

Novel alpha-1-microglobulin derived proteins and their use

Field of the invention

The present invention relates to modified variants of human alpha-1-microglobulin protein with improved properties and the use of such variants in medical treatment and diagnostics. The inventors have surprisingly found that introduction of specific amino acid substitutions and / or addition of specific N-terminal extensions confer improved properties to alpha-1-microglobulin with regard to stability, solubility, and binding of heme.

10 Background of the invention

A1M (α_1 -microglobulin) is a low molecular weight protein with an extracellular tissue-cleaning function (Ekström et al., 1977; Åkerström and Gram, 2014). It is present in all tissues and organs in fish, birds, rodents, mammals and other vertebrates. A1M is synthesized mainly in the liver but also at a lower rate in most other cells in the body. It is encoded by the α_1 -microglobulin-bikunin precursor gene (*AMBP*) and translated in all cells and species as a continuous peptide precursor together with another protein, bikunin (Kaumeyer et al., 1986). However, the two proteins are separated by protease cleavage, processed and secreted into the blood as two different proteins with different functions (Lindqvist et al., 1992; Bratt et al., 1993). The reason for the ubiquitous co-synthesis of A1M and bikunin is still unknown. In the blood, about 50% of A1M is found in free, monomeric form, and the remaining 50% as high-molecular weight complexes covalently bound with immunoglobulin A, albumin and prothrombin (Berggård et al., 1997).

25 A1M is a one-domain, 183-amino acid, glycosylated protein. The crystal structure of a large fragment of A1M expressed in *E.coli* was recently published (Meining and Skerra, 2012). Based on its structure, A1M belongs to a protein family, Lipocalins, with 50 or more members from animals, plants and bacteria (Flower, 1996; Åkerström et al., 2006). The lipocalins have a common three-dimensional structure which consists of eight antiparallel β -strands forming a barrel with one closed end (bottom) and an open end (top). The barrel functions as a pocket for hydrophobic ligands in most lipocalins. Four loops (loop 1-4), which make up the rim of the open end of the barrel, vary highly in length and composition between the various lipocalins. In A1M, a handful of amino acid side-groups located on these loops, or on the inside of the pocket, have been shown to be important for the identified functions of the protein. Thus, a free cysteine, C34, located on a short helix on loop 1, has a negative reduction potential and gives

A1M reductase properties (Allhorn et al., 2005). Two tyrosine residues, Y22 and 132, were shown to be covalently modified by radical oxidation products in vitro (Åkerström et al., 2007). Four lysine residues, K69, 92, 118 and 130, regulate the reductase activity (Allhorn et al., 2005), influence the binding of free heme groups (Rutardottir et al., 2015), and are covalently modified on A1M purified from human urine and amniotic fluid with low molecular weight yellow-brown, heterogeneous substances (Berggård et al., 1999; Sala et al., 2004).

Employing the reductase activity, radical scavenging and heme-binding properties, A1M acts as an antioxidant that protects cells and tissues from oxidative damage. A1M was shown to protect in vitro blood cell cultures, placenta tissue and skin against oxidative damage from hemoglobin, heme and reactive oxygen species (ROS) (Olsson et al., 2008; May et al., 2011; Olsson et al., PLoS One 2011; Olsson et al., ARS 2012). A1M also showed in vivo protective effects in rats and rabbits against placenta and kidney tissue damage after hemoglobin infusion (Sverrisson et al., 2014; Nääv et al., 2015). In a series of reports, hemoglobin and oxidative stress were shown to be involved in the pathogenesis of preeclampsia, a serious complication of pregnancy (Centlow et al., 2008; Hansson et al., 2013), and the levels of A1M-mRNA and protein in liver, placenta and plasma were elevated in pregnant women with preeclampsia (Olsson et al., 2010; Anderson et al., 2011). Therefore, it was suggested that A1M may be employed as a therapeutic agent to treat pregnant women with preeclampsia to ameliorate the oxidative damage and thus the clinical symptoms of the disease (Olsson et al., 2010).

A production process of recombinant human A1M was developed in *E.coli* and shown to possess the reductase, radical-binding and heme-binding properties (Kwasek et al., Allhorn et al., 2002; 2005; Åkerström et al., 2007). Indeed, this recombinant A1M variant showed in vivo therapeutic effects in a sheep model of preeclampsia (Wester-Rosenlöf et al., 2014). However, although functional, *E.coli*-expressed recombinant A1M lacks glycosylation, and has poor solubility and stability compared to human A1M purified from urine or plasma. The lack of stability and solubility of the protein limits its use as a drug for human use, mainly due to difficulties to obtain highly concentrated solutions and long-term storage conditions in buffers at physiological pH and salt conditions.

Accordingly, there is a need for developing a protein with structural similarities with human A1M, but with improved properties regarding stability and solubility. Moreover, the protein should be relatively easy to prepare in amounts suitable for therapeutic use.

5 Detailed description of the invention

The present invention relates to novel alpha-1-microglobulin proteins with improved stability and solubility profiles.

10 Site-directed and additive mutagenesis was used to engineer A1M-species with retained, or possibly enhanced functional properties, but with improved protein stability and solubility that allow long-term storage at high concentrations in physiological buffers.

As it appears from the Examples herein, four lines of reasoning were followed when selecting the positions and identities of mutated amino acid side-groups:

- 1) Animal homologues from a variety of species with different expected environmental pressure in terms of oxidative stress, temperature, oxygen pressure were expressed (N=12);
- 2) Single amino acid substitutions that occur frequently among the 56 sequenced A1M-homologues at positions located in loops 1-4 or the interior surface of the hydrophobic pocket, were introduced into the human gene construct and expressed (N=3); and
- 3) Addition or removal of favourably located lysyl or tyrosyl residues, based on the hypothesis that these may influence pKa of the C35 thiolyl (Allhorn et al., 2005) or serve as radical-trapping sites (Berggård et al., 1999; Sala et al., 2004; Åkerström et al. 2007) (N=5); 4).

In addition, the influence of N-terminal, charged and hydrophilic extensions were tested on some A1M-variants. The rationale behind the design of the tested N-terminal extensions was to add 1) a tag for purification (e.g. His-tag), 2) a linker to separate the tag from the core of the A1M protein, 3) several (1-5) charged amino acid side-groups conferring hydrophilic properties to the protein in order to gain maximal stability and solubility in water-solutions, 4) without compromising the physiological functions of A1M.

The project was divided into three major phases:

- 35 Phase I) Expression of the 27 A1M-variants described above followed by analysis of solubility, stability and function;

Phase II) Design, expression and analysis of a few A1M-variants with expected optimal properties based on the outcome of phase 1; and

Phase III) Design, expression and analysis of non-mutated wildtype (wt)-A1M and the most successful mutated A1M-variant equipped with or without N-terminal, charged

5 and hydrophilic extensions.

As will be explained in more details herein, the present invention provides an A1M-derived protein having the amino acid sequence of formula I:

10 X¹-X²-X³-X⁴-X⁵-X⁶-X⁷-X⁸-X⁹-X¹⁰-X¹¹-X¹²-X¹³-X¹⁴-GPVPTPPDN IQVQENF-X¹⁵-IS
RIYGKWNLA IGSTCPWLKK I-X¹⁶-DRMTVSTL VLGEGATEAE ISMTST-X¹⁷-
WRK GVCEETSGAY EKTDTDGKFL YHKS KW-X¹⁸-ITM ESYVVHTNYD EY-
AIFLTCKF SRHHGPTITA KLYGRAPQLR ETLLQDFRVV AQGVGIPEDS IFT-
MADRGECE VPGEQEPEPI LIPR (formula I)

15

wherein at least one X is present and (in parentheses are suggestions for further substitutions)

X¹ is absent or represents Met or N-formyl Met;

X² is absent or represents His;

20 X³ is absent or represents His;

X⁴ is absent or represents His;

X⁵ is absent or represents His;

X⁶ is absent or represents His;

X⁷ is absent or represents His;

25 X⁸ is absent or represents His;

X⁹ is absent or represents His;

X¹⁰ is absent or selected from Asp and Glu, Lys and Arg;

X¹¹ is absent or selected from Asp and Glu, Lys and Arg;

X¹² is absent or selected from Asp and Glu, Lys and Arg;

30 X¹³ is absent or selected from Asp and Glu, Lys and Arg;

X¹⁴ is absent or represents Lys Glu, Asp or Arg or Met or N-formyl Met;

X¹⁵ represents Asp or Asn or Glu;

X¹⁶ represents Met or Lys or Arg;

X¹⁷ represents Arg or His or Lys;

35 X¹⁸ represents Asp or Asn or Glu;

or a pharmaceutically acceptable salt thereof,

with the proviso that when all X¹-X¹⁴ are absent, X¹⁵ represents Asn, X¹⁶ represents Met, and X¹⁷ represents Arg, then X¹⁸ cannot represent Asn.

- 5 The invention also provides a derivative of A1M, i.e. X¹-X¹⁴-A1M, wherein A1M may be any A1M obtained from the species mentioned in Table 2 (i.e. human, mouse, naked mole-rat, frog, chicken, rabbit, squirrel monkey, walrus, manatee, plaice and orangutan). The present inventors have found that inclusion of X¹-X¹⁴ imparts improved properties at least to human A1M, and accordingly, it is contemplated that this start sequence also can impart important improved properties to A1M from other species or to species-recombinant A1M.

The present invention also provides an A1M-derived protein having the amino acid sequence of formula II:

15

X¹-X²-X³-X⁴-X⁵-X⁶-X⁷-X⁸-X⁹-X¹⁰-X¹¹-X¹²-X¹³-X¹⁴-GPVPTPPDN IQVQENF-X¹⁵-IS
RIYGKWWYNLA IGSTCPWLKK I-X¹⁶-DRMTVSTL VLGEGATEAE ISMTST-X¹⁷-
WRK GVCEETSGAY EKTDTDGKFL YHKS KW-X¹⁸-ITM ESYVVHTNYD EY-
AIFLTKKF SRHHGPTITA KLYGRAPQLR ETLLQDFRVV AQGVGIPEDS IFT-
20 MADRGEC VPGEQEPEPI (formula II)

wherein at least one X is present and (in parentheses are suggestions for further substitutions)

X¹ is absent or represents Met or N-formyl Met;

X² is absent or represents His;

25

X³ is absent or represents His;

X⁴ is absent or represents His;

X⁵ is absent or represents His;

X⁶ is absent or represents His;

X⁷ is absent or represents His;

30

X⁸ is absent or represents His;

X⁹ is absent or represents His;

X¹⁰ is absent or selected from Asp and Glu, Lys and Arg;

X¹¹ is absent or selected from Asp and Glu, Lys and Arg;

X¹² is absent or selected from Asp and Glu, Lys and Arg;

35

X¹³ is absent or selected from Asp and Glu, Lys and Arg;

X¹⁴ is absent or represents Lys or Glu, Asp or Arg or Met or N-formyl Met;

X¹⁵ represents Asp or Asn or Glu;

X¹⁶ represents Met or Lys or Arg;

X¹⁷ represents Arg or His or Lys;

X¹⁸ represents Asp or Asn or Glu;

5 or a pharmaceutically acceptable salt thereof,

with the proviso that when all X¹-X¹⁴ are absent, X¹⁵ represents Asn, X¹⁶ represents Met, and X¹⁷ represents Arg, then X¹⁸ cannot represent Asn.

10 The invention also provides a derivative of A1M, i.e. X¹-X¹⁴-A1M, wherein A1M may be any A1M obtained from the species mentioned in Table 2 (i.e. human, mouse, naked mole-rat, frog, chicken, rabbit, squirrel monkey, walrus, manatee, plaice and orangutan), wherein A1M is truncated C-terminally, so that the C-terminal tetrapeptide sequence LIPR does not form part of the protein.

15

As it appears from the Examples herein, the present inventors have found that:

- i) The initial sequence X¹-X¹⁴ seems to impart improved properties to A1M,
- ii) Point mutation M41K, R66H or N17,96D of the A1M molecule imparts improved stability with maintained function of A1M,
- 20 iii) Mutations (M41K + R66H), (M41K + N17,96D), (R66H + N17,96D), and/or (M41K + R66H + N17,96D) show increased solubility and/or stability with maintained function,
- iv) Mutation (R66H + N17,96D) had best overall performance in the experiments performed,
- 25 v) A1M with mutations (R66H + N17,96D) and initial sequences MHHHHHHHHGSGGIEGR (M8H5GIEGR); MHHHHHHHHDDDDK (M8H4DK), MHHHHHHHHDDDDK (M6H4DK) or MHHHHHHHH (M8H) as N-terminal sequences have been tested, and the protein variants with
- 30 M8H4DK as N-terminal sequence showed higher solubility and/or stability compared with the other N-terminal extensions.
- vi) Truncation of the C-terminal of A1M seems to impart improved heme binding and degradation.

35 It is interesting to note that this N-terminal extension sequence resembles a His-tag with an enterokinase cleavage site, but it is without the ability to cleave the His-tag from

A1M as the amino acid Ala is not included. The presence of Ala in a DDDKA indicate the enterokinase cleavage site. Thus, it is contemplated that the N-terminal sequence itself imparts improved protein stability and solubility properties to A1M when the repeated His-residues are followed by five charged amino acids.

5

Based on these observations, it is contemplated that variation of an A1M protein along the lines indicated above will provide proteins with A1M functionality, but with improved characteristics regarding stability and/or solubility.

- 10 Thus, the present invention relates to all possible combinations of A1M containing X¹ – X¹⁸ as described above.

More specifically, the following A1M derived proteins are within the scope of the present invention, such that the all proteins may be full-length corresponding to human

- 15 wild type A1M; or may be truncated C-terminally, i.e. without LIPR:

P017240DK1

Position in hA1M		Compound			
Tag	Mutations	X1	X2	X3	X4
Broadest claim		Met/Absent	His/absent	His/absent	His/absent
6-His	'in sequence' mutations can be any	(f)Met	His	His	His
8-His	in sequence' mutations can be any	(f)Met	His	His	His
No tag;	in sequence' mutations can be any	Absent	Absent	Absent	Absent
Any tag	M41K	Met/Absent	His/absent	His/absent	His/absent
Any tag	N17,96D	Met/Absent	His/absent	His/absent	His/absent
No tag	N17D	Absent	Absent	Absent	Absent
No tag	N17,96D	Absent	Absent	Absent	Absent
No tag	N96D	Absent	Absent	Absent	Absent
No tag	N17,96D, R66H	Absent	Absent	Absent	Absent
No tag	M41K	Absent	Absent	Absent	Absent
No tag	R66H	Absent	Absent	Absent	Absent
No tag	N17,96D, M41K, R66H	Absent	Absent	Absent	Absent
No tag	N17D, R66H	Absent	Absent	Absent	Absent
No tag	R66H, N96D	Absent	Absent	Absent	Absent
No tag	M41K, R66H	Absent	Absent	Absent	Absent
No tag	N17,96D, M41K	Absent	Absent	Absent	Absent
No tag	N17D, M41K	Absent	Absent	Absent	Absent
No tag	M41K, N96D	Absent	Absent	Absent	Absent
Any tag	N17,96D, R66H	Met/Absent	His/absent	His/absent	His/absent
No tag	N17D, M41K, R66H	Absent	Absent	Absent	Absent
No tag	M41K, R66H, N96D	Absent	Absent	Absent	Absent
6His	N17D	(f)Met	His	His	His
6His	N17,96D	(f)Met	His	His	His
6His	N96D	(f)Met	His	His	His
6His	N17,96D, R66H	(f)Met	His	His	His

P017240DK1

Position in hA1M						
Tag	Mutations	Compound	X1	X2	X3	X4
6His	M41K	27	(f)Met	His	His	His
6His	R66H	28	(f)Met	His	His	His
6His	N17,96D, M41K, R66H	29	(f)Met	His	His	His
6His	N17D, R66H	30	(f)Met	His	His	His
6His	R66H, N96D	31	(f)Met	His	His	His
6His	M41K, R66H	32	(f)Met	His	His	His
6His	N17,96D, M41K	33	(f)Met	His	His	His
6His	N17D, M41K	34	(f)Met	His	His	His
6His	M41K, N96D	35	(f)Met	His	His	His
6His	N17D, M41K, R66H	36	(f)Met	His	His	His
6His	M41K, R66H, N96D	37	(f)Met	His	His	His
8His	N17D	38	(f)Met	His	His	His
8His	N17,96D	39	(f)Met	His	His	His
8His	N96D	40	(f)Met	His	His	His
8His	N17,96D, R66H	41	(f)Met	His	His	His
8His	M41K	42	(f)Met	His	His	His
8His	R66H	43	(f)Met	His	His	His
8His	N17,96D, M41K, R66H	44	(f)Met	His	His	His
8His	N17D, R66H	45	(f)Met	His	His	His
8His	R66H, N96D	46	(f)Met	His	His	His
8His	M41K, R66H	47	(f)Met	His	His	His
8His	N17,96D, M41K	48	(f)Met	His	His	His
8His	N17D, M41K	49	(f)Met	His	His	His
8His	M41K, N96D	50	(f)Met	His	His	His
8His	N17D, M41K, R66H	51	(f)Met	His	His	His
8His	M41K, R66H, N96D	52	(f)Met	His	His	His

