
RYELEMQSYVPPKGAV

Figure 6b continued

> > SEQUENCE 369 L52V

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMTVDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 370 D53N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLNKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 371 D53Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLQKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 372 K54N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDNEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 373 K54Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDQEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 374 E55Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKQKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 375 E55H

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKHKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 376 E55N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKNKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 377 K56N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKENQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 378 K56Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEQKQRFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 379 R58H

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQHFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 380 R58Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQQFHEMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 381 F59I

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRIHEMAEKDKA
RYELEMQSYVPPKGAV

> > SEQUENCE 382 F59V

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRVHEMAEKDK
ARYELEMQSYVPPKGAV

Figure 6b continued**> > SEQUENCE 383 E61Q**

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHQMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 384 E61H

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHHMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 385 E61N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHNMAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 386 M62I

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEIAEKDKA
RYELEMQSYVPPKGAV

> > SEQUENCE 387 M62V

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEVAEKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 388 E64Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAQKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 389 E64H

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAHKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 390 E64N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMANKDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 391 K65N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAENDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 392 K65Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEQDK
ARYELEMQSYVPPKGAV

> > SEQUENCE 393 D66N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKNK
ARYELEMQSYVPPKGAV

> > SEQUENCE 394 D66Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKQK
ARYELEMQSYVPPKGAV

> > SEQUENCE 395 K67N

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDN
ARYELEMQSYVPPKGAV

> > SEQUENCE 396 K67Q

GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDQ

Figure 6b continued**ARYELEMQSYVPPKGAV****> > SEQUENCE 397 R69H****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
AHYELEMQSYVPPKGAV****> > SEQUENCE 398 R69Q****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
AQYELEMQSYVPPKGAV****> > SEQUENCE 399 Y70H****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARHELEMQSYVPPKGAV****> > SEQUENCE 400 Y70I****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARIELEMQSYV****> > SEQUENCE 401 E71Q****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYQLEMQSYVPPKGAV****> > SEQUENCE 402 E71H****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYHLEMQSYVPPKGAV****> > SEQUENCE 403 E71N****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYNLEMQSYVPPKGAV****> > SEQUENCE 404 L72I****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPKGAV****> > SEQUENCE 405 L72V****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYEVEMQSYVPPKGAV****> > SEQUENCE 406 E73Q****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELQMMSYVPPKGAV****> > SEQUENCE 407 E73H****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELHMQSYVPPKGAV****> > SEQUENCE 408 E73N****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELNMMSYVPPKGAV****> > SEQUENCE 409 M74I****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEIQSYVPPKGAV****>> SEQUENCE 410 M74V****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK**

Figure 6b continued**ARYELEVQSYVPPKGAV****> > SEQUENCE 411 Y77H****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSHVPPKGAV****> > SEQUENCE 412 Y77I****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSIVPPKGAV****> > SEQUENCE 413 P79A****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVAPKGAV****> > SEQUENCE 414 P79S****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVSPKGAV****> > SEQUENCE 415 P80A****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPAKGAV****> > SEQUENCE 416 P80S****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPSKGAV****> > SEQUENCE 417 K81N****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPNGAV****> > SEQUENCE 418 K81Q****GKVKDNKPRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDK
ARYELEMQSYVPPQGAV**

38/56

Figure 7a**Box A 77 amino acid of HMGB1** *Anopheles gambia* (XP_311154)# Protection against proteolysis
If sequence:**PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMA
EKDKARYELEMQSYVPPKGAV**

In bold amino acids sensitive to proteases proteolysis

Figure 7b**Box A 77 amino acid of HMGB1** *Anopheles gambia* (XP_311154)

Mutant list:

P1A	E26N	R51Q	P73A
P1S	F30I	F52I	P73S
R2H	F30V	F52V	K74N
R2Q	E32Q	E54Q	K74Q
R4H	E32H	E54H	
R4Q	E32N	E54N	
M5I	F33I	M55I	
M5V	F33V	M55V	
Y8H	R35H	E57Q	
Y8I	R35Q	E57H	
F10I	K36N	E57N	
F10V	K36Q	K58N	
F11I	E39Q	K58Q	
F11V	E39H	D59N	
R16H	E39N	D59Q	
R16Q	R40H	K60N	
E17Q	R40Q	K60Q	
E17H	W41Y	R62H	
E17N	W41S	R62Q	
E18Q	K42N	Y63H	
E18H	K42Q	Y63I	
E18N	M44I	E64Q	
K20N	M44V	E64H	
K20Q	L45I	E64N	
K21N	L45V	L65I	
K21Q	D46N	L65V	
K22N	D46Q	E66Q	
K22Q	K47N	E66H	
P24A	K47Q	E66N	
P24S	E48Q	M67I	
E25Q	E48H	M67V	
E25H	E48N	Y70H	
E25N	K49N	Y70I	
E26Q	K49Q	P72A	
E26H	R51H	P72S	

39/56

Figure 7b continued**> SEQUENCE 419 Wild type**

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> > SEQUENCE 420 P1A

ARGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 421 P1S

SRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 422 R2H

PHGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 423 R2Q

PQGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 424 R4H

PRGHMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 425 R4Q

PRGQMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 426 M5I

PRGRITAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 427 M5V

PRGRVTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 428 Y8H

PRGRMTAHAFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 429 Y8I

PRGRMTAIAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 430 F10I

PRGRMTAYAIFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 431 F10V

PRGRMTAYAVFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

Figure 7b continued**> SEQUENCE 432 F11I**

PRGRMTAYAFIVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
SYVPPKGAV

> SEQUENCE 433 F11V

PRGRMTAYAFVVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 434 R16H

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 435 R16Q

PRGRMTAYAFFVQTCQEEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 436 E17Q

PRGRMTAYAFFVQTCRQEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 437 E17H

PRGRMTAYAFFVQTCRHEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 438 E17N

PRGRMTAYAFFVQTCRNEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 439 E18Q

PRGRMTAYAFFVQTCREQHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 440 E18H

PRGRMTAYAFFVQTCREHHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 441 E18N

PRGRMTAYAFFVQTCRENHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 442 K20N

PRGRMTAYAFFVQTCREEHNKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 443 K20Q

PRGRMTAYAFFVQTCREEHQKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 444 K21N

PRGRMTAYAFFVQTCREEHKNKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 445 K21Q

PRGRMTAYAFFVQTCREEHKQKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

Figure 7b continued**> SEQUENCE 446 K22N**

PRGRMTAYAFFVQTCREEHKKNHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 447 K22Q

PRGRMTAYAFFVQTCREEHKKQHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 448 P24A

PRGRMTAYAFFVQTCREEHKKKHAEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 449 P24S

PRGRMTAYAFFVQTCREEHKKKHSEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 450 E25Q

PRGRMTAYAFFVQTCREEHKKKHPQEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 451 E25H

PRGRMTAYAFFVQTCREEHKKKHPHEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 452 E25N

PRGRMTAYAFFVQTCREEHKKKHPNEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 453 E26Q

PRGRMTAYAFFVQTCREEHKKKHPEQQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 454 E26H

PRGRMTAYAFFVQTCREEHKKKHPEHQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 455 E26N

PRGRMTAYAFFVQTCREEHKKKHPENQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 456 F30I

PRGRMTAYAFFVQTCREEHKKKHPEEQVIAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 457 F30V

PRGRMTAYAFFVQTCREEHKKKHPEEQVIVAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 458 E32Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAQFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 459 E32H

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAHFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

Figure 7b continued**> SEQUENCE 460 E32N**

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFANFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 461 F33I

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEISRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 462 F33V

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEVSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 463 R35H

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSHKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 464 R35Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSQKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 465 K36N

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRNCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 466 K36Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRQCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 467 E39Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAQRWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 468 E39H

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAHRWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 469 E39N

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCANRWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 470 R40H

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAEHWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 471 R40Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAEQWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 472 W41Y