Position in hA1M		Con'd	N-terminal 'tag'							
Tag	Mutations	Compound	X5	X6	X7	X8	X9	X10		
Broadest claim		1	His/absent	His/absent	His/absent	His/absent	His/absent	His/absent	Asp/absent/c	
6-His	'in sequence' mutations can be any	2	His	His	His	Absent	Absent	Absent	Asp	
8-His	in sequence' mutations can be any	3	His	His	His	His	His	His	Asp	
No tag;	in sequence' mutations can be any	4	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
Any tag	M41K	5	His/absent	His/absent	His/absent	His/absent	His/absent	His/absent	Asp/absent/Glu/Lys/	
Any tag	N17,96D	6	His/absent	His/absent	His/absent	His/absent	His/absent	His/absent	Asp/absent/Glu/Lys/	
No tag	N17D	7	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	N17,96D	8	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	N96D	9	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	N17,96D, R66H	10	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	M41K	11	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	R66H	12	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	N17,96D, M41K, R66H	13	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	N17D, R66H	14	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	R66H, N96D	15	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	M41K, R66H	16	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	N17,96D, M41K	17	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	N17D, M41K	18	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	M41K, N96D	19	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
Any tag	N17,96D, R66H	20	His/absent	His/absent	His/absent	His/absent	His/absent	His/absent	Asp/absent/Glu/Lys/	
No tag	N17D, M41K, R66H	21	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
No tag	M41K, R66H, N96D	22	Absent	Absent	Absent	Absent	Absent	Absent	Absent	
6His	N17D	23	His	His	His	Absent	Absent	Absent	Asp	
6His	N17,96D	24	His	His	His	Absent	Absent	Absent	Asp	
6His	N96D	25	His	His	His	Absent	Absent	Absent	Asp	
6His	N17,96D, R66H	26	His	His	His	Absent	Absent	Absent	Asp	

P017240DK1

Tag	Position in hA1M Mutations	on'd Compound	N-terminal 'tag'							
			X5	X6	X7	X8	X9	X10		
6His	M41K	27	His	His	His	Absent	Absent	Asp		
6His	R66H	28	His	His	His	Absent	Absent	Asp		
6His	N17,96D, M41K, R66H	29	His	His	His	Absent	Absent	Asp		
6His	N17D, R66H	30	His	His	His	Absent	Absent	Asp		
6His	R66H, N96D	31	His	His	His	Absent	Absent	Asp		
6His	M41K, R66H	32	His	His	His	Absent	Absent	Asp		
6His	N17,96D, M41K	33	His	His	His	Absent	Absent	Asp		
6His	N17D, M41K	34	His	His	His	Absent	Absent	Asp		
6His	M41K, N96D	35	His	His	His	Absent	Absent	Asp		
6His	N17D, M41K, R66H	36	His	His	His	Absent	Absent	Asp		
6His	M41K, R66H, N96D	37	His	His	His	Absent	Absent	Asp		
8His	N17D	38	His	His	His	His	His	Asp		
8His	N17,96D	39	His	His	His	His	His	Asp		
8His	N96D	40	His	His	His	His	His	Asp		
8His	N17,96D, R66H	41	His	His	His	His	His	Asp		
8His	M41K	42	His	His	His	His	His	Asp		
8His	R66H	43	His	His	His	His	His	Asp		
8His	N17,96D, M41K, R66H	44	His	His	His	His	His	Asp		
8His	N17D, R66H	45	His	His	His	His	His	Asp		
8His	R66H, N96D	46	His	His	His	His	His	Asp		
8His	M41K, R66H	47	His	His	His	His	His	Asp		
8His	N17,96D, M41K	48	His	His	His	His	His	Asp		
8His	N17D, M41K	49	His	His	His	His	His	Asp		
8His	M41K, N96D	50	His	His	His	His	His	Asp		
8His	N17D, M41K, R66H	51	His	His	His	His	His	Asp		
8His	M41K, R66H, N96D	52	His	His	His	His	His	Asp		

P017240DK1

Position in hA1M		Con'd				
Tag	Mutations		X11	X12	X13	
Broadest claim		1	Asp/absent/Glu/Lys/Arg	Asp/absent/Glu/Lys/Arg	Asp/absent/Glu/Lys/Arg	
6-His	'in sequence' mutations can be any	2	Asp	Asp	Asp	
8-His	in sequence' mutations can be any	3	Asp	Asp	Asp	
No tag;	in sequence' mutations can be any	4	Absent	Absent	Absent	
Any tag	M41K	5	Asp/absent/Glu/Lys/Arg	Asp/absent/Glu/Lys/Arg	Asp/absent/Glu/Lys/Arg	
Any tag	N17,96D	6	Asp/absent/Glu/Lys/Arg	Asp/absent/Glu/Lys/Arg	Asp/absent/Glu/Lys/Arg	
No tag	N17D	7	Absent	Absent	Absent	
No tag	N17,96D	8	Absent	Absent	Absent	
No tag	N96D	9	Absent	Absent	Absent	
No tag	N17,96D, R66H	10	Absent	Absent	Absent	
No tag	M41K	11	Absent	Absent	Absent	
No tag	R66H	12	Absent	Absent	Absent	
No tag	N17,96D, M41K, R66H	13	Absent	Absent	Absent	
No tag	N17D, R66H	14	Absent	Absent	Absent	
No tag	R66H, N96D	15	Absent	Absent	Absent	
No tag	M41K, R66H	16	Absent	Absent	Absent	
No tag	N17,96D, M41K	17	Absent	Absent	Absent	
No tag	N17D, M41K	18	Absent	Absent	Absent	
No tag	M41K, N96D	19	Absent	Absent	Absent	
Any tag	N17,96D, R66H	20	Asp/absent/Glu/Lys/Arg	Asp/absent/Glu/Lys/Arg	Asp/absent/Glu/Lys/Arg	
No tag	N17D, M41K, R66H	21	Absent	Absent	Absent	
No tag	M41K, R66H, N96D	22	Absent	Absent	Absent	
6His	N17D	23	Asp	Asp	Asp	
6His	N17,96D	24	Asp	Asp	Asp	
6His	N96D	25	Asp	Asp	Asp	
6His	N17,96D, R66H	26	Asp	Asp	Asp	

P017240DK1

Position in hA1M		Con'd			
Tag	Mutations		X11	X12	X13
6His	M41K	27	Asp	Asp	Asp
6His	R66H	28	Asp	Asp	Asp
6His	N17,96D, M41K, R66H	29	Asp	Asp	Asp
6His	N17D, R66H	30	Asp	Asp	Asp
6His	R66H, N96D	31	Asp	Asp	Asp
6His	M41K, R66H	32	Asp	Asp	Asp
6His	N17,96D, M41K	33	Asp	Asp	Asp
6His	N17D, M41K	34	Asp	Asp	Asp
6His	M41K, N96D	35	Asp	Asp	Asp
6His	N17D, M41K, R66H	36	Asp	Asp	Asp
6His	M41K, R66H, N96D	37	Asp	Asp	Asp
8His	N17D	38	Asp	Asp	Asp
8His	N17,96D	39	Asp	Asp	Asp
8His	N96D	40	Asp	Asp	Asp
8His	N17,96D, R66H	41	Asp	Asp	Asp
8His	M41K	42	Asp	Asp	Asp
8His	R66H	43	Asp	Asp	Asp
8His	N17,96D, M41K, R66H	44	Asp	Asp	Asp
8His	N17D, R66H	45	Asp	Asp	Asp
8His	R66H, N96D	46	Asp	Asp	Asp
8His	M41K, R66H	47	Asp	Asp	Asp
8His	N17,96D, M41K	48	Asp	Asp	Asp
8His	N17D, M41K	49	Asp	Asp	Asp
8His	M41K, N96D	50	Asp	Asp	Asp
8His	N17D, M41K, R66H	51	Asp	Asp	Asp
8His	M41K, R66H, N96D	52	Asp	Asp	Asp

P017240DK1

Tag	Mutations	Position in hA1M	Con'd	tag'/first AA					17	41	66	96
				X14	X15	X16	X17	X18				
Broadest claim			Compound									
6-His	'in sequence' mutations can be any		1	Lys/(f)Met/Absent/Glu/Asp/Arg	Asp/Asn	Met/Lys/Arg	Arg/His/Lys	Asp/A				
8-His	in sequence' mutations can be any		2	Lys	Asp/Asn	Met/Lys/Arg	Arg/His/Lys	Asp/A				
No tag;	in sequence' mutations can be any		3	Lys	Asp/Asn	Met/Lys/Arg	Arg/His/Lys	Asp/Asn				
Any tag			4	Absent	Asp/Asn	Met/Lys/Arg	Arg/His/Lys	Asp/Asn				
Any tag	M41K		5	Lys/(f)Met/Absent/Glu/Asp/Arg	Asn	Lys	Arg	Asn				
Any tag	N17,96D		6	Lys/(f)Met/Absent/Glu/Asp/Arg	Asp	Met	Arg	Asp				
No tag	N17D		7	(f)Met	Asp	Met	Arg	Asn				
No tag	N17,96D		8	(f)Met	Asp	Met	Arg	Asp				
No tag	N96D		9	(f)Met	Asn	Met	Arg	Asp				
No tag	N17,96D, R66H		10	(f)Met	Asp	Met	His	Asp				
No tag	M41K		11	(f)Met	Asn	Lys	Arg	Asn				
No tag	R66H		12	(f)Met	Asp	Met	His	Asn				
No tag	N17,96D, M41K, R66H		13	(f)Met	Asp	Lys	His	Asp				
No tag	N17D, R66H		14	(f)Met	Asp	Met	His	Asn				
No tag	R66H, N96D		15	(f)Met	Asn	Met	His	Asp				
No tag	M41K, R66H		16	(f)Met	Asn	Lys	His	Asn				
No tag	N17,96D, M41K		17	(f)Met	Asp	Lys	Arg	Asp				
No tag	N17D, M41K		18	(f)Met	Asp	Lys	Arg	Asn				
No tag	M41K, N96D		19	(f)Met	Asn	Lys	Arg	Asp				
Any tag	N17,96D, R66H		20	Lys/(f)Met/Absent/Glu/Asp/Arg	Asp	Met	His	Asp				
No tag	N17D, M41K, R66H		21	(f)Met	Asp	Lys	His	Asn				
No tag	M41K, R66H, N96D		22	(f)Met	Asn	Lys	His	Asp				
6His	N17D		23	Lys	Asp	Met	Arg	Asn				
6His	N17,96D		24	Lys	Asp	Met	Arg	Asp				
6His	N96D		25	Lys	Asn	Met	Arg	Asp				
6His	N17,96D, R66H		26	Lys	Asp	Met	His	Asp				

P017240DK1

Position in hA1M		Con'd	tag' / first AA		17	41	66	96
Tag	Mutations	Compound	X14	X15	X16	X17	X18	
6His	M41K	27	Lys	Asn	Lys	Arg	Asn	Asn
6His	R66H	28	Lys	Asn	Met	His	Asn	Asn
6His	N17,96D, M41K, R66H	29	Lys	Asp	Lys	His	Asp	Asp
6His	N17D, R66H	30	Lys	Asp	Met	His	Asn	Asn
6His	R66H, N96D	31	Lys	Asn	Met	His	Asp	Asp
6His	M41K, R66H	32	Lys	Asn	Lys	His	Asn	Asn
6His	N17,96D, M41K	33	Lys	Asp	Lys	Arg	Asp	Asp
6His	N17D, M41K	34	Lys	Asp	Lys	Arg	Asn	Asn
6His	M41K, N96D	35	Lys	Asn	Lys	Arg	Asp	Asp
6His	N17D, M41K, R66H	36	Lys	Asp	Lys	His	Asn	Asn
6His	M41K, R66H, N96D	37	Lys	Asn	Lys	His	Asp	Asp
8His	N17D	38	Lys	Asp	Met	Arg	Asn	Asn
8His	N17,96D	39	Lys	Asp	Met	Arg	Asp	Asp
8His	N96D	40	Lys	Asn	Met	Arg	Asp	Asp
8His	N17,96D, R66H	41	Lys	Asp	Met	His	Asp	Asp
8His	M41K	42	Lys	Asn	Lys	Arg	Asn	Asn
8His	R66H	43	Lys	Asn	Met	His	Asn	Asn
8His	N17,96D, M41K, R66H	44	Lys	Asp	Lys	His	Asp	Asp
8His	N17D, R66H	45	Lys	Asp	Met	His	Asn	Asn
8His	R66H, N96D	46	Lys	Asn	Met	His	Asp	Asp
8His	M41K, R66H	47	Lys	Asn	Lys	His	Asn	Asn
8His	N17,96D, M41K	48	Lys	Asp	Lys	Arg	Asp	Asp
8His	N17D, M41K	49	Lys	Asp	Lys	Arg	Asn	Asn
8His	M41K, N96D	50	Lys	Asn	Lys	Arg	Asp	Asp
8His	N17D, M41K, R66H	51	Lys	Asp	Lys	His	Asn	Asn
8His	M41K, R66H, N96D	52	Lys	Asn	Lys	His	Asp	Asp

P017240DK1

Position in hA1M		Con'd Compound	Comment
Tag	Mutations		
Broadest claim		1	
6-His	'in sequence' mutations can be any	2	when all X1-X14 are absent, X15 represents Asn, X16
8-His	in sequence' mutations can be any	3	and X17 represents Arg, then X18 cannot represent A:
No tag;	in sequence' mutations can be any	4	
Any tag	M41K	5	Proviso: when X15 represents Asn, X16 represents Met,
Any tag	N17,96D	6	and X17 represents Arg, then X18 cannot represent Asn
No tag	N17D	7	
No tag	N17,96D	8	
No tag	N96D	9	
No tag	N17,96D, R66H	10	Preferred mutation without tag
No tag	M41K	11	
No tag	R66H	12	
No tag	N17,96D, M41K, R66H	13	All mutations
No tag	N17D, R66H	14	
No tag	R66H, N96D	15	
No tag	M41K, R66H	16	
No tag	N17,96D, M41K	17	
No tag	N17D, M41K	18	
No tag	M41K, N96D	19	
Any tag	N17,96D, R66H	20	
No tag	N17D, M41K, R66H	21	
No tag	M41K, R66H, N96D	22	
6His	N17D	23	
6His	N17,96D	24	
6His	N96D	25	
6His	N17,96D, R66H	26	Preferred mutation with 6His

P017240DK1

Tag	Position in hA1M		Con'd Compound	Comment
	Tag	Mutations		
6His		M41K	27	
6His		R66H	28	
6His		N17,96D, M41K, R66H	29	
6His		N17D, R66H	30	
6His		R66H, N96D	31	
6His		M41K, R66H	32	
6His		N17,96D, M41K	33	
6His		N17D, M41K	34	
6His		M41K, N96D	35	
6His		N17D, M41K, R66H	36	
6His		M41K, R66H, N96D	37	
8His		N17D	38	
8His		N17,96D	39	
8His		N96D	40	
8His		N17,96D, R66H	41	Preferred mutation with 8His
8His		M41K	42	
8His		R66H	43	
8His		N17,96D, M41K, R66H	44	
8His		N17D, R66H	45	
8His		R66H, N96D	46	
8His		M41K, R66H	47	
8His		N17,96D, M41K	48	
8His		N17D, M41K	49	
8His		M41K, N96D	50	
8His		N17D, M41K, R66H	51	
8His		M41K, R66H, N96D	52	

Alpha-1-microglobulin – a general background

A1M is synthesized in the liver at a high rate, secreted into the blood stream and transported across the vessel walls to the extravascular compartment of all organs. The protein is also synthesized in other tissues (blood cells, brain, kidney, skin) but at a lower rate. Due to the small size, free A1M is rapidly filtered from blood in the kidneys.

A1M is a member of the lipocalin superfamily, a group of proteins from animals, plants and bacteria with a conserved three-dimensional structure but very diverse functions. Each lipocalin consists of a 160-190-amino acid chain that is folded into a β -barrel pocket with a hydrophobic interior. At least twelve human lipocalin genes are known. A1M is a 26 kDa plasma and tissue protein that so far has been identified in mammals, birds, fish and frogs. The three-dimensional structure of A1M determined by X-ray crystallography is shown in Figure 10. A1M is synthesized in the liver at a high rate, secreted into the blood stream and rapidly ($T_{1/2} = 2-3$ min) transported across the vessel walls to the extravascular compartment of all organs. A1M is found both in a free, monomeric form and as covalent complexes with larger molecules (IgA, albumin, prothrombin) in blood and interstitial tissues. Due to the small size, free A1M is rapidly filtered from blood in the kidneys. The major portion is then reabsorbed, but significant amounts are excreted to the urine.

Antioxidants are protective factors that eliminate oxidants or prevent harmful oxidation reactions. The human organism can produce antioxidants in response to oxidative stress. Such endogenous antioxidants include the superoxide-degrading enzyme superoxide dismutase (SOD), the hydrogen peroxide-degrading enzymes catalase and glutathione peroxidase, and the heme-degrading enzyme heme oxygenase-1 (HO-1). A1M was recently shown to be involved in protecting against oxidative tissue damage by functioning both as a scavenger of radicals and heme as well as a reductase and inhibitor of oxidation. Several recent papers demonstrate that A1M protects cell cultures and organ explants against oxidative damage, partly by accumulating in mitochondria and protecting mitochondrial function. Indeed, infusion of human recombinant A1M has been successfully employed for *in vivo* treatment of the oxidative stress-related diseases preeclampsia and hemoglobin-induced glomerular injuries in animal models.

Sequence and structural properties of A1M

The full sequence of human A1M is known. The protein consists of a polypeptide with 183 amino acid residues. Many additional A1M cDNAs and/or proteins have been detected, isolated and/or sequenced from other mammals, birds, amphibians, and fish. The length of the peptide chain of A1M differs slightly among species, due mainly to variations in the C-terminus. Alignment comparisons of the different deduced amino acid sequences show that the percentage of identity varies from approximately 75-80% between rodents or ferungulates and man, down to approximately 45% between fish and mammals. A free cysteine side-chain at position 34 is conserved. This group has been shown to be involved in redox reactions (see below), in complex formation with other plasma proteins and in binding to a yellow-brown chromophore. The three-dimensional structure of A1M shows that C34 is solvent exposed and located near the opening of the lipocalin pocket (see Figure 10).

In the present context the term " α_1 -microglobulin" intends to cover α_1 -microglobulin as identified in SEQ ID NO: 1 (human A1M) , SEQ ID NO: 2 (human recombinant A1M) and A1M from other species, including homologues, fragments or variants thereof having similar therapeutic activities. Thus, A1M as used herein is intended to mean a protein having at least 80% sequence identity with SEQ ID NO:1 or SEQ ID NO:2. It is preferred that A1M as used herein has at least 90% sequence identity with SEQ ID NO:1 or SEQ ID NO:2. It is even more preferred that A1M as used herein has at least 95% such as 99% or 100% sequence identity with SEQ ID NO:1 or SEQ ID NO:2. In a preferred aspect, the α_1 -microglobulin is in accordance with SEQ ID NO: 1 or 2 as identified herein. In the sequence listing the amino acid sequences of human A1M and human recombinant A1M (SEQ ID NOs 1 and 2, respectively) are given. However, homologues, variants and fragments of A1M having the important parts of the proteins as identified in the following are also comprised in the term A1M as used herein. Regarding alignments/identity see the following paragraph.

Details on alignment/identity

Positions of amino acid residues herein refer to the positions in human wt A1M as it is found in human blood plasma (SEQ ID NO:1). When referring to amino acid residues in recombinant A1M, which harbors a methionine or N-formyl methionine residue N-terminally linked to the glycine residue that is the initial residue in wt-A1M (SEQ ID NO: 2), or in mutated human A1M or A1M from other species a person skilled in the art will understand how to identify residues corresponding to residues in human wt-A1M (SEQ ID NO:1) even when positions are shifted due to e.g. deletions or insertions.

When recombinant proteins are produced they most often start with an initial Met residue, which may be removed using e.g. a methionine aminopeptidase or another enzyme with a similar activity. The A1M variants presented here may be with or without an initial Met residue.

5

Homologues of A1M

As mentioned above homologues of A1M can also be used in accordance with the description herein. In theory A1M from all species can be used for the purposes described herein including the most primitive found so far, which is from fish (plaice). A1M is also
 10 available in isolated form from human, orangutan, squirrel monkey, rat, naked mole rat, mouse, rabbit, guinea pig, cow, frog, chicken, walrus, manatee and plaice.

Considering homologues, variants and fragments of A1M, the following has been identified as important parts of the protein for the anti-oxidative effect:

- 15 Y22 (Tyrosine, pos 22, basepairs 64-66)
- C34 (Cystein, position 34, basepairs 100-102)
- K69 (Lysine, pos 69, basepairs 205-207)
- K92 (Lysine, pos 92, basepairs 274-276)
- K118 (Lysine, pos 118, basepairs 352-354)
- 20 K130 (Lysine, pos 130, basepairs 388-390)
- Y132 (Tyrosine, pos 132, basepairs 394-396)
- L180 (Leucine, pos 180, basepairs 538-540)
- I181 (Isoleucine, pos 181, basepairs 541-543)
- P182 (Proline, pos 182, basepairs 544-546)
- 25 R183 (Arginine, pos 183, basepairs 547-549)

Numbering of amino acids and nucleotides throughout the document refers to SEQ ID 1; if other A1M from other species, A1M analogs or recombinant sequences thereof are employed, a person skilled in the art will know how to identify the amino acids corresponding to the amino acids in SEQ ID NO: 1.
 30

Thus, in those cases, where A1M eg has 80% (or 90% or 95%) sequence identity with one of SEQ ID NO: 1 or 2, it is preferred that the amino acids mentioned above are present at the appropriate places in the molecule.

35

Human A1M is substituted with oligosaccharides in three positions, two sialylated complex-type, probably diantennary carbohydrates linked to N17 and N96 and one more simple oligosaccharide linked to T5. The carbohydrate content of A1M proteins from different species varies greatly, though, ranging from no glycosylation at all in *Xenopus*
5 *leavis* over a spectrum of different glycosylation patterns. However, one glycosylation site, corresponding to N96 in man, is conserved in mammals, suggesting that this specific carbohydrate may be a more important constituent of the protein than the other two oligosaccharides.

10 A1M is yellow-brown-coloured when purified from plasma or urine. The colour is caused by heterogeneous compounds covalently bound to various amino acid side groups mainly located at the entrance to the pocket. These modifications represent the oxidized degradation products of organic oxidants covalently trapped by A1M *in vivo*, for example heme, kynurenine and tyrosyl radicals.

15 A1M is also charge- and size-heterogeneous and more highly brown-coloured A1M-molecules are more negatively charged. The probable explanation for the heterogeneity is that different side-groups are modified to a varying degree with different radicals, and that the modifications alter the net charge of the protein. Covalently linked coloured
20 substances have been localized to C34, and K92, K118 and K130, the latter with molecular masses between 100 and 300 Da. The tryptophan metabolite kynurenine was found covalently attached to lysyl residues in A1M from urine of haemodialysis patients and appears to be the source of the brown colour of the protein in this case [6]. Oxidized fragments of the synthetic radical ABTS (2,2'-azino-di-(3-ethylbenzothiazoline)-6-
25 sulfonic acid) was bound to the side-chains of Y22 and Y132.

C34 is the reactive center of A1M. It becomes very electronegative, meaning that it has a high potential to give away electrons, by the proximity of the positively charged side-chains of K69, K92, K118 and K130, which induce a deprotonization of the C34 thiol
30 group which is a prerequisite of oxidation of the sulphur atom. Preliminary data shows that C34 is one of the most electronegative groups known.

Theoretically, the amino acids that characterize the properties of A1M (C34, Y22, K92, K118, K130, Y132, L180, I181, P182, R183), which will be described in more detail below, can be arranged in a similar three-dimensional configuration on another frame-
35

work, for instance a protein with the same global folding (another lipocalin) or a completely artificial organic or inorganic molecule such as a plastic polymer, a nanoparticle or metal polymer.

- 5 The three-dimensional arrangement of some of these amino acids (blue ovals, the lysines are depicted by a „+“), the A1M-framework (barrel), the electron-flow and the radical-trapping, are illustrated in Figure 10.

10 Accordingly, homologues, fragments or variants comprising a structure including the reactive centre and its surroundings as depicted above, are preferred.

Modifications and changes have been made in the structure of the polypeptides of this disclosure and still resulted in a molecule having similar functional characteristics as the original polypeptide (e.g., a conservative amino acid substitution). For example,
15 certain amino acids can be substituted for other amino acids in a sequence without appreciable loss of activity. Because it is the interactive capacity and nature of a polypeptide that defines that polypeptide's biological functional activity, certain amino acid sequence substitutions can be made in a polypeptide sequence and nevertheless obtain a polypeptide with like functional properties.

20

In making such changes, the hydropathic index of amino acids can be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a polypeptide is generally understood in the art. It is known that certain amino acids can be substituted for other amino acids having a similar hydropathic index or
25 score and still result in a polypeptide with similar biological activity. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. Those indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cysteine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).
30

It is believed that the relative hydropathic character of the amino acid determines the secondary structure of the resultant polypeptide, which in turn defines the interaction of
35 the polypeptide with other molecules, such as enzymes, substrates, receptors, antibodies, antigens, and the like. It is known in the art that an amino acid can be substituted

by another amino acid having a similar hydropathic index and still obtain a functionally equivalent polypeptide. In such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

5

Substitution of like amino acids can also be made on the basis of hydrophilicity, particularly where the biologically functional equivalent polypeptide or peptide thereby created is intended for use in immunological embodiments. The following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate
 10 (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); proline (-0.5 ± 1); threonine (-0.4); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically
 15 equivalent, and in particular, an immunologically equivalent polypeptide. In such changes, the substitution of amino acids the hydrophilicity values of which are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

20 As outlined above, amino acid substitutions are generally based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take one or more of the foregoing characteristics into consideration are well known to those of skill in the art and include, but are not limited to (original residue: exemplary substitution): (Ala: Gly,
 25 Ser), (Arg: Lys), (Asn: Gln, His), (Asp: Glu, Cys, Ser), (Gln: Asn), (Glu: Asp), (Gly: Ala), (His: Asn, Gln), (Ile: Leu, Val), (Leu: Ile, Val), (Lys: Arg), (Met: Leu, Tyr), (Ser: Thr), (Thr: Ser), (Trp: Tyr), (Tyr: Trp, Phe), and (Val: Ile, Leu). Embodiments of this disclosure thus contemplate functional or biological equivalents of a polypeptide as set forth above. In particular, embodiments of the polypeptides can include variants having
 30 about 50%, 60%, 70%, 80%, 90%, and 95% sequence identity to the polypeptide of interest.

In the present context, the homology between two amino acid sequences or between two nucleic acid sequences is described by the parameter "identity" (see also above).
 35 Alignments of sequences and calculation of homology scores may be done using a full Smith-Waterman alignment, useful for both protein and DNA alignments. The default

scoring matrices BLOSUM50 and the identity matrix are used for protein and DNA alignments respectively. The penalty for the first residue in a gap is -12 for proteins and -16 for DNA, while the penalty for additional residues in a gap is -2 for proteins and -4 for DNA. Alignment may be made with the FASTA package version v20u6.

5

Multiple alignments of protein sequences may be made using "ClustalW". Multiple alignments of DNA sequences may be done using the protein alignment as a template, replacing the amino acids with the corresponding codon from the DNA sequence.

- 10 Alternatively different software can be used for aligning amino acid sequences and DNA sequences. The alignment of two amino acid sequences is e.g. determined by using the Needle program from the EMBOSS package (<http://emboss.org>) version 2.8.0. The Needle program implements the global alignment algorithm described in. The substitution matrix used is BLOSUM62, gap opening penalty is 10, and gap extension penalty is 0.5.
- 15

The degree of identity between an amino acid sequence; e.g. SEQ ID NO: 1 and a different amino acid sequence (e.g. SEQ ID NO: 2) is calculated as the number of exact matches in an alignment of the two sequences, divided by the length of the "SEQ ID NO: 1" or the length of the "SEQ ID NO: 2", whichever is the shortest. The result is expressed in percent identity. See above regarding alignment and identity.

20

An exact match occurs when the two sequences have identical amino acid residues in the same positions of the overlap.

25

If relevant, the degree of identity between two nucleotide sequences can be determined by the Wilbur-Lipman method using the LASER- GENE™ MEGALIGN™ software (DNASTAR, Inc., Madison, WI) with an identity table and the following multiple alignment parameters: Gap penalty of 10 and gap length penalty of 10. Pairwise alignment parameters are Ktuple=3, gap penalty=3, and windows=20.

30

The percentage of identity of an amino acid sequence of a polypeptide with, or to, amino acids of SEQ ID NO: 1 may be determined by i) aligning the two amino acid sequences using the Needle program, with the BLOSUM62 substitution matrix, a gap opening penalty of 10, and a gap extension penalty of 0.5; ii) counting the number of exact matches in the alignment; iii) dividing the number of exact matches by the length

35

of the shortest of the two amino acid sequences, and iv) converting the result of the division of iii) into percentage. The percentage of identity to, or with, other sequences of the invention is calculated in an analogous way.

- 5 By way of example, a polypeptide sequence may be identical to the reference sequence, that is be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the % identity is less than 100%. Such alterations are selected from: at least one amino acid deletion, substitution (including conservative and non-conservative substitution), or insertion,
- 10 and wherein said alterations may occur at the amino- or carboxy-terminus positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence, or in one or more contiguous groups within the reference sequence.
- 15 Conservative amino acid variants can also comprise non-naturally occurring amino acid residues. Non-naturally occurring amino acids include, without limitation, trans-3-methylproline, 2,4-methanoproline, cis-4-hydroxyproline, trans-4-hydroxyproline, N-methyl-glycine, allo-threonine, methylthreonine, hydroxy-ethylcysteine, hydroxyethylhomocysteine, nitro-glutamine, homoglutamine, pipercolic acid, thiazolidine carboxylic acid,
- 20 dehydroproline, 3- and 4-methylproline, 3,3-dimethylproline, tert-leucine, norvaline, 2-azaphenyl-alanine, 3-azaphenylalanine, 4-azaphenylalanine, and 4- fluorophenylalanine. Several methods are known in the art for incorporating non- naturally occurring amino acid residues into proteins. For example, an *in vitro* system can be employed wherein nonsense mutations are suppressed using chemically aminoacylated suppressor tRNAs. Methods for synthesizing amino acids and aminoacylating tRNA are known
- 25 in the art. Transcription and translation of plasmids containing nonsense mutations is carried out in a cell-free system comprising an *E. coli* S30 extract and commercially available enzymes and other reagents. Proteins are purified by chromatography. In a second method, translation is carried out in *Xenopus* oocytes by microinjection of mutated mRNA and chemically aminoacylated suppressor tRNAs. Within a third method,
- 30 *E. coli* cells are cultured in the absence of a natural amino acid that is to be replaced (e.g., phenylalanine) and in the presence of the desired non-naturally occurring amino acid(s) (e.g., 2-azaphenylalanine, 3- azaphenylalanine, 4-azaphenylalanine, or 4-fluorophenylalanine). The non-naturally occurring amino acid is incorporated into the protein
- 35 in place of its natural counterpart. Naturally occurring amino acid residues can be converted to non-naturally occurring species by *in vitro* chemical modification. Chemical

modification can be combined with site-directed mutagenesis to further expand the range of substitutions. Alternative chemical structures providing a 3-dimensional structure sufficient to support the antioxidative properties of A1M may be provided by other technologies e.g. artificial scaffolds, amino-acid substitutions and the like. Furthermore, structures mimicking the active sites of A1M as listed above are contemplated as having the same therapeutic or physiologic function as A1M.

Pharmaceutical compositions and dosage

The present invention also provides a kit comprising:

- 10 i) a pharmaceutical composition comprising a contrast medium, and
 - ii) a pharmaceutical composition comprising A1M, any of the SEQ ID Nos: 1–10 and 17, or any of the A1M derived proteins mentioned herein (or a mutant, analogue, fragment or variant as defined herein).
- 15 In the following a listing of the sequences are given. The invention encompasses all possible variations eg such as those illustrated herein.

- SEQ ID NO: 1: wt hA1M
- SEQ ID NO: 2: rhA1M (i.e. Met-A1M)
- 20 SEQ ID NO: 3: Preferred mutation without “extension” – N17,96D, R66H
 - SEQ ID NO: 4: No extension, M41K
 - SEQ ID NO: 5: Preferred mutation with 6 His, N17,96D, R66H
 - SEQ ID NO: 6: 6His, M41K
 - SEQ ID NO: 7: preferred mutation with 8 His extension, N17,96D, R66H
 - 25 SEQ ID NO: 8: 8 His, M41K
 - SEQ ID NO: 9: Extension + wt hA1M
 - SEQ ID NO: 10: Omnibus A1M with possible extensions.
 - SEQ ID NO: 11-16: segments of wt hA1M
 - SEQ ID NO: 17: Omnibus C-terminally truncated A1M with possible extensions
 - 30 SEQ ID NO: 18: Preferred mutation without “extension” – N17,96D, R66; C-terminally truncated.
 - SEQ ID NO: 19: No extension, M41K; C-terminally truncated.
 - SEQ ID NO: 20: Preferred mutation with 6 His, N17,96D, R66H; C-terminally truncated.
 - SEQ ID NO: 21: 6His, M41K; C-terminally truncated.
 - 35 SEQ ID NO: 22: preferred mutation with 8 His extension, N17,96D, R66H; C-terminally truncated.

SEQ ID NO: 23: 8 His, M41K; C-terminally truncated.