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERYKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 473 W41S

Figure 7b continued

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERSKTMLDKEKQRFHEMAEKDKARYELEMQ
SYVPPKGAV

> SEQUENCE 474 K42N

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWNTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 475 K42Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWQTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 476 M44I

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTILDKEKQRFHEMAEKDKARYELEMQ
SYVPPKGAV

> SEQUENCE 477M44V

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTVLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 478 L45I

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQ
SYVPPKGAV

> SEQUENCE 479 L45V

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 480 D46N

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLNKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 481 D46Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLQKEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 482 K47N

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDNEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 483 K47Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDQEKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 484 E48Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKQKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 485 E48H

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKHKQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 486 E48N

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKNKQRFHEMAEKDKARYELEM
QSYVPPKGAV

Figure 7b continued

> SEQUENCE 487 K49N

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKENQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 488 K49Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEQQRFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 489 R51H

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQHFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 490 R51Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQQFHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 491 F52I

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRIHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 492 F52V

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRVHEMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 493 E54Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHQMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 494 E54H

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHHMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 495 E54N

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHNMAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 496 M55I

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEIAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 497 M55V

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEVAEKDKARYELEM
QSYVPPKGAV

> SEQUENCE 498 E57Q

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAQKDKARYELEM
QSYVPPKGAV

> SEQUENCE 499 E57H

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAHKDKARYELEM
QSYVPPKGAV

> SEQUENCE 500 E57N

PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMANKDKARYELEM
QSYVPPKGAV

Figure 7b continued

> SEQUENCE 501 K58N
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAENDKARYELEM
QSYVPPKGAV

> SEQUENCE 502 K58Q
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEQDKARYELEM
QSYVPPKGAV

> SEQUENCE 503 D59N
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKNKARYELEM
QSYVPPKGAV

> SEQUENCE 504 D59Q
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKQKARYELEM
QSYVPPKGAV

> SEQUENCE 505 K60N
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDNARYELEM
QSYVPPKGAV

> SEQUENCE 506 K60Q
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDQARYELEM
QSYVPPKGAV

> SEQUENCE 507 R62H
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKAHYELEM
QSYVPPKGAV

> SEQUENCE 508 R62Q
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKAQYELEM
QSYVPPKGAV

> SEQUENCE 509 Y63H
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARHELEM
QSYVPPKGAV

> SEQUENCE 510 Y63I
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARIELEM
QSYVPPKGAV

> SEQUENCE 511 E64Q
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYQLEM
QSYVPPKGAV

> SEQUENCE 512 E64H
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYHLEM
QSYVPPKGAV

> SEQUENCE 513 E64N
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYNLEM
QSYVPPKGAV

> SEQUENCE 514 L65I
PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYEIMQ

Figure 7b continued**SYVPPKGAV****> SEQUENCE 515 L65V****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYEVEM
QSYVPPKGAV****> SEQUENCE 516 E66Q****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELQM
QSYVPPKGAV****> SEQUENCE 517 E66H****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELHM
QSYVPPKGAV****> SEQUENCE 518 E66N****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELNM
QSYVPPKGAV****> SEQUENCE 519 M67I****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEIQ
SYVPPKGAV****> SEQUENCE 520 M67V****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEV
QSYVPPKGAV****> SEQUENCE 521 Y70H****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSHVPPKGAV****> SEQUENCE 523 Y70I****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSIVPPKGAV****> SEQUENCE 524 P72A****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVAPKGAV****> SEQUENCE 525 P72S****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVSPKGAV****> SEQUENCE 526 P73A****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPAKGAV****> SEQUENCE 527 P73S****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPSKGAV****> SEQUENCE 528 K74N****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM
QSYVPPNGAV****> SEQUENCE 529 K74Q****PRGRMTAYAFFVQTCREEHKKKHPEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEM**

Figure 7b continued

QSYVPPQGAV

48/56

Figure 8a**Box A 54 amino acid of HMGB1** *Anopheles gambia* (XP_311154)# Protection against proteolysis
If sequence:**PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV**