The kit is in the form of one package containing the above-mentioned two compositions.

5

The pharmaceutical composition comprising a contrast medium is typically a composition already on the market.

10 The pharmaceutical composition comprising A1M (or an analogue, fragment or variant thereof as defined herein) is intended for i.v. administration. Accordingly, A1M can be formulated in a liquid, e.g. in a solution, a dispersion, an emulsion, a suspension etc.

15 For parenteral use suitable solvents include water, vegetable oils, propylene glycol and organic solvents generally approved for such purposes. In general, a person skilled in the art can find guidance in "Remington's Pharmaceutical Science" edited by Gennaro et al. (Mack Publishing Company), in "Handbook of Pharmaceutical Excipients" edited by Rowe et al. (PhP Press) and in official Monographs (e.g. Ph.Eur. or USP) relating to relevant excipients for specific formulation types and to methods for preparing a specific formulation.

20

A1M will be administrated in one or several doses in connection to the administration of contrast medium. Preferably, each dose will be administrated i.v. either as a single dose, as a single dose followed by slow infusion during a short time-period up to 60 minutes, or only as a slow infusion during a short time-period up to 60 minutes. The first dose may be administrated at the same time as the contrast medium, or within a period of 0-60 minutes before to 0-30 minutes after injection of the contrast medium. Additional A1M-doses can be added, but may not be necessary, after injection of the contrast medium. Each dose contains an amount of A1M which is related to the body-weight of the patient: 1-15 mg A1M/kg of the patient.

30

Sequence Listing Free Text

SEQ ID NO: 1

<223> Wildtype human A1M, no mutations

35

SEQ ID NO: 2

<223> rhA1M, ie N-terminal Met

SEQ ID NO: 3

<223> hA1M, No tag, N-terminal Met, N17,96D; R66H

SEQ ID NO: 4

5 <223> hA1M, not tag, N-terminal Met, M41K

SEQ ID NO: 5

<223> 6His, N17,96D; R66H

10 SEQ ID NO: 6

<223> hA1M, 6His, M41K

SEQ ID NO: 7

<223> 8His, N17,96D; R66H

15

SEQ ID NO: 8

<223> hA1M, 8His, M41K

SEQ ID NO: 9

20 <223> hA1M, 8His, no mut

SEQ ID NO: 10

<211> 193

<212> PRT

25 <213> Homo sapiens

<220>

<221> VARIANT

<222> 1

30 <223> Xaa = Met or absent

<220>

<221> VARIANT

<222> 1

35 <223> Xaa = Met or absent

<220>

<221> VARIANT

<222> 2

40 <223> Xaa = His or absent

<220>

<221> VARIANT

<222> 2

45 <223> Xaa = His or absent

<220>

<221> VARIANT

- <222> 3
<223> Xaa = His or absent
- <220>
5 <221> VARIANT
<222> 3
<223> Xaa = His or absent
- <220>
10 <221> VARIANT
<222> 4
<223> Xaa = His or absent
- <220>
15 <221> VARIANT
<222> 4
<223> Xaa = His or absent
- <220>
20 <221> VARIANT
<222> 5
<223> Xaa = His or absent
- <220>
25 <221> VARIANT
<222> 5
<223> Xaa = His or absent
- <220>
30 <221> VARIANT
<222> 6
<223> Xaa = His or absent
- <220>
35 <221> VARIANT
<222> 6
<223> Xaa = His or absent
- <220>
40 <221> VARIANT
<222> 7
<223> Xaa = His or absent
- <220>
45 <221> VARIANT
<222> 7
<223> Xaa = His or absent

<220>
<221> VARIANT
<222> 8
<223> Xaa = His or absent
5
<220>
<221> VARIANT
<222> 8
<223> Xaa = His or absent
10
<220>
<221> VARIANT
<222> 9
<223> Xaa = His or absent
15
<220>
<221> VARIANT
<222> 9
<223> Xaa = His or absent
20
<220>
<221> VARIANT
<222> 10
<223> Xaa = Asp, Glu, Lys, Arg, or absent
25
<220>
<221> VARIANT
<222> 10
<223> Xaa = Asp, Glu, Lys, Arg, or absent
30
<220>
<221> VARIANT
<222> 11
<223> Xaa = Asp, Glu, Lys, Arg, or absent
35
<220>
<221> VARIANT
<222> 11
<223> Xaa = Asp, Glu, Lys, Arg, or absent
40
<220>
<221> VARIANT
<222> 12
<223> Xaa = Asp, Glu, Lys, Arg, or absent
45
<220>
<221> VARIANT
<222> 12

<223> Xaa = Asp, Glu, Lys, Arg, or absent

<220>

<221> VARIANT

5 <222> 13

<223> Xaa = Asp, Glu, Lys, Arg, or absent

<220>

<221> VARIANT

10 <222> 13

<223> Xaa = Asp, Glu, Lys, Arg, or absent

<220>

<221> VARIANT

15 <222> 14

<223> Xaa = Asp, Glu, Lys, Arg, or absent

<220>

<221> VARIANT

20 <222> 14

<223> Xaa = Asp, Glu, Lys, Arg, or absent

<220>

<221> VARIANT

25 <222> 31

<223> Xaa = Asn or Asp

<220>

<221> VARIANT

30 <222> 55

<223> Xaa = Met or Lys

<220>

<221> VARIANT

35 <222> 80

<223> Xaa = Arg or His

<220>

<221> VARIANT

40 <222> 110

<223> Xaa = Asn or Asp

SEQ ID NO: 11

<223> Y1

45

SEQ ID NO: 12

<223> Y2

SEQ ID NO: 13
<223> Y3

5 SEQ ID NO: 14
<223> Y4

SEQ ID NO: 15
<223> Y5

10 SEQ ID NO: 16
<223> Y5

SEQ ID NO: 17
<211> 197
15 <212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
20 <222> 1
<223> Xaa = Met or absent

<220>
<221> VARIANT
25 <222> 1
<223> Xaa = Met or absent

<220>
<221> VARIANT
30 <222> 2
<223> Xaa = His or absent

<220>
<221> VARIANT
35 <222> 2
<223> Xaa = His or absent

<220>
<221> VARIANT
40 <222> 3
<223> Xaa = His or absent

<220>
<221> VARIANT
45 <222> 3
<223> Xaa = His or absent

<220>

<221> VARIANT
<222> 4
<223> Xaa = His or absent

5 <220>
<221> VARIANT
<222> 4
<223> Xaa = His or absent

10 <220>
<221> VARIANT
<222> 5
<223> Xaa = His or absent

15 <220>
<221> VARIANT
<222> 5
<223> Xaa = His or absent

20 <220>
<221> VARIANT
<222> 6
<223> Xaa = His or absent

25 <220>
<221> VARIANT
<222> 6
<223> Xaa = His or absent

30 <220>
<221> VARIANT
<222> 7
<223> Xaa = His or absent

35 <220>
<221> VARIANT
<222> 7
<223> Xaa = His or absent

40 <220>
<221> VARIANT
<222> 8
<223> Xaa = His or absent

45 <220>
<221> VARIANT
<222> 8
<223> Xaa = His or absent

- <220>
<221> VARIANT
<222> 9
5 <223> Xaa = His or absent
- <220>
<221> VARIANT
<222> 9
10 <223> Xaa = His or absent
- <220>
<221> VARIANT
<222> 10
15 <223> Xaa = Asp, Glu, Lys, Arg, or absent
- <220>
<221> VARIANT
<222> 10
20 <223> Xaa = Asp, Glu, Lys, Arg, or absent
- <220>
<221> VARIANT
<222> 11
25 <223> Xaa = Asp, Glu, Lys, Arg, or absent
- <220>
<221> VARIANT
<222> 11
30 <223> Xaa = Asp, Glu, Lys, Arg, or absent
- <220>
<221> VARIANT
<222> 12
35 <223> Xaa = Asp, Glu, Lys, Arg, or absent
- <220>
<221> VARIANT
<222> 12
40 <223> Xaa = Asp, Glu, Lys, Arg, or absent
- <220>
<221> VARIANT
<222> 13
45 <223> Xaa = Asp, Glu, Lys, Arg, or absent
- <220>
<221> VARIANT

<222> 13

<223> Xaa = Asp, Glu, Lys, Arg, or absent

<220>

5 <221> VARIANT

<222> 14

<223> Xaa = Asp, Glu, Lys, Arg, or absent

<220>

10 <221> VARIANT

<222> 14

<223> Xaa = Asp, Glu, Lys, Arg, or absent

<220>

15 <221> VARIANT

<222> 31

<223> Xaa = Asn or Asp

<220>

20 <221> VARIANT

<222> 55

<223> Xaa = Met or Lys

<220>

25 <221> VARIANT

<222> 80

<223> Xaa = Arg or His

<220>

30 <221> VARIANT

<222> 110

<223> Xaa = Asn or Asp

SEQ ID NO: 18

35 <223> hA1M, No tag, N-terminal Met, N17,96D; R66H; truncated

SEQ ID NO: 19

<223> hA1M, not tag, N-terminal Met, M41K; truncated

40 SEQ ID NO: 20

<223> 6His, N17,96D; R66H; truncated

SEQ ID NO: 21

<223> hA1M, 6His, M41K; truncated

45

SEQ ID NO: 22

<223> 8His, N17,96D; R66H; truncated

SEQ ID NO: 23
<223> hA1M, 8His, M41K; truncated

5 SEQ ID NO: 24
<223> 1. M8H5GIEGR-Mouse

SEQ ID NO: 25
<223> 2. M8H5GIEGR-Naked Mole rat

10 SEQ ID NO: 26
<223> 3. M8H5GIEGR-Frog

15 SEQ ID NO: 27
<223> 4. M8H5GIEGR-Chicken

SEQ ID NO: 28
<223> 5. M8H5GIEGR-Rabbit

20 SEQ ID NO: 29
<223> 6. M8H5GIEGR-SQ Monkey

SEQ ID NO: 30
<223> 7. M8H5GIEGR-Walrus

25 SEQ ID NO: 31
<223> 8. M8H5GIEGR-Manatee

30 SEQ ID NO: 32
<223> 9. M8H5GIEGR-Plaice

SEQ ID NO: 33
<223> 10. M8H5GIEGR-Orangutan

35 SEQ ID NO: 34
<223> 11. M8H5GIEGR-Human P35K

SEQ ID NO: 35
<223> 12. M8H5GIEGR-Human M41K

40 SEQ ID NO: 36
. M8H5GIEGR-Human R66H

45 SEQ ID NO: 37
<223> 14. M8H5GIEGR-Human T75K

SEQ ID NO: 38
<223> 15. M8H5GIEGR-Human T75Y

SEQ ID NO: 39

<223> 16. M8H5GIEGR-Human M99K

SEQ ID NO: 40

5 <223> 17. M8H5GIEGR-Human S101Y

SEQ ID NO: 41

<223> 18. M8H5GIEGR-Human K69.92.118.130R

10 SEQ ID NO: 42

<223> 19. M8H5GIEGR-Coelacanth

SEQ ID NO: 43

<223> 21. M8H5GIEGR-Human L89T

15

SEQ ID NO: 44

<223> 22. M8H5GIEGR-Human N1796D

SEQ ID NO: 45

20 <223> 23. M8H5GIEGR-Human T45K

SEQ ID NO: 46

<223> 24. M8H5GIEGR-Human A135E

25 SEQ ID NO: 47

<223> 25. M8H5GIEGR-Human V170S

SEQ ID NO: 48

<223> 26. M8H5GIEGR-Human

30

SEQ ID NO: 49

<223> 27. M8H5GIEGR-Human G172Q

SEQ ID NO: 50

35 <223> 33. M8H4DK-Human M41K+

SEQ ID NO: 51

<223> 34. M8H4DK-Human M41K+N1796D 34

40 SEQ ID NO: 52

<223> 35. M8H4DK-Human R66H+N1796D

SEQ ID NO: 53

<223> 36. M8H4DK-Human M41K+R66H+N1796D

45

SEQ ID NO: 54

<223> 38. M8H4DK-Human R66H

SEQ ID NO: 55
<223> 39.M8H4DK-Human

5 SEQ ID NO: 56
<223> 40. M8H-Human wt

SEQ ID NO: 57
<223> 41. M8H-Human R66H+N1796D

10 SEQ ID NO: 58
<223> 60. M8H4DK-Human wt

15 SEQ ID NO: 59
<223> 1.M8H5GIEGR-Mouse

SEQ ID NO: 60
<223> 2. M8H5GIEGR-Naked Mole

20 SEQ ID NO: 61
<223> 3. M8H5GIEGR-Frog

SEQ ID NO: 62
<223> 4. M8H5GIEGR-Chicken

25 SEQ ID NO: 63
<223> 5. M8H5GIEGR-Rabbit

30 SEQ ID NO: 64
<223> 6. M8H5GIEGR-SQ Monkey

SEQ ID NO: 65
<223> 7. M8H5GIEGR-Walrus

35 SEQ ID NO: 66
<223> 8. M8H5GIEGR-Manatee

SEQ ID NO: 67
<223> 9. M8H5GIEGR-Plaice

40 SEQ ID NO: 68
<223> 10. M8H5GIEGR-Orangutan

45 SEQ ID NO: 69
<223> 11. M8H5GIEGR-Human P35K

SEQ ID NO: 70
<223> 12. M8H5GIEGR-Human M41K

SEQ ID NO: 71
<223> 13. M8H5GIEGR-Human R66H

5 SEQ ID NO: 72
<223> 14. M8H5GIEGR-Human T75K

SEQ ID NO: 73
<223> 15. M8H5GIEGR-Human T75Y

10 SEQ ID NO: 74
<223> 16. M8H5GIEGR-Human M99K

15 SEQ ID NO: 75
<223> 17. M8H5GIEGR-Human S101Y

SEQ ID NO: 76
<223> 18. M8H5GIEGR-Human K69.92.118.

20 SEQ ID NO: 77
<223> 19. M8H5GIEGR-Coelacanth

SEQ ID NO: 78
<223> 21. M8H5GIEGR-Human L89T

25 SEQ ID NO: 79
<223> 22. M8H5GIEGR-Human N1796D

30 SEQ ID NO: 80
<223> 23. M8H5GIEGR-Human T45K

SEQ ID NO: 81
<223> 24. M8H5GIEGR-Human A135E

35 SEQ ID NO: 82
<223> 25. M8H5GIEGR-Human V170S

SEQ ID NO: 83
<223> 26. M8H5GIEGR-Human V148D

40 SEQ ID NO: 84
<223> 27. M8H5GIEGR-Human G172Q

45 SEQ ID NO: 85
<223> 33. M8H4DK-Human M41K+R66H

SEQ ID NO: 86
<223> 34. M8H4DK-Human M41K+N1796D

SEQ ID NO: 87

<223> 35. M8H4DK-Human R66H+N1796D

SEQ ID NO: 88

5 <223> 36. M8H4DK-Human M41K+R66H+N1796D

SEQ ID NO: 89

<223> 37. M8H4DK-Human M41K

10 SEQ ID NO: 90

<223> 38. M8H4DK-Human R66H

SEQ ID NO: 91

<223> 39. M8H4DK-Human N1796D

15

SEQ ID NO: 92

<223> 40. M8H-Human wt

SEQ ID NO: 93

20 <223> 41. M8H-Human R66H+N1796D

SEQ ID NO: 94

<223> 42. untagged-Human R66H+N1796D

25 SEQ ID NO: 95

<223> 61. untagged-Human wt

Abbreviations

30 A1M: alpha-1-microglobulin, IB: inclusion bodies, wt: wildtype, R66H: point mutation in A1M-gene leading to expression of His instead of Arg at position 66, N17,96D: point mutations in A1M-gene leading to expression of Asp instead of Asn at positions 17 and 96, M8H4DK: peptide with amino acid sequence MHHHHHHHHDDDDK, M6H4DK: peptide with amino acid sequence MHHHHHHHHDDDDK, M8H: peptide with amino acid sequence MHHHHHHHHH, M8H5GIEGR: peptide with amino acid sequence MHHHHHHHHHGGGGGIEGR, CV: column volume, SEC: size-exclusion chromatography, DLS: dynamic light scattering, DSF: differential scanning fluorimetry, Gu-HCl: guanidine hydrochloride, ORAC: oxygen radical antioxidant capacity, SD: standard deviation, PAGE: polyacrylamide gel electrophoresis

40

Definitions

In describing and claiming the disclosed subject matter, the following terminology will be used in accordance with the definitions set forth below. Unless defined otherwise, all

technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a substantial difference over what is generally understood in the art. The techniques and procedures described or referenced herein are generally well understood and commonly employed using conventional methodology by those skilled in the art. As appropriate, procedures involving the use of commercially available kits and reagents are generally carried out in accordance with manufacturer defined protocols and/or parameters unless otherwise noted. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present disclosure, the preferred methods and materials are described herein.

15 In this specification, unless otherwise specified, “a” or “an” means “one or more”.

The terms “*treatment or prophylaxis*” in their various grammatical forms in relation to the present invention refer to preventing, curing, reversing, attenuating, alleviating, ameliorating, inhibiting, minimizing, suppressing, or halting (1) the deleterious effects of a disorder, (2) disorder progression, or (3) disorder causative agent.

The term “*effective amount*” in relation to the present invention refers to that amount which provides a therapeutic effect for a given condition and administration regimen. This is a predetermined quantity of active material calculated to produce a desired therapeutic effect in association with the required additives and diluents; i.e., a carrier, or administration vehicle. Further, it is intended to mean an amount sufficient to reduce and most preferably prevent a clinically significant deficit in the activity and response of the host. Alternatively, a therapeutically effective amount is sufficient to cause an improvement in a clinically significant condition in a host. As is appreciated by those skilled in the art, the amount of a compound may vary depending on its specific activity. Suitable dosage amounts may contain a predetermined quantity of active composition calculated to produce the desired therapeutic effect in association with the required diluents; i.e., carrier, or additive. Further, the dosage to be administered will vary depending on the active principle or principles to be used, the age, weight etc. of the patient to be treated but will generally be within the range from 0,001 to 1000 mg/kg body weight/day. Moreover, the dose depends on the administration route.

The term "polypeptides" includes proteins and fragments thereof. Polypeptides are disclosed herein as amino acid residue sequences. Those sequences are written left to right in the direction from the amino to the carboxy terminus. In accordance with standard nomenclature, amino acid residue sequences are denominated by either a three letter or a single letter code as indicated as follows: Alanine (Ala, A), Arginine (Arg, R), Asparagine (Asn, N), Aspartic Acid (Asp, D), Cysteine (Cys, C), Glutamine (Gln, Q), Glutamic Acid (Glu, E), Glycine (Gly, G), Histidine (His, H), Isoleucine (Ile, I), Leucine (Leu, L), Lysine (Lys, K), Methionine (Met, M), Phenylalanine (Phe, F), Proline (Pro, P), Serine (Ser, S), Threonine (Thr, T), Tryptophan (Trp, W), Tyrosine (Tyr, Y), and Valine (Val, V).

"Variant" refers to a polypeptide or polynucleotide that differs from a reference polypeptide or polynucleotide, but retains essential properties. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall (homologous) and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more modifications (e.g., substitutions, additions, and/or deletions). A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polypeptide may be naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally.

The invention is further illustrated in – but not limited to – the figures.

Legend to figures

Figure 1. Amino acid sequence alignment of A1M from 12 different species.

The amino acid sequence of human wt A1M and 11 additional species investigated in project phase I were aligned using the <http://www.ebi.ac.uk/Tools/msa/clustalw2/> software. The identity of the different sequences to the human sequence is presented as percent. Amino acids identical between all species in the set are marked with yellow. Additionally, amino acids believed to be important in human A1M are marked: the unpaired cysteine residue (C34) important for reduction and antioxidant properties (Allhorn et al., 2005) as well as heme binding is marked with pink (Mening and Skerra, 2012). The asparagines known to be glycosylated (N17 and 96) are marked with green (Escribano et al, 1990). The four lysines (K69, 92, 118 and 130) that have been found

modified in urine A1M (Åkerström et al., 1995; Berggård et al., 1999) and are believed to be important for the reductase activity (Allhorn et al., 2005) are marked with light blue. The H123 suggested to take part in the heme-binding (Meining and Skerra, 2012) is marked with grey and finally the tyrosines (Y22 and 132) shown to be involved in radical scavenging (Åkerström et al., 2007) are marked in dark blue.

Figure 2. SDS-PAGE of expressed proteins in phase III.

Equal amount of bacterial lysate from uninduced samples (marked 0) and samples taken one to four hours after induction (marked 1, 2, 3, 4) were separated by SDS-PAGE. In these samples a band slightly below 25kDa is expressed at increasing amounts by time. The intensity of the bands culminates after 3 hours of induction. M8H-tagged wt-A1M and R66K+N17,96D-A1M are expressed with molecular weights slightly smaller than for the M8H4DK bands, but also here the expression level culminates around 3 hours. Untagged wt-A1M and R66H+N17,96D-A1M appear as even smaller bands, but here the intensity of the bands is stronger at 4 hours than 3 hours indicating a slightly delayed expression, in particular for wt A1M.

Figure 3. Yield, purity and aggregation of phase III variants.

A. The yields of purified A1M of all phase III variants were compared. All variants resulted in good yields, with a slightly lower yield of the untagged wt-A1M. B. The purity was investigated by separation of 10µg of all variants on SDS-PAGE. The purity was determined by densitometric analysis of the main monomeric band compared to all bands using the Image software from Bio-Rad. All histidine-tagged variants showed very high purity while the purities of the untagged variants were a little bit lower (around 90%). C. The presence of large aggregates were analysed by separating 20µg of each variant by native PAGE. The intensity of the main monomeric band and the material remaining in the application slit (very large aggregated) were determined by densitometry. Most variants showed low amounts of large aggregates with exception of M8H-wt-A1M and the untagged variants, which show a slightly higher percentage. All variants are coded by a number given in panel A.

Figure 4. Aggregation analysis of 100µM and 1mM solutions of phase III variants in various buffers.

The tendency to aggregate after concentration from 100µM to 1mM in different buffers was investigated by native PAGE analysis. 20µg of protein were separated in each lane. The percent large aggregates were calculated using densitometry analysis of the

individual band intensities.

- A. All variants were concentrated from 100µM to 1mM in Tris-buffer pH 8.0 or 7.4 and separated by PAGE. The M8H4DK-wt, R66H+N1796D, M41K, R66H, N1796D are labelled 60, 35, 37, 38 and 39 respectively, the M8H-wt and M8H-R66H+N1796D are labelled 40 and 41 and the untagged wt and R66H+N1796D are labelled 61 and 42.
- B. The variants were concentrated to 1mM in TRIS-buffer pH 8.0 and 7.4 and in PBS pH7.4, subjected to one freeze-thaw cycle, and then separated by PAGE and analysed as described.

10 *Figure 5. Analysis of reductase and antioxidant capacity of phase III variants.*

- A. The reductase activity was analysed as the reduction of the ABTS radical. 0-2µM A1M were added in duplicates to 56µM ABTS radical. The reduction was followed as decrease of absorbance at 405nm during 95seconds. The area under the curve (AUC) for each concentration was calculated and the Net AUC was calculated by subtracting the AUC of buffer only. The average Net AUC+/- SEM of the duplicates was plotted against concentration. All M8H4DK-variants have about the same activity as wt A1M with a tendency to lower activity for M41K and R66H. The M8H and untagged variants also show full activity with a small tendency of higher activity of the R66H+N1796D variants compared to wt A1M.
- B. The antioxidation ability was investigated in the ORAC assay. The activities of the A1M-variants were compared to a Trolox standard and expressed as number of Trolox equivalents. Each assay was done in triplicates and the result of M84DK-wt was set to 100%. The antioxidation capacities of all other variants were expressed in relation. The M8H4DK-R66H+N17,96D variant showed significantly higher capacity than the M8H4DK-wt A1M. When the tags were shortened or removed the difference was smaller. Data presented are the combined results of two independent experiments.

Figure 6. Analysis of reductase activity and heme binding of phase III variants.

- A. The ability of A1M-variants to reduce cytochrome c was investigated by mixing dilution series (0-10µM) of A1M with 100µM cytochrome c+100µM NADH and following the increase in absorbance at 550nm for 20 minutes. The assay was done in duplicates. The AUC was calculated for each concentration and the Net AUC was calculated by subtraction of the AUC of buffer only. Data are presented as the Net AUC +/- SEM of two independent experiments. Ovalbumin was used as a negative control. Most A1M variants showed a slightly lower reduction capacity compared to wt at the lower concentrations. For N17,96D-A1M and R66H+N17,96D-A1M it is significant (at 0.3-0.6µM).

Shortening and removal of the tag had no influence of this property.

B. The incorporation of heme into the A1M-variants was analysed by the appearance of an absorbance peak between 410-420nm, and the magnitude of the red-shift of the peak (Karnaukhova et al., 2014; Rutardottir et al., 2016). 44µM protein solutions were mixed with 40µM heme in duplicates and incubated for 2 hours at RT. The mixtures were analysed by wavelength scan between 270-450nm. The position of the maximal peak and the 413:386 ratio were calculated. All M8H4DK-variants have about the same red-shift and ratio as wt-A1M, while the M8H-tagged variants have significantly higher ratio. The untagged variants lack red-shift activity, confirming previous results (Karnaukhova et al., 2014).

C. The binding of A1M to heme agarose. The specific binding of A1M was analysed by mixing dilution series of A1M, or the control ovalbumin, with heme-agarose or control agarose. The assay was made in duplicates. Protein quantification of the starting material and the flow-throughs from heme-agarose and control agarose incubations was determined using BCA. From these data the amount of protein specifically bound to heme agarose could be determined for each sample. Bound protein was plotted against added protein, and a linear correlation was seen for all variants. The slope of the line was calculated with linear regression and the average between the duplicates is shown. All variants bind heme in this method to about the same extent.

20

Figure 7. Rescue of K562 cells from heme-induced cell death.

K562 cells were exposed to 100µM heme in the presense of a dilution series of A1M (0-10µM) for 1 hour. Then, cell death was monitored as release of LDH into the medium. The absorbance value from live cells was subtracted and the signal of heme-incubated cells without A1M was set to 100%. The values of the A1M incubations were calculated in relation to this. The assay was made in duplicates. The average result from three independent experiments (mean +/-SEM) is shown in the figure. No significant difference could be seen between any of the variants except for untagged R66H+N17,96D-A1M that has lower activity than untagged wt-A1M.

30

Figure 8. Size-distribution and reductase activity of M8H4DK-R66H+N17,96D-A1M and M8H4DK-wt-A1M after storage at +4°C and room-temperature.

M8H4DK-Wt-A1M and M8H4DK-R66H+N17,96D-A1M are depicted as "wt" and "variant 35" and shown in the upper and lower panels, respectively. Aggregation was analysed by SEC-FPLC. Monomeric A1M and large aggregates were eluted around 15ml and 8ml, respectively. The small shoulder seen around 13-14ml most likely is dimeric

35

A1M. The percentage of large aggregates was calculated from the area under the 8ml peak compared to the total peak area. The recovery of protein (%; shown in *italic*) after stress exposure was calculated from the total peak areas compared to the total peak area of the starting material. The reduction activity of ABTS was analysed in 2µM A1M solutions. Data are shown as average of duplicates. Freshly thawed 100µM A1M solutions are shown in (A), 100µM A1M exposed to five freeze-thaw cycles (B), 1mM A1M stored at +4°C over-night (C), 1mM A1M stored at +4°C for a week (D), 100µM A1M stored at RT for a week (E), and 1mM A1M stored at RT for a week (F). Data show that M8H4DK-R66H+N1796D better tolerates concentration to 1mM (C & D) and storage at RT (E and F).

Figure 9. Size-distribution and reductase activity of 1mM M8H4DK-wt-A1M and M8H4DK-R66H+N17,96D-A1M after storage at +37°C.

Wt-A1M and R66H+N17,96D-A1M are depicted as “ht2014” and “35” and shown in the upper and lower panels, respectively. Aggregation was analysed by SEC-FPLC. Monomeric A1M and large aggregates are eluted around 20min and 12min, respectively. The small shoulder seen around 18-19min most likely is dimeric A1M. The percentage of large aggregates was calculated from the area under the 12 min peak compared to the total peak area. The recovery of protein (%; shown in *italic*) after stress exposure was calculated from the total peak areas compared to the total peak area of the starting material. The reduction activity of ABTS was analysed in 2µM A1M solutions. Data are shown as average of duplicates. 1mM A1M solutions stored for 1.5hours (B), 2.5hours (C) and 4.5hours (D) were compared to 1mM A1M start material (A). The M8H4DK-wt at 4.5hours was completely precipitated and could not be analysed. Data show that M8H4DH-R66H+N1796D better tolerates storage at +37°C than M8H4DK-wt.

Figure 10. Model of the structure of A1M. The model was prepared as described (ref 29). The eight β-strands, shown as ribbons, form a slightly cone-shaped cylinder with a hydrophobic interior: the “lipocalin pocket”. One side of the lipocalin pocket is open (shown by the arrow), i.e. it permits entrance of small molecules. The opposite side is closed. Two α-helices are shown as cylinders. The positions of three carbohydrate groups (T5; N17; N96) and four side-chains involved in reductase activity (C34; K92; K118; K130) are shown.

Figure 11

Heme-binding analysed by migration shift/fluorescence quenching on native PAGE (**A** and **B**), and UV-absorbance spectrophotometry (**C**). **A**. Fifteen μg M8H4DK-wt A1M (wt-A1M) or M8H4DK-35-A1M (35-A1M) were incubated with different amounts of heme for 30 min at 20°C, separated by native PAGE, and the gel analysed by tryptophan fluorescence (Flourescence) and densitometry scanning after Coomassie staining (Stain). **B**. The images were digitalized by using Image Lab™ Software (Bio-Rad). Heme binding, measured as fluorescence quenching (black) and migration distance (blue) were plotted against the molar ratio A1M:heme. Mean values of duplicate experiments are shown, wt-A1M (filled symbols), 35-A1M (open symbols). **C**. A1M and heme were mixed (32 and 19 μM , respectively), incubated for 2h at 20°C and scanned. The absorbance of the proteins alone at 32 μM are shown as comparison. The absorbance of the buffer (20 mM Tris-HCl, pH 8.0 + 0.15M NaCl) was subtracted from all scans as blank.

Figure 12

Comparison of the enzymatic properties of M8H4DK-wt A1M (wt-A1M) and M8H4DK-35-A1M (35-A1M). **A**. Freshly purified wt-A1M (■) or 35-A1M (○) at various concentrations were mixed with ABTS-radical at 56 μM in 25mM sodium phosphate buffer pH 8.0 in microtiter plate wells, and the rate of reduction was followed by reading the absorbance at 405 nm during 95 seconds. The absorbance for each concentration was plotted against time and the area under the curve (AUC) between 0 and 95 s was calculated for each concentration. The net AUC was calculated by subtracting the AUC of buffer only. Mean of triplicates +/- SEM are shown. **B**. The ABTS-reduction rate was determined as described in **A**, but using wt-A1M (■) or 35-A1M (○) after storage for 7 days at 4°C or room-temperature and 0.1 or 1 mM. Single experiments are shown. **C**. The reduction of cytochrome c was investigated by mixing dilution series (0-10 μM) of wt-A1M (■) or 35-A1M (○) with 100 μM cytochrome c+100 μM NADH and following the increase in absorbance at 550nm for 20 minutes. The assay was done in duplicates. The AUC was calculated for each concentration and the Net AUC was calculated by subtraction of the AUC of buffer only. Data are presented as the Net AUC +/- SEM of two independent experiments. Ovalbumin was used as a negative control. **D**. The antioxidant ability was investigated in the ORAC assay. The activities of the A1M-variants at 5 μM were compared to a Trolox standard and expressed as number of Trolox equivalents. Each assay was done in triplicates and the result of wt-A1M was set to 100%. Data presented are the mean of two independent experiments +/- SEM.

Figure 13

K562 cells cultured at 10^5 cells per well in a 96-well microtiter plate, were exposed to 100 μ M heme in the presence of a dilution series of M8H4DK-wt A1M (wt-A1M) or M8H4DK-35-A1M (35-A1M) (0-10 μ M) for 1 hour. Cell death was monitored as release of LDH into the medium. The LDH-value from live cells was subtracted and the signal of heme-incubated cells without A1M was set to 100% and the values of the A1M incubations were calculated in relation to this. The assay was made in duplicates. The average result from three independent experiments (mean $\pm$ SEM) is shown. Wt-A1M (■) or 35-A1M (○).

10

Figure 14

HK-2 cells were exposed to a mixture of 200 μ M $(\text{NH}_4)\text{Fe}(\text{SO}_4)_2$, 400 μ M hydrogen peroxide, and 2 mM ascorbate (the Fenton reaction, displayed in **A** and **B**) or 0-30 μ M heme (displayed in **C** and **D**) with or without the simultaneous addition of 0-20 μ M M8H4DK-wt A1M (wt-A1M) (displayed as ■ and black columns) or M8H4DK-35-A1M (35-A1M) (displayed as ○ and white columns) for 6 hours. After incubation, cells were analyzed for cell viability using WST-1 (displayed in **A** and **C**) or mRNA expression of HO-1 and Hsp70 (displayed in **B** and **D**) as described in materials and methods. The cell viability (**A** and **C**) was normalized against control samples from untreated cells. Results are from triplicate experiments and presented as mean $\pm$ SEM. The mRNA expression of HO-1 and Hsp70 (**B** and **D**) was normalized against GAPDH and is given as fold change. The fold-change values were calculated by normalizing against control samples from untreated cells. Results are from triplicate experiments and presented as mean $\pm$ SEM. Differences between the respective exposures and control conditions were analyzed using One way ANOVA with *post hoc* Bonferroni correction. * indicates statistical comparison vs. Fenton (displayed in **B**) or heme (displayed in **D**). *P < 0.05, **P < 0.01, ***P < 0.001. No significant difference was observed when comparing wt-A1M vs. 35-A1M.