In bold amino acids sensitive to proteases proteolysis

Figure 8b**Box A 54 amino acid of HMGB1** *Anopheles gambia* (XP_311154)

Mutant list:

P1A	K24N	E43Q
P1S	K24Q	E43H
E2Q	E25Q	E43N
E2H	E25H	M44I
E2N	E25N	M44V
E3Q	K26N	Y47H
E3H	K26Q	Y47I
E3N	R28H	P49A
F7I	R28Q	P49S
F7V	F29I	P50A
E9Q	F29V	P50S
E9H	E31Q	K51N
E9N	E31H	K51Q
F10I	E31N	
F10V	M32I	
R12H	M32V	
R12Q	E34Q	
K13N	E34H	
K13Q	E34N	
E16Q	K35N	
E16H	K35Q	
E16N	D36N	
R17H	D36Q	
R17Q	K37N	
W18Y	K37Q	
W18S	R39H	
K19N	R39Q	
K19Q	Y40H	
M21I	Y40I	
M21V	E41Q	
L22I	E41H	
L22V	E41N	
D23N	L42I	
D23Q	L42V	

Figure 8b continued

> SEQUENCE 530 Wild type

5 PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 531 P1A

AEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

10 > SEQUENCE 532 P1S

SEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 533 E2Q

PQEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

15 > SEQUENCE 534 E2H

PHEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 535 E2N

20 PNEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 536 E3Q

PEQQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 537 E3H

25 PEHQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 538 E3N

30 PENQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 539 F7I

PEEQVIIAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 540 F7V

35 PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 541 E9Q

PEEQVIFAQFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

40 > SEQUENCE 542 E9H

PEEQVIFAHFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 543 E9N

45 PEEQVIFANFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 544 F10I

PEEQVIFAEISRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 545 F10V

50 PEEQVIFAEVSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 546 R12H

PEEQVIFAEFSHKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

55 > SEQUENCE 547 R12Q

PEEQVIFAEFSQKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

Figure 8b continued

> SEQUENCE 548 K13N
PEEQVIFAEFSRNCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

5 > SEQUENCE 549 K13Q
PEEQVIFAEFSRQCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 550 E16Q
PEEQVIFAEFSRKCAQRWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

10 > SEQUENCE 551 E16H
PEEQVIFAEFSRKCAHRWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 552 E16N
PEEQVIFAEFSRKCANRWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

15 > SEQUENCE 553 R17H
PEEQVIFAEFSRKCAEHWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 554 R17Q
PEEQVIFAEFSRKCAEQWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

20 > SEQUENCE 555 W18Y
PEEQVIFAEFSRKCAERYKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 556 W18S
PEEQVIFAEFSRKCAERSKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 557 K19N
PEEQVIFAEFSRKCAERWNTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

30 > SEQUENCE 558 K19Q
PEEQVIFAEFSRKCAERWQTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 559 M21I
PEEQVIFAEFSRKCAERWKTILDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

35 > SEQUENCE 560 M21V
PEEQVIFAEFSRKCAERWKTVLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 561 L22I
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

40 > SEQUENCE 562 L22V
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

45 > SEQUENCE 563 D23N
PEEQVIFAEFSRKCAERWKTMLNKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 564 D23Q
PEEQVIFAEFSRKCAERWKTMLQKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

50 > SEQUENCE 565 K24N
PEEQVIFAEFSRKCAERWKTMLDNEKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 566 K24Q
PEEQVIFAEFSRKCAERWKTMLDQEKQRFHEMAEKDKARYELEMQSYVPPKGAV

55 > SEQUENCE 567 E25Q

Figure 8b continued

PEEQVIFAEFSRKCAERWKTMLDKQKQRFHEMAEKDKARYELEMQSYVPPKGAV

5 > SEQUENCE 568 E25H
PEEQVIFAEFSRKCAERWKTMLDKHKQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 569 E25N
PEEQVIFAEFSRKCAERWKTMLDKNKQRFHEMAEKDKARYELEMQSYVPPKGAV