Figure 15

Plasma clearance (pharmacokinetics, displayed in **A**) and biodistribution (displayed in **B**) of M8H4DK-wt A1M (wt-A1M) and M8H4DK-35-A1M (35-A1M) injected intravenously in animals. **A**. Wt-A1M (■) or 35-A1M (○) was injected (5mg/kg) in Wistar rats and blood was collected at regular intervals. A1M-concentrations were determined by RIA using the particular A1M-variant as standard. Each point are from three animals and presented as mean $\pm$ SEM. **B**. Wt- or 35-A1M, 5 mg/kg, was injected intravenously

35

in C57BL/6NRj-mice which were sacrificed 10 and 30 min post-injection. Organs were sampled, weighed and homogenized. Concentrations of injected (human) A1M were determined by sandwich-ELISA. Each bar are from three animals and presented as mean $\pm$ SD. Wt-A1M, 10 min (black) and 30 min (dark gray); 35-A1M, 10 min (white) and 30 min (light gray).

Figure 16

Female C57BL/6 mice were exposed to Glycerol (2.0 ml/kg, i.m.) followed by i.v. administration of either M8H4DK-wt A1M (wt-A1M) (dark grey bars, n=10), M8H4DK-35-A1M (35-A1M) (white bars, n=10) or vehicle buffer (sham control, grey bars, n=6) 30 minutes post-glycerol injections. At 4 hours (post-glycerol administration) animals were euthanized and kidneys excised, snap frozen and subsequently analyzed for mRNA expression of HO-1 (A) and Hsp70 (B) using real-time PCR as described in materials and methods. mRNA expression were normalized against those of GAPDH and fold change values were calculated by normalizing against control samples from untreated animals (controls). Results are presented as as box plots, displaying medians and 25th and 75th percentiles. Statistical comparison between groups were performed by ANOVA with *post hoc* Bonferroni correction. * indicates statistical comparison vs. Glycerol. *P < 0.05, **P < 0.01. No significant difference was observed when comparing wt-A1M vs. 35-A1M.

The invention is further illustrated in the following examples. The examples are illustrative and do not limit the scope of the invention in any manner.

Experimental

The experimental work was divided into three major phases:

- Phase I) Expression of the 27 A1M-variants described above followed by analysis of solubility, stability and function;
- Phase II) Design, expression and analysis of a few A1M-variants with expected optimal properties based on the outcome of phase 1;
- Phase III) Design, expression and analysis of wt-A1M and the most successful mutated A1M-variant equipped with or without N-terminal tags.

In Phase I, four lines of reasoning were used, when selecting the positions and identities of mutated amino acid side-groups: 1) Animal homologues from a variety of species with different expected environmental pressure in terms of oxidative stress, temperature, oxygen pressure were expressed (N=12); 2) Single amino acid substitutions that occur frequently among the 56 sequenced A1M-homologues at positions located in loops 1-4 or the interior surface of the hydrophobic pocket, were introduced into the human gene construct and expressed (N=3); 3) Addition or removal of favourably located lysyl or tyrosyl residues, based on the hypothesis that these may influence pKa of the C35 thiolyl (Allhorn et al., 2005) or serve as radical-trapping sites (Berggård et al., 1999; Sala et al., 2004; Åkerström et al. 2007) (N=5); 4) Hydrophobic→hydrophilic substitutions on the surface of the protein, with no predicted influence on function or folding (N=7).

Materials and methods

Expression

The sequences of the A1M variants were provided to DNA2.0, Inc. (USA) which synthesized the genes and cloned them into their PJ401Express vector (T5 promoter, kanamycin resistance). The DNA sequence was confirmed by sequencing. The vectors were transformed into competent E. coli (BL21 Star(DE3) (Invitrogen, Life technologies corp, USA)) according to the manufacturer's instructions and four individual clones of each variants were tested for micro-expression. The clone with highest expression of each variant was prepared as a glycerol stock which was used for production expression.

A1M variant expression clones were grown in complete NYAT (15mM $(\text{NH}_4)_2\text{SO}_4$, 84mM K_2HPO_4 , 23mM $\text{NaH}_2\text{PO}_4 \times \text{H}_2\text{O}$, 2.2mM $(\text{NH}_4)_2\text{Hcitrate}$, 1% (w/v) glucose, 2mM MgSO_4 , 9 μM $\text{CaCl}_2 \times 2\text{H}_2\text{O}$, 85 μM $\text{FeCl}_3 \times 6\text{H}_2\text{O}$, 1.3 μM $\text{ZnSO}_4 \times 7\text{H}_2\text{O}$, 1.3 μM $\text{CuSO}_4 \times 5\text{H}_2\text{O}$, 1.8 μM $\text{MnSO}_4 \times \text{H}_2\text{O}$, 1.5 μM $\text{CoCl}_2 \times 6\text{H}_2\text{O}$, 108 μM EDTA, 50 $\mu\text{g/ml}$ kanamycin) to an OD_{600} of 1.5. Then protein expression was induced by addition of 1mM IPTG. The production went on for 4 hours. Samples for SDS-PAGE analysis were taken before induction, and 1, 2, 3 and 4 hours after induction.

Purification

Bacteria from the cultures were collected by centrifugation at 5000 rpm, 15 minutes. The bacterial pellets were lysed by five freeze-thaw cycles, diluted 3 times in 20 mM

- Tris-HCl pH 8.0 and sonicated. The inclusion bodies (IB's) were collected by centrifugation at 6000 rpm, 30 minutes, and washed by three more cycles of resuspension and centrifugation. For extraction, the IB's were resuspended in 6M guanidine-hydrochloride, 20 mM Tris-HCl pH 8.0 (6M Gu-HCl) and incubated with stirring overnight at +4°C.
- 5 The extract was clarified by high-speed centrifugation at 26000 g, 60minutes. The supernatant was saved for further clarification, while the pellet went through another cycle of extraction. The supernatants of the extractions were combined and further clarified by depth filter filtration (K700P filter laid on top of KS50P filter, Pall Corp., USA). A1M in the clarified extract was purified, using a Ni-agarose resin (Sigma-Aldrich, USA).
- 10 Briefly, the resin was packed into a 10-ml disposable chromatography column (Bio-Rad, USA) and equilibrated in 6M Gu-HCl. The A1M extract was applied onto the column using free-flow and the flow-through collected. The column was washed with five column volumes of 6M Gu-HCl and then eluted with four volumes of 6M Gu-HCl +0.5 M imidazole. Starting material, flow-through and eluted fractions were precipitated with
- 15 ethanol, resolved in 1xSDS-PAGE loading buffer and separated by SDS-PAGE. The purification procedure was repeated if the flow through contained significant amounts of A1M. The protein content of the extract was determined by absorbance 280nm.

- The Ni-agarose eluates were diluted to an approximate A_{280} of 5.0 and cooled to +4°C
- 20 before refolding. The eluate was then mixed with 2/3 volumes of 0.275 M L-cystein in 20 mM Tris-HCl pH 9.5 + 0.1M NaCl. Then 16.7 volumes refolding buffer were quickly added. The final buffer concentration was: (0.2 mg/ml A1M, 0.1M Tris, 0.6 M NaCl, 0.45 M L-arginine, 2 mM EDTA, 10 mM L-cystein and 1 mM L-cystein, pH 9.5). The mixture was then stirred for 1 hour at +4°C, and the solution concentrated to the initial
- 25 volume of the A1M solution using Centricon plus 70, 10K ultrafiltration devices (Merck Millipore; USA). After concentration, the A1M solution was diafiltrated to 20mM Tris-HCl pH 8.0 using the same devices, by 10 consecutive 2.5x dilution/concentration cycles. After diafiltration the solution was clarified by centrifugation at 15000g for 15 minutes and then run through a 0.2µm filter.

30

- The refolded A1M was immediately applied to a 5ml Bio-Scale Mini UNOsphere Q Cartridge (Bio-Rad), equilibrated with 20mM Tris-HCl pH 8.0. The column was run on an ÄKTA purifier 10 instrument (GE Healthcare, USA) according to Bio-Rad's instruction for the cartridges. After sample application the column was washed with five column
- 35 volumes (CV) of 20mM Tris-HCl pH 8.0 before elution with a 20 CV linear gradient from 0-0.35M NaCl. Finally, the column was washed with three CVs of 1 M NaCl. The flow-

though and selected fractions collected during the linear gradient were analysed by SDS-PAGE. Flow-through with remaining A1M was immediately re-run on a new column and the A1M containing fractions were pooled, concentrated to 100µM, sterile filtered and frozen at -20°C in aliquotes.

5

Gel electrophoresis

SDS-PAGE was run according to Laemmli (Laemmli, 1970) using standard protocols. Proteins were separated on stain-free 4-20% TGX gels (Bio-Rad) at 300V for 17 minutes. Native PAGE was run without SDS and without reducing agents on stain-free
10 4-20% TGX gels at 200V for 40 minutes. The gels were analysed on a Chemidoc MP instrument (Bio-Rad).

Circular Dichroism

The circular dichroism spectra were recorded on a Jasco-J180 spectrofluorimeter instrument (JASCO Inc., Japan) of 10 µM solutions in 20 mM Tris-HCl pH 8.0 + 0.15 M
15 NaCl in a 2-mm cuvette. The solutions were scanned at +22°C between 190-260 nm. Three runs were overlayed for each sample. The percentage of α-helix and β-sheet structure was calculated using the <http://k2d3.ogic.ca> software.

20 *SEC-FPLC*

Proteins were analysed by size exclusion on an ÄKTA purifier 10 instrument using a 24-ml Superose 12 10/30 GL column (GE Healthcare). The column was equilibrated with 20 mM Tris-HCl pH 8.0 + 0.15 M NaCl using a flow-rate of 1ml/min. 100-200 µg of protein were loaded onto the column in a volume of 100 µl and eluted with 20 mM Tris-
25 HCl pH 8.0 + 0.15 M NaCl using a flow-rate of 0.75 ml/min. Typically, monomeric A1M was eluted after 15 ml/20 min, dimeric after 13-14 ml/18-19 min, and large aggregates were eluted after 8 ml/12 min. The percentage of large aggregates was calculated from the area under the 8-ml peak compared to the total peak area. The percentage of total protein retrieved on the column after stress treatments (see below) was calculated by
30 comparing the total peak areas of treated vs. non-treated samples.

RP-HPLC

Reversed phase HPLC was run on an Agilent 1260 Infinity Binary LC system using an Aeris Wildpore 3.6 µM XP-C8 column (Phenomenex Inc., USA). The column was run at
35 +25°C using a flow rate of 1 ml/min and equilibrated with a mixture of 70% H₂O+0.1% TFA and 30% Acetonitrile+0.1% TFA. 10 µl (=10 µg of protein) were loaded and eluted

with a linear gradient of 30-50% acetonitrile over 20 minutes. The column was regenerated by washing with 95% acetonitrile for 10 minutes.

Dynamic light scattering

- 5 Dynamic light scattering (DLS) analysis of non-stressed and shearing-stressed samples was done using the service of SARomics Biostructure AB, Lund. A1M samples, diluted to 10 μ M in 10mM Tris-HCl pH 8.0 + 0.125 M NaCl, were analysed on a Malvern APS instrument at +20°C. Samples were prepared in duplicates and each sample was monitored three times.

10

Differential scanning fluorimetry

- Thermostability of the A1M variants was analysed by differential scanning fluorimetry (DSF) using the service of SARomics Biostructure AB, Lund. A1M diluted to 4.4 μ M in 10 mM HEPES pH 8.0 + 0.125 M NaCl was mixed with SYPRO orange (1000x dilution of SYPRO orange in total). The analysis was made in duplicates and the average melting temperature (T_m) was calculated.
- 15

Introduction of stress by shearing forces

- 10 μ l of 100- μ M A1M solutions were exposed to shearing force stress by 80 pipettings with a multiple channel pipett using 0-10 μ l pipett tips. The stress-treatment was performed in duplicates. After pipetting, the duplicates were combined and diluted 10 times before analysis with DLS as described.
- 20

High concentration-induced stress

- 25 Solubility and stability of A1M was analysed at high concentration. 500 μ l of 100 μ M A1M solutions were concentrated tenfold to 50 μ l using Amicon Ultra-0.5, 10K devices (Merck Millipore) by centrifugation at 14000g for 10 min. After concentration the volumes were corrected to exactly 50 μ l using the respective flowthroughs. Concentrated and non-concentrated samples (10 μ g) were compared side-by-side on native PAGE.
- 30 The influence of different buffers were examined by diafiltration of the samples before concentration. This was done by five cycles 10-time dilution/concentration in Amicon Ultra-15, 10K devices.

Quantification of free thiol groups

- 35 The free thiol groups of the A1M variants were determined with the "Thiol and sulfide quantification kit" from Molecular probes. The assay was performed in 96-well plates

according to the kit manual. A volume of 3µl standard or A1M (100µM) was mixed with 3µl cystamine work solution, 100µl papain-SSCH, 100µl L-BAPNA. The assay was read as absorbance 405nm.

5 *ABTS reduction assay*

The assay is a modification of (Åkerström et al 2007). 7mM of 2,2 Azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) di-ammonium salt (Sigma-Aldrich) was oxidized with 2.45mM K₂S₂O₈ overnight, and then diluted to 56µM in 25mM sodium phosphate buffer pH 8.0. 100µl of the ABTS working solution was added per well in a 96-well plate. A
10 time-point zero measurement was done at A₄₀₅ using a Perkin Elmer Plate reader. 2µl of an A1M solution (0-100 µM) were quickly added by a multiple channel pipett. The kinetics of the decrease in absorbance at 405nm was quickly followed for 95s. For practical reasons only 8wells were analysed at the time. The A1M dilution series was run in duplicate or triplicates. If the number of samples to be analysed required several
15 plates, new ABTS stock was diluted into working stock for each plate and an wt-A1M reference sample was included in all plates. The absorbance for each concentration was plotted against time and the area under the curve (AUC) was calculated for each concentration. The net AUC was calculated by subtracting the AUC of buffer only.

20 *Oxygen radical antioxidant capacity (ORAC) assay*

The commercial kit OxySelect™ Oxygen radical antioxidant capacity (ORAC) activity assay (Cell Biolabs, Inc. USA) is based on the oxidation and destruction of a fluorescent probe by peroxy radicals. When an antioxidant is present this destruction is inhibited. As standard the water soluble vitamin E derivate, Trolox is used. Performance,
25 analysis and calculations followed the kit manual and an A1M concentrations of 2.5-5µM fitted nicely in to the standard curve. Ovalbumin (Sigma-Aldrich) was used as a negative protein control.

Cytochrome c reduction assay

30 The assay is modified from (Allhorn et al. 2005). A working solution was made by mixing 100µM cytochrome c (Sigma-Aldrich) and 100µM NADH in 10mM Tris-HCl pH 8.0 + 0.125M NaCl. 11µl of an A1M solution (0-100 µM) were added to a 96-well plate in duplicates. 100µl of the cytochrome c working solution was quickly added per well using a multichannel pipett. The kinetics of the increase in absorbance at 550nm was followed
35 for 20 minutes. One plate was analysed at the time. No biases over time could be observed. If several plates were to be analysed the same working solution was used for

all without any observable artefacts caused by this procedure. After measurements the results were analysed as described for the ABTS assay.