10 > SEQUENCE 570 K26N
PEEQVIFAEFSRKCAERWKTMLDKENQRFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 571 K26Q
PEEQVIFAEFSRKCAERWKTMLDKEQQRFHEMAEKDKARYELEMQSYVPPKGAV

15 > SEQUENCE 572 R28H
PEEQVIFAEFSRKCAERWKTMLDKEKQHFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 573 R28Q
20 PEEQVIFAEFSRKCAERWKTMLDKEKQQFHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 574 F29I
PEEQVIFAEFSRKCAERWKTMLDKEKQRIHEMAEKDKARYELEMQSYVPPKGAV

25 > SEQUENCE 575 F29V
PEEQVIFAEFSRKCAERWKTMLDKEKQRVHEMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 576 E31Q
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHQMAEKDKARYELEMQSYVPPKGAV

30 > SEQUENCE 577 E31H
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHHMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 578 E31N
35 PEEQVIFAEFSRKCAERWKTMLDKEKQRFHNMAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 579 M32I
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEIAEKDKARYELEMQSYVPPKGAV

40 > SEQUENCE 580 M32V
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEVAEKDKARYELEMQSYVPPKGAV

> SEQUENCE 581 E34Q
45 PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAQKDKARYELEMQSYVPPKGAV

> SEQUENCE 582 E34H
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAHKDKARYELEMQSYVPPKGAV

> SEQUENCE 583 E34N
50 PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMANKDKARYELEMQSYVPPKGAV

> SEQUENCE 584 K35N
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAENDKARYELEMQSYVPPKGAV

55 > SEQUENCE 585 K35Q
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEQDKARYELEMQSYVPPKGAV

> SEQUENCE 586 D36N

Figure 8b continued

PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKNKARYELEMQSYVPPKGAV

5 > SEQUENCE 587 D36Q
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKQKARYELEMQSYVPPKGAV

> SEQUENCE 588 K37N
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDNARYELEMQSYVPPKGAV

10 > SEQUENCE 590 K37Q
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDQARYELEMQSYVPPKGAV

> SEQUENCE 591 R39H
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKAHYELEMQSYVPPKGAV

15 > SEQUENCE 592 R39Q
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKAQYELEMQSYVPPKGAV

> SEQUENCE 593 Y40H
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARHELEMQSYVPPKGAV

20 > SEQUENCE 594 Y40I
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARIELEMQSYVPPKGAV

> SEQUENCE 595 E41Q
25 PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYQLEMQSYVPPKGAV

> SEQUENCE 596 E41H
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYHLEMQSYVPPKGAV

30 > SEQUENCE 597 E41N
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYNLEMQSYVPPKGAV

> SEQUENCE 598 L42I
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPKGAV

35 > SEQUENCE 599 L42V
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYEVEMQSYVPPKGAV

> SEQUENCE 600 E43Q
40 PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELQMMSYVPPKGAV

> SEQUENCE 601 E43H
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELHMQSYVPPKGAV

45 > SEQUENCE 602 E43N
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELNMMSYVPPKGAV

> SEQUENCE 603 M44I
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEIQSYVPPKGAV

50 > SEQUENCE 604 M44V
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEVQSYVPPKGAV

> SEQUENCE 605 Y47H
55 PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSHVPPKGAV

> SEQUENCE 606 Y47I
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSIVPPKGAV

Figure 8b continued

> SEQUENCE 607 P49A

PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVAPKGAV

5

> SEQUENCE 608 P49S

PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVSPKGAV

> SEQUENCE 609 P50A

PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPAKGAV

10

> SEQUENCE 610 P50S

PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPSKGAV

> SEQUENCE 611 K51N

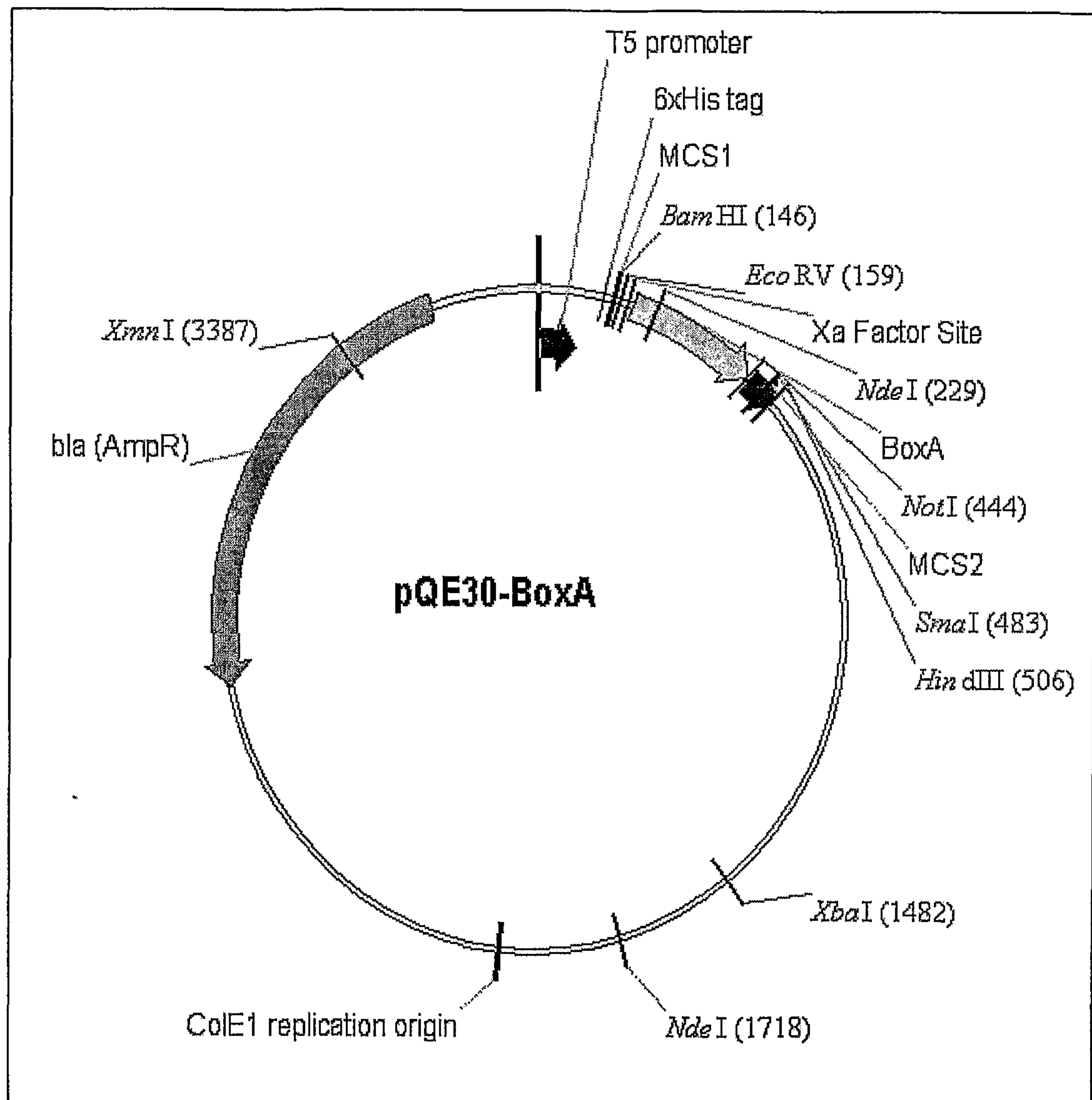
15

PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPNGAV

> SEQUENCE 612 K51Q

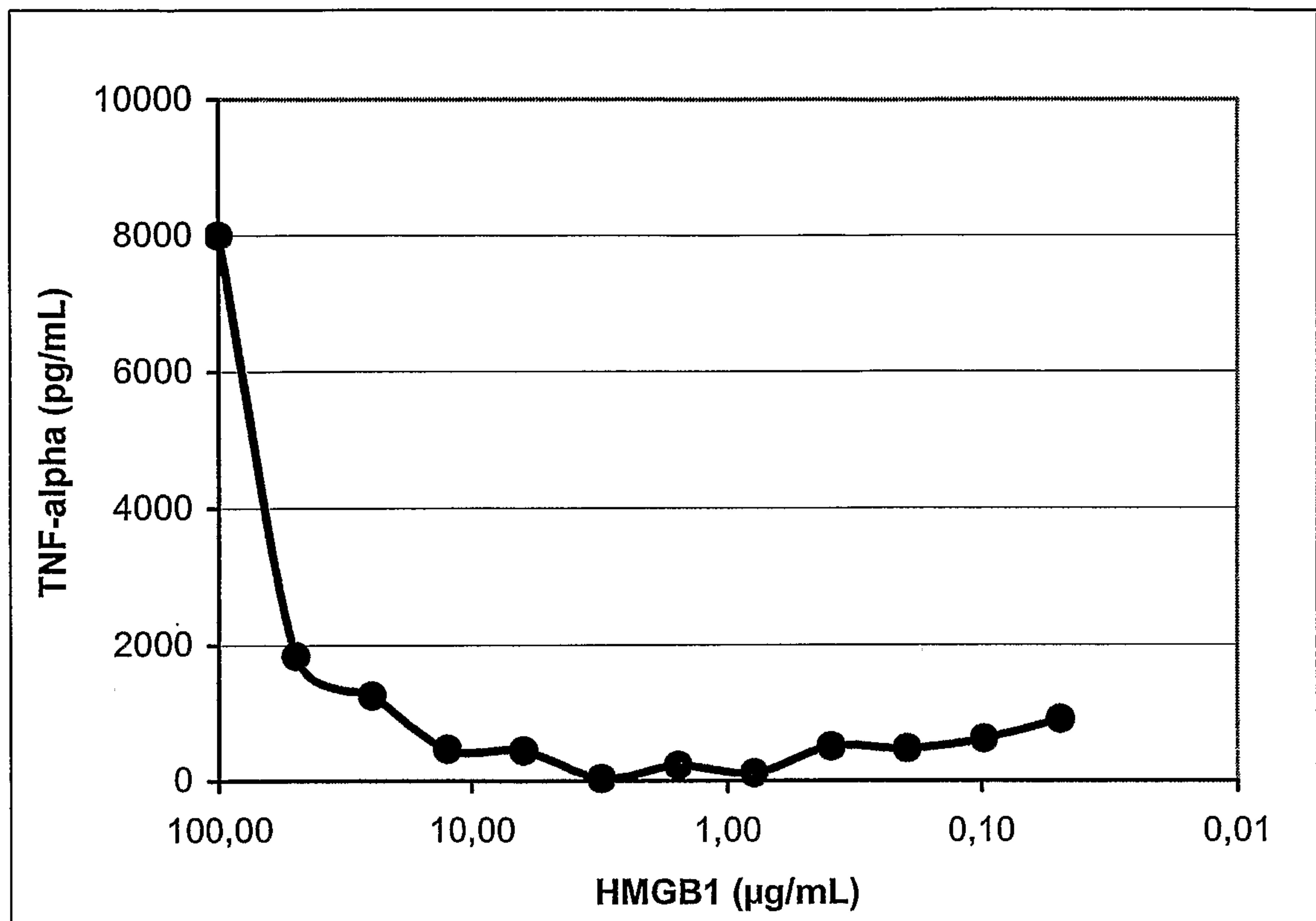
PEEQVIFAEFSRKCAERWKTMLDKEKQRFHEMAEKDKARYELEMQSYVPPQGAV

Figure 9



55/56

Figure 10



56/56

Figure 11