The red-shift assay

- 5 The incorporation of free heme in A1M yields a red-shift of the Soret band absorbance peak (Karnaukhova et al., 2014; Rutardottir et al., 2016), and was evaluated in the A1M-mutants as the A_{413}/A_{386} ratio. 44 μ M A1M in 20mM Tris-HCl pH 8.0 + 0.25M NaCl was mixed with 40 μ M free heme (Applichem). The incubations were done in duplicates and incubated for 2 hours at room temperature. The red-shift was analysed in four
- 10 times diluted samples by wavelength scan in a Beckman Spectrophotometer. The average maximal peak of the duplicates as well as the average ratio between the absorbances at 413 and 386nm was calculated. Ovalbumin and buffer only were used as negative controls.
- 15 *Specific binding of A1M to heme-agarose*
- Binding of A1M to heme-agarose was analysed by the method of (Larsson et al., 2004), modified as described (Rutardottir et al., 2016). Heme-agarose (Sigma-Aldrich) and control Sepharose 4B (Sigma-Aldrich) were equilibrated with 20mM Tris-HCl pH 8.0 + 0.25M NaCl and prepared as 50% slurries. 75 μ l of an A1M dilution series (0-
- 20 13.3 μ M) were added to duplicate 96-well plates (one plate for heme agarose and one for control Sepharose). 20 μ l of heme-agarose och control Sepharose slurry were added to the wells by careful pipetting to assure transfer of similar amounts, and incubated for 30 minutes at RT during rotational stirring. The mixtures were carefully transferred to an AcroPrep Advance 96-well filter plate, 1.2 μ m Supor membrane (Pall Corp).
- 25 The plates were centrifuged for 2min at 1000g collecting the flow-through in a low-binding microplate. 25 μ l of each flow-through, as well as the non-incubated starting material, were assayed for protein content using Pierce BCA Protein Assay kit (Thermo Scientific Inc., USA). Protein amounts were recalculated as amount bound (added minus amount in flow-through) for both the heme- and control Sepharose-incubated samples.
- 30 After subtraction of the amount bound to the control Sepharose, the amount specifically bound to heme agarose was plotted against the added amount. The slope of the curve was calculated with linear regression. The SD of the duplicates were used to evaluate the significance of the differences between the A1M variants.
- 35 *Heme binding of M8H4DK-wt A1M (Wt-A1M) and M8H4DK-35-A1M (35-A1M)*

Heme binding was analysed as previously described by native PAGE (Karnaukhova et al., 2014) and UV-spectrophotometry (Ruttarsdottir et al., 2016). Briefly, for native PAGE, A1M and various concentrations of heme were incubated in Tris-buffer, pH 8.0 for 30 min at room temperature, separated by native PAGE on stain-free 12% Criterion
5 TGX gels at 200V for 40 minutes. The gel bands were analysed on a Chemidoc MP instrument (Bio-Rad) for tryptophan fluorescence using the Stain-free setting, stained with Coomassie Brilliant Blue, destained and imaged again on the Chemidoc using the Coomassie setting. Both sets of bands were then quantified using Image Lab™ Software (Bio-Rad). Heme binding was then estimated as the amount of quenching of
10 the tryptophan fluorescence relative to total protein amounts after Coomassie staining. Absorbance spectra were measured on a Beckman (Beckman Instruments, Fullerton, CA) DU 800 spectrophotometer using a scan rate of 600 nm/min in the UV-VIS region between 250 and 700 nm at 22°C. Concentrations of A1M and heme were 32 and 19 µM, respectively, in 20 mM Tris-HCl, pH 8.0, 0.15 M NaCl. Heme was added from a
15 stock solution of 10 mM in DMSO. Protein solutions were scanned 2 h after mixing.

Plasma clearance and biodistribution

For the plasma clearance studies, each A1M-variant was injected intravenously (i.v) in six male Wistar rats (5.0 mg/kg; stock-solutions in 20 mM Tris-HCl, pH 8.0) and blood
20 samples were taken in EDTA-tubes at 1, 5, 15, 30 min, and 1, 3, 6, 16 and 24 h post-injection, using different sampling intervals in groups of three rats to avoid over-sampling. Plasma was aspirated after centrifugation 140xg for 10 min, and the concentration of A1M was determined by radioimmunoassay (RIA) as described (Gram et al. 2015) using the specific A1M-variant as standard. For the biodistribution studies, the
25 A1M-variants were injected i.v. in C57BL/6NRj mice (5.0 mg/kg; stock-solutions in 20 mM Tris-HCl, pH 8.0). The mice were sacrificed 10 min post-injection (n=3), and after 30 min (n=3). Organs were sampled, weighed and homogenized in 5:1 (vol:weight) Cell Extraction Buffer (Invitrogen, cat. No. FNN0011), containing 50 µl/ml cOmplete Mini, EDTA-free proteinase inhibitor cocktail tablets (Roche, cat. no. 11836170001). A1M-
30 concentrations were determined by an in-house sandwich ELISA. Briefly, 96-well microtiter plates were coated overnight at 4°C with mouse monoclonal anti-A1M (clone 35.14, 5 µg/ml in PBS), washed, and then incubated with A1M-standards (human urinary A1M, purified as described at our laboratory (Åkerström et al., 1995) or homogenised tissue samples, diluted in incubation buffer (PBS + 0.05% Tween 20 + 0.5% bo-
35 vine serum albumin), for 60 min at RT. After washing, the wells were incubated with horseradish peroxidase-conjugated mouse monoclonal anti-A1M (clone 57.10, 5 ng/ml

in incubation buffer) for 60 min at RT. The plates were washed and developed by incubating with SureBlue TMB Microwell Peroxidase Substrate (KPL) in the dark for 20 min, and finally stopped with 1M sulfuric acid. Absorbance was read at 450 nm in a Wallac 1420 Multilabel Counter. The two mouse monoclonal antibodies were raised by

- 5 AgriSera AB (Vännäs, Sweden) against human urinary A1M. The ELISA was specific for human A1M, did not cross-react with endogenous mouse plasma A1M, and reacted with human urinary A1M, M8H4DK-wt A1M (wt-A1M) and M8H4DK-35-A1M (35-A1M) equally well.

10 *Rhabdomyolysis induced kidney damage*

- This study was approved by the ethical committee for animal studies in Malmö-Lund, no. M21-15. Female C57BL/6 mice with a body weight of 20.5 ± 0.7 g were obtained from Taconic (Denmark), housed in Type III cages with wire lids, at a constant room temperature with 12 hour light dark cycles. Temperature ($20 \pm 0.5^\circ\text{C}$) and relative moist
- 15 ($50 \pm 5\%$) was maintained throughout the studies. All animals had free access to food (RM1(E) SQC, SDS, England), tap water and cage enrichment. After overnight water deprivation (15 hours) animals were weighed and anaesthetized using isoflurane and then allocated to the following four groups: 1) Control (n=6), received no intramuscular (i.m.) or intravenous (i.v.) administration; 2) Glycerol (n=10), received 50% sterile glycerol (Teknova, Hollister, CA, USA) i.m. (2.0 ml/kg body weight, single dose, divided into
- 20 both hind limbs); 3) Glycerol + M8H4DK-wt A1M (wt-A1M) (n=10), received wt-A1M i.v. (7 mg/kg body weight, single dose) 30 minutes after glycerol i.m. (2.0 ml/kg body weight) administration; and 4) Glycerol + M8H4DK-35-A1M (35-A1M) (n=10), received 35-A1M i.v. (7 mg/kg body weight, single dose) 30 minutes after glycerol i.m. (2.0 ml/kg
- 25 body weight) administration. Following i.m. administration animals were placed on a heat pad during awakening and then put back to their cages and supplied with free access to food and water. After 4 hours (post-glycerol injection) animals were anaesthetized using isoflurane and kidney collected for RNA extraction followed by mRNA evaluation as described below.

30

RNA isolation and real-time PCR

- Total RNA was isolated from HK-2 cells, using Direct-zol™ RNA MiniPrep supplied by Zymo Research (Irvine, CA, USA), or mouse kidneys, using NucleoSpin RNA/Protein (Machery-Nagel, Duren, Germany) followed by RNeasy® Mini Kit (QIAGEN, German-
- 35 town, MD, USA). The OD ratio (optical density at 260 nm/280 nm) of RNA was always higher than 1.9. Reverse transcription was performed according to the manufacturer on

1.0 µg total RNA using iScript™ cDNA Synthesis Kit (Bio-Rad, CA, USA). RT² qPCR Primer Assay (human (HK-2 cells) and mouse (kidneys) primers from QIAGEN) were used to quantify the mRNA expression of heme oxygenase 1 (HO-1) and heat shock protein 70 (Hsp70). Data were normalized to glyceraldehyde-3-phosphate dehydrogenase (human (HK-2 cells) and mouse (kidney) GAPDH, RT² qPCR Primer Assay from QIAGEN). Data are presented as columns, displaying mean ± SEM, for *in vitro* data and box plots, displaying medians and 25th and 75th percentiles, for *in vivo* data. The fold change values were calculated by normalizing against control samples from untreated cells or animals (controls). Expression was analyzed using iTaq™ Universal SYBR® Green Supermix (Bio-Rad). Amplification was performed as described by the manufacturer (Bio-Rad) for 40 cycles in an iCycler Thermal Cycler (Bio-Rad) and data analyzed using iCycler iQ Optical System Software (Bio-Rad).

Rescue of K562 cells from heme induced cell death

A1M was previously shown to inhibit heme-induced cell-death of human erythroid K562 cells (Olsson et al., 2008). The cells were cultured in DMEM with glutamax + 10% FCS and antibiotics (Gibco, Life Technologies Corp., USA) according to the instructions at ATCC. Cells were washed and resuspended in DMEM without phenol red and FCS but supplemented with glutamax I and antibiotics (Gibco). Cells were seeded into 96-well plates, 10⁵ cells per well, and exposed to 100µM heme in the presence of a 0-10 µM A1M dilution series. As a positive control for cell death, 10µl of lysis solution from the LDH detection kit (see below) was added. The cells were incubated in a 37°C CO₂-incubator for 1 hour. The plates were quickly centrifuged at 350g for 4 minutes before 50µl of the medium was transferred to a 96-well microplate for analysis of LDH release using the cytoTox 96® Non-Radio. Cytotoxicity Assay (Promega Biotech AB, Sweden) according to the manufacturer's instructions. Heme-induced cells typically gave 7 times higher signal compared to live cells and 70% of the signal of completely lysed cells. The average signal of cells incubated without heme- or A1M-addition was subtracted from all and the signal of heme only-incubated cells was set to 100%. All other signals were related to this value. This procedure enabled comparison of several independent experiments.

Protection of HK-2 cells

Human kidney cortex proximal tubule epithelial cells (HK-2, ATCC® CRL-2190, ATCC, Teddington, UK) were cultured in keratinocyte serum free medium (K-SFM) supplemented with bovine pituitary extract (BPE, 0.05 mg/ml) and epidermal growth

factor (5 ng/mL)(all from Invitrogen, Paisley, UK). When cells reached approximately 80-90% confluence, heme (0-30 μ M, from a freshly prepared 10 mM stock solution) or a mixture of $(\text{NH}_4)\text{Fe}(\text{SO}_4)_2$, hydrogen peroxide, and ascorbate (0-200 μ M, the Fenton reaction), with or without the simultaneous addition of A1M (0-20 μ M, M8H4DK-wt A1M (wt-A1M) and M8H4DK-35-A1M (35-A1M)), were added and cells were incubated for 6 hours. After incubation, cells were analyzed for cell viability using WST-1 (the measured metabolic activity of cells, e.g. measurement of cellular cleavage of the WST-1 stable tetrazolium salt to the soluble formazan dye, is a direct correlate to the number of viable cells)(Roche Diagnostics GmbH, Mannheim, Germany) according to the instructions from the manufacturer. The cell viability was normalized against control samples from untreated cells. Parallel cultures were harvested using Qiazol™ Lysis reagent (for RNA extraction, QIAGEN, Germantown, MD, USA). Total RNA was extracted from cells to evaluate mRNA expression as described below.

15 Statistics

Comparisons between unrelated groups were performed by ANOVA with *post hoc* Bonferroni correction. P-values <0.05 were considered significant.

Results and Discussion

20

Project overview

The purpose of this investigation is to identify more stable and soluble variants of A1M with preserved functional properties. The project was performed in three phases. In phase I, A1M of different species and point mutations of human A1M were screened to identify individual amino acids that had a positive effect on stability without compromising function. In phase II, amino acids shown to have positive effect in phase I, were combined to find even better combinations. Finally in phase III, data from phase I and II were confirmed and the effect of different N-terminal tags was investigated. In all three phases, the proteins were expressed in the same E coli system, using the same vector and purification protocol. All variants were expressed, purified and analysed in parallel, using similar protocols and procedure, within each phase. The analysis panel of each phase is summarized in Table 1. Amino acid and DNA sequence of all constructs can be found below.

35 Table 1. Analyses performed in the different phases of the A1M variant project

	Analysis	phase I	phase II	phase III
1	SDS-PAGE (identify, purity)	x	x	x
2	Absorbance 280nm (quantity)	x	x	x
3	Circular Dichroism (identity)	x		
4	PAGE (aggregation analysis)	x	x	x
5	SEC-FPLC (aggregation analysis)		x	
6	DLS (aggregation analysis)	x		x
7	RP-HPLC (purity, aggregation analysis)			x
8	DSF (thermostability)	x		x
9	Shearing stability (DLS analysis)	x		x
10	Concentration stability (PAGE analysis)	x	x	x
11	Concentration stability other buffers (PAGE analysis)			x
12	Stability freeze thaw, prolonged storage in fridge, RT, +37°C (SEC-FPLC, ABTS reduction)*		x	
13	Quantification of free thiol groups	x		
14	Reduction capacity of the ABTS radical	x	x	x
15	Oxygen radical anti-oxidant capacity (ORAC)			x
16	Reduction capacity of cytochrome c			x
17	Reduction/binding capacity of free heme	x	x	x
18	Binding capacity of heme-agarose	x	x	x
19	Rescue of K562 cells from heme-induced cell death		x	x

*Only performed on human wt and the M8H4DK-R66H+N1796D variant.

Amino acid sequence of all constructs:

- 5 60. M8H4DK-Human wt (rhA1M) (SEQ ID NO: 9)
MHHHHHHHHHDDDDKGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLKKIM-
DRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTCKFSRHHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPVGEQEPEPILIPR
- 10 1. M8H5GIEGR-Mouse (SEQ ID NO: 24)
MHHHHHHHHHGGGGGIEGRDPASTLPDIQVQENFSESRI-
YGKWYNLAVGSTCPWLSRIKDKMSVQTLVLQEGATETEISMTSTRWRRGVCEEIT-
GAYQKTDIDGKFLYHKSKWNITLESYVVHTNYDEYAIFLTCKSSHHHGLTI-
15 TAKLYGREPQLRDSLLQEFKDVALNVGISENSIIFMPDRGECPVGDREVEPTSIAR
2. M8H5GIEGR-Naked Mole rat (SEQ ID NO: 25)
MHHHHHHHHHGGGGGIEGRNPVPMPPDNIQVQENFDESRI-
20 YGKWFNLATGSTCPWLKRIKDRLSVSTMVLGKGTTTE-
TQISTTHTHWRQGVQCETSGVYKKTDTAGKFLYHKSKWNVTMESYV-
VHTNYDEYAIILTCKFSHHHGPTITAKLYGREPRLRDSLLQEFREMGVGVGIPEDSIFT-
MANRGECPVGDQAPESTPAPR

3. M8H5GIEGR-Frog (SEQ ID NO: 26)
MHHHHHHHHHGGGGGIEGRCSPIQPEDNIQIQENFDLQRIYGKWDIAIGSTCK-
WLKHHKEKFNMGTLELSDGETDGEVRIVNTRMRHGTCSQIVGSYQKTETPGKFDYF-
5 NARWGTTIQNYIVFTNYNEYVIMQMRKKKGSETTTTVKLYGRSPDLRPT-
LVDEFRRQFALAQGIPEDSIVMLPNNGECSPGIEVRPRR
4. M8H5GIEGR-Chicken (SEQ ID NO: 27)
10 MHHHHHHHHHGGGGGIEGRTPVGDQDEDIQVQENFEPERMYGKWYDVAVGTTCK-
WMKNYKEKFMSGTLVLGPGPSADQISTISTRRLRQGDCKRVSGEYQKTDTPGKYTYYN
PKWDVSIKSYVLRNTYEEYAVILMKKTSNFGPTTTTLKLYGRSPELREEL-
TEAFQQLALEMGIPADSVFILANKGECVPQETATAPER
- 15 5. M8H5GIEGR-Rabbit (SEQ ID NO: 28)
MHHHHHHHHHGGGGGIEGRDPVPTLPDDIQVQENFELSRI-
YGKWYNLAVGSTCPWLKRIKDRMAVSTLVLGEGTSETEISMTSTHWRRGVCEE-
ISGAYEKTDTDGKFLYHKAKWNLTMESYVVHTNYDEYAIFLTCKFSRRHGPTI-
TAKLYGREPQLRESLLQEFREVALGVGIPENSIFTMIDRGECVPGQQEPKPAPVLR
6. M8H5GIEGR-SQ Monkey (SEQ ID NO: 29)
MHHHHHHHHHGGGGGIEGRSPVPTPPEGIQVQENFNLSRIYGKWYNLAIGSTCPWLK-
KIMDRCLKVSTLVLEEGATEAEISMTSTRWRKGFCEQTSWAY-
25 EKTDTDGKFLYHEPKWNVTMESYVAHTNYEEYAIFLTCKFSRHHGPTI-
TAKLYGREPQLRESLLQDFRVVAQGVGIPEDSIFTMANRGECVPGEQEPQPILHRR
7. M8H5GIEGR-Walrus (SEQ ID NO: 30)
30 MHHHHHHHHHGGGGGIEGRSPVLTPPDIAIQVQENFDISRIYGKWFH-
VAMGSTCPWLKKFMDRMSMSTLVLGEGATDGEISMTSTRWRRGTCEEISGAY-
EKTSTNGKFLYHNPKWNITMESYVVHTDYDEYAIFLTCKFSRHHGPTI-
TAKLYGRQPQLRESLLEEFRELALGVGIPEDSIFTMANKGECVPGEQEPEPSPHMR
- 35 8. M8H5GIEGR-Manatee (SEQ ID NO: 31)
MHHHHHHHHHGGGGGIEGRSPVKTPLNIDIQVQENFDLPRIYGKWFNIAIG-
STCQWLKRLKAGPTMSTLVLGEGATDTEISTTSTRWRKGFCEEISGAYEKTD-
TAGKFLYHGSKWNVLTLESYVVHTNYDEYAIFLTCKFSRYGLTI-
TAKLYGRQPQVRESLLEEFREFALGVGIPEDSIFTTADKGECVPGEQEPEPTAALR
9. M8H5GIEGR-Plaice (SEQ ID NO: 32)
MHHHHHHHHHGGGGGIEGRLPVLPEP-
LYPTQENFDLTRFVGTWHDVALTSSCPHMQRN-
45 RADAAIGKLVLEKDTGNKLVTRTRLRHGTCEVMSGEYELTSTPGRIFYHIDRW-
DADVDAYVVHTNYDEYAIIMSKQKTSNGENSTSLKLYSRTMSVRDVLDDFKTLVRHQ
GMSDDTIIIKQNKGDICIPGEQVEEAPSQPEPKR
10. M8H5GIEGR-Orangutan (SEQ ID NO: 33)
50 MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIM-
DRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHKSKWNIT

MESYVVHTNYDEYAIFLTKKFSRRHGPTITAKLYGRAPQLRETLLQDFRVVAQGVGI-
PEDSIFTMADRGECPGEGEPEPILIPR

11. M8H5GIEGR-Human P35K (SEQ ID NO: 34)

5 MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCKWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRRHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

10

12. M8H5GIEGR-Human M41K (SEQ ID NO: 35)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRRHGPTI-
15 TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

13. M8H5GIEGR-Human R66H (SEQ ID NO: 36)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
20 KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRRHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

25 14. M8H5GIEGR-Human T75K (SEQ ID NO: 37)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEEKSGAYEKDSTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRRHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

15. M8H5GIEGR-Human T75Y (SEQ ID NO: 38)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEEYSGAYEKDSTDGKFLYHK-
35 SKWNITMESYVVHTNYDEYAIFLTKKFSRRHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

16. M8H5GIEGR-Human M99K (SEQ ID NO: 39)

40 MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRRHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

45

17. M8H5GIEGR-Human S101Y (SEQ ID NO: 40)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRRHGPTI-
50 TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

18. M8H5GIEGR-Human K69.92.118.130R (SEQ ID NO: 41)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRGVCETSGAY-
EKTDTDGKFLYHRSKWNITMESYVVHTNYDEYAIFLTKRFSRHHGPTITAR-
LYGRAPQLRETLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

19. M8H5GIEGR-Coelacanth (SEQ ID NO: 42)

MHHHHHHHHHGGGGGIEGRGSPRLRDEDIQVQENFDLPRIYGKWEIAI-
ASTCPWVKNHKDKMFMGMTMVLQEGEQSDRISTTSTRIRDGTCSQITGYTLTTTPGK-
10 FAYHNSKWNLDVNSYVVHTNYDEYSIVMMQKYKSS-
NSTTTVRLYGRTQELRDSLHAEFKKFALDQGIDEDSIYILPKRDECPGEPKAESLMAR

21. M8H5GIEGR-Human L89T (SEQ ID NO: 43)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
15 KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFTYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

20 22. M8H5GIEGR-Human N1796D (SEQ ID NO: 44)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHK-
SKWDITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

23. M8H5GIEGR-Human T45K (SEQ ID NO: 45)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMKVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHK-
30 SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

24. M8H5GIEGR-Human A135E (SEQ ID NO: 46)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
35 KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGREPQLRETLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

40

25. M8H5GIEGR-Human V170S (SEQ ID NO: 47)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
45 TAKLYGRAPQLRETLQDFRVVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

26. M8H5GIEGR-Human V148D (SEQ ID NO: 48)

MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
50 KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLQDFRDVAQGVGIPEDSIFTMADRGECPGEGEPEPILIPR

27. M8H5GIEGR-Human G172Q (SEQ ID NO: 49)
MHHHHHHHHHGGGGGIEGRGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLK-
KIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKDSTDGKFLYHK-
5 SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPVGEQEPEPILIPR
33. M8H4DK-Human M41K+R66H (SEQ ID NO: 50)
10 MHHHHHHHHHDDDDKGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLKKIK-
DRMTVSTLVLGEGATEAEISMTSTHWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPVGEQEPEPILIPR
- 15 34. M8H4DK-Human M41K+N1796D (SEQ ID NO: 51)
MHHHHHHHHHDDDDKGPVPTPPDNIQVQENFDISRIYGKWYNLAIGSTCPWLKKIK-
DRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWDITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
20 TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPVGEQEPEPILIPR
35. M8H4DK-Human R66H+N1796D (SEQ ID NO: 52)
MHHHHHHHHHDDDDKGPVPTPPDNIQVQENFDISRIYGKWYNLAIGSTCPWLKKIM-
DRMTVSTLVLGEGATEAEISMTSTHWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWDITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
25 TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPVGEQEPEPILIPR
36. M8H4DK-Human M41K+R66H+N1796D (SEQ ID NO: 53)
MHHHHHHHHHDDDDKGPVPTPPDNIQVQENFDISRIYGKWYNLAIGSTCPWLKKIK-
DRMTVSTLVLGEGATEAEISMTSTHWRKGVCEETSGAYEKDSTDGKFLYHK-
30 SKWDITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPVGEQEPEPILIPR
37. M8H4DK-Human M41K (SEQ ID NO: 8)
MHHHHHHHHHDDDDKGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLKKIK-
35 DRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPVGEQEPEPILIPR
38. M8H4DK-Human R66H (SEQ ID NO: 54)
40 MHHHHHHHHHDDDDKGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLKKIM-
DRMTVSTLVLGEGATEAEISMTSTHWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPVGEQEPEPILIPR
- 45 39. M8H4DK-Human N1796D (SEQ ID NO: 55)
MHHHHHHHHHDDDDKGPVPTPPDNIQVQENFDISRIYGKWYNLAIGSTCPWLKKIM-
DRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKDSTDGKFLYHK-
SKWDITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
50 TAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPVGEQEPEPILIPR

40. M8H-Human wt (SEQ ID NO: 56)

MHHHHHHHHHGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLKKIM-
DRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHK-
5 SKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLQDFRVVAQGVGIPEDSIFTMADRGECPGEQEPEPILIPR

41. M8H-Human R66H+N1796D (SEQ ID NO: 57)

MHHHHHHHHHGPVPTPPDNIQVQENFDISRIYGKWYNLAIGSTCPWLKKIM-
10 DRMTVSTLVLGEGATEAEISMTSTHWRKGVCEETSGAYEKTDTDGKFLYHK-
SKWDITMESYVVHTNYDEYAIFLTKKFSRHHGPTI-
TAKLYGRAPQLRETLQDFRVVAQGVGIPEDSIFTMADRGECPGEQEPEPILIPR

42. untagged-Human R66H+N1796D (SEQ ID NO: 3)

15 MGPVPTPPDNIQVQENFDISRIYGKWYNLAIGSTCPWLKKIMDRMTVSTLVLGEGA-
TEAEISMTSTHWRKGVCEETSGAYEKTDTDGKFLYHKSKWDITMESYV-
VHTNYDEYAIFLTKKFSRHHGPTITAKLYGRAPQLRETLQDFRVVAQGVGIPEDSIFT-
MADRGECPGEQEPEPILIPR

20 61 untagged-Human wt (SEQ ID NO: 2)

MGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLKKIMDRMTVSTLVLGEGA-
TEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHKSKWNITMESYV-
VHTNYDEYAIFLTKKFSRHHGPTITAKLYGRAPQLRETLQDFRVVAQGVGIPEDSIFT-
MADRGECPGEQEPEPILIPR

25

DNA sequence of all constructs:

60. M8H4DK-Human wt (SEQ ID NO: 58)

ATGCATCACCATCACCATCACCATCACGATGACGATGACAAGGGCCCTGT-
30 GCCAACGCCGCCCCGACAACATCCAAGTGCAGGAAAACCTTCAATATCTCTCG-
GATCTATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCCTGG-
CTGAAGAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGG-
CGCTACAGAGGCGGAGATCAGCATGACCAGCACTCGTTGG-
CGGAAAGGTGTCTGTGAGGAGACGTCTGGAGCTTATGAGAAAACAGATACT-
35 GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTGAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
40 GGTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

1.M8H5GIEGR-Mouse (SEQ ID NO: 59)

45 ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGACCCTGCGTCAACACTGCCAGATATCCAGGTTGAGGAGAACTTCAGTGAG
TCCCGGATCTATGGAAAATGGTACAACCTGGCGGTGGGATCCACCTGCCCGTGG-
CTGAGCCGCATTAAGGACAAGATGAGCGTGAGCACGCTGGTGCTGCAGGAGGGG

GCGACAGAAACAGAGATCAGCATGACCAGTACTCGATGGCGGAGAGGTGTCTGT-
GAGGA-
GATCACTGGGGCGTACCAGAAGACGGACATCGATGGAAAGTTCCTCTACCACAAA
TCCAAATGGAACATAACCTTGGAATCCTATGT-
5 GGTCCACACCAACTATGACGAATATGC-
CATTTTCCTTACCAAGAAGTCCAGCCACCACCACGGGCTCACCATCACTGCCAAG
CTCTATGGTCGGGAGCCACAGCTGAGGGACAGCCTTCTGCAGGAGTTCAAG-
GATGTGG-
CCCTGAATGTGGGCATCTCTGAGAACTCCATCATTTTTATGCCTGACAGAGGGGA
10 ATGTGTCCCTGGGGATCGGGAGGTGGAGCCACATCAATTGCCAGATGA

2. M8H5GIEGR-Naked Mole rat (SEQ ID NO: 60)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCAATCCTGTGCCGATGCCGCCAGACAACATCCAAGTGCAGGA-
15 GAACTTTGATGAATCCCGGATCTATGGGAAATGGTTCAACCTGGCTACGGG-
CTCCACGTGCCCCGTGGCTGAAGAGGATCAAAGACAGGCTGAGTGT-
GAGCACAATGGTGCTGGGCAAGGGGACCACGGA-
GACACAGATCAGCACAACCCACACCCACTGGCGGCAAGGGGTGTGCCAGGA-
GACCTCAGGGGTTTACAAGAAAACAGACACGG-
20 CTGGGAAGTTCCTCTACCACAAGTCCAAATGGAATGTAACCATGGAGTCCTATGT-
GGTCCACACCAACTATGATGAGTATGC-
CATCATTCTAACTAAGAAGTTCAGCCACCACCATGGACCGAC-
CATTACTGCCAAGCTCTATGGGAGAGAGCCGCGGCTGAGA-
GACAGCCTCCTGCAGGAATTCAGGGAGATGGCCCTGGGCGTAGG-
25 CATCCCCGAGGATTCCATCTTCACAATGGCCAACAGAGGGGAATGT-
GTCCCTGGTGACCAGGCACCAGAGTCCACCCCAGCCCCGAGGTGA

3. M8H5GIEGR-Frog (SEQ ID NO: 61)

30 ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCTGCAGCCCAATCCAGCCAGAGGACAATATCCAGATCCAGGA-
GAACTTTGATCTCCAGAGGATTTATGGCAAATGGTACGACATTGCCATCGG-
CTCCACCTGCAAATGG-
CTGAAGCACCACAAGGAAAAGTTCAACATGGGGACACTGGAGCTTAGCGATGGG-
35 GAGACCGACGGGGAGGTGCGGATTGTGAACACAAG-
GATGAGGCACGGAACCTGCTCTCAGATTGTTGGGTCCTATCAGAAGACAGA-
GACCCCAGGGGAAGTTCGACTATTTCAACGCACGGTGGGGAACACGATCCAAAA
CTACATTGTCTTCACTAACTACAATGAG-
TATGTCATCATGCAGATGAGGAAGAAGAAGGGATCGGA-
40 GACCACCACGACCGTCAAGCTGTATGGGCGGAGCCCA-
GACTTGCGTCCGACCCTCGTTGATGAATTCAGGCAGTTTGCCTTGGCTCAGGG-
CATTCTGAAGACTCCATCGTGATGCTACCTAACAATGGTGAGT-
GCTCTCAGGGGAAATAGAAGTGAGACCACGGAGATGA

45

4. M8H5GIEGR-Chicken (SEQ ID NO: 62)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCACGCCTGTTGGGGACCAGGATGAGGACATTCAAGTGCAAGAGAATTTGA

GCCTGAGCGGATGTATGGGAAATGGTATGACGTAGCTGTTGG-
CACCACCTGCAAGT-
GGATGAAGAACTACAAGGAGAAGTTCAGCATGGGCACACTGGTGCTGGGCCCCG
GCCCCAGCGCTGACCAGATCAGTACCATCAGCACCAGGCTGCGG-
5 CAAGGTGACTGCAAACGT-
GTCTCAGGAGAGTACCAGAAAAGTACACCCCTGGCAAATACACCTACTATAACC
CCAAGTGGGATGTGTCTATCAAGTCCTACGT-
GCTTCGCACCAACTATGAAGAATACGCAGTCATTCTGATGAAGAAGACAAGTAATT
TTGGCCCAACCACCACACTGAAGCTGTATGGGAGAAGCCCAGAGCTGCGGGAA-
10 GAGCTCACCGAGGCTTTCCAGCAGCTGGCTCTGGAGATGGGCATCCCTGCAGAT
TCCGTCTTCATCCTGGCCAACAAAGGTGAATGTGTCCCACAGGA-
GACTGCCACTGCCCCTGAGAGGTGA

5. M8H5GIEGR-Rabbit (SEQ ID NO: 63)
15 ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGACCCCGTGCCACCCCTGCCGGACGACATCCAAGTGCAGGA-
GAACTTCGAGCTCTCTCGGATCTACGGGAAATGGTACAACCTGGCTGT-
GGGGTCCACCTGCCCCTGGCTGAAGAGGATCAAGGACAGGATGGCCGT-
GAGCACGCTGGTGCTGGGAGAGGGGACGAGCGAGACGGA-
20 GATCAGCATGACCAGCACGCACTGGCGGAGGGGCGTCTGTGAGGA-
GATCTCCGGGGCCTATGA-
GAAAACGGACACTGACGGGAAGTTCCTGTACCACAAAGCCAAATGGAACCTAAC-
CATGGAGTCCTACGTGGTGACACCAACTACGATGAGTATGC-
CATTTTTCTCACCAAGAAATTCAGCCGCCGCCACGGCCCCAC-
25 CATCACCGCCAAGCTCTATGGGCGGGAGCCGCAGCTGAGGGA-
GAGCCTCCTGCAGGAGTTCAGGGAGGTGGCTCTCGGGGTGGGGATCCCCGA-
GAACTCCATCTTCACCATGATCGACAGAGGGGAATGTGTGCCCGGG-
CAGCAGGAACCAAAGCCTGCCCCCGTGTTGAGATGA

30
6. M8H5GIEGR-SQ Monkey (SEQ ID NO: 64)
ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGAGCCAGTGCCGACGCCGCCGAAGGCATTCAAGT-
GCAGGAAACTTCAATCTCTCTCGGATCTACGGCAAGTGGTACAACCTGG-
35 CCATCGGTTCCACCTGCCCCTGGCTAAAGAAGATCATGGACAGGTTGAAAGT-
GAGCACGCTGGTGCTGGAAGAGGGCGCCACGGAGGCGGA-
GATCAGCATGACCAGCACTCGCTGGCGGAAAGGTTTCTGTGAGCA-
GACCTCTTGGGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACGAACCCAAATGGAACGTAAC-
40 CATGGAGTCCTATGTGGCCACACCAACTATGAGGAGTATGC-
CATTTTCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTATGGGCGG-
GAGCCACAGCTGAGGGAAAGCCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGATTCCATCTTCACCATGG-
45 CTAACCGAGGTGAATGCGTCCCTGGG-

7. M8H5GIEGR-Walrus (SEQ ID NO: 65)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCAGTCCCGTGCTGACGCCGCCTGACGCCATCCAAGTGCAAGA-
GAACTTCGACATCTCTCGGATCTACGGGAAGTGGTTTCATGTGGCCATGGG-
CTCCACCTGCCCGTGGCTGAAGAAGTTCATGGACAG-
5 GATGTCCATGAGCACGCTGGTGCTGGGCGAGGGGGCGACGGATGGGGA-
GATCAGCATGACCAGCACACGTTGGCGGAGAGGCACCTGTGAGGA-
GATCTCTGGGGCTTATGAGAAAACCAGCACTAACGGAAAGTTCCTCTATCA-
TAATCCCAAATGGAACATCACCATGGAGTCCTATGTGGTCCACACCGACTAT-
GATGAGTACGCCATCTTTCTGACCAAGAAATTCAGCCGCCACCATGGGCCCAC-
10 CATTACTGCCAAGCTCTATGGGCGACAGCCGCAGCTTCGAGAAAGCCTGCTG-
GAGGAGTTCAGGGAGCTTGCCCTGGGTGTGGG-
CATCCCCGAGGACTCCATCTTCACCATGGCCAACAAAGGTGAGTGT-
GTCCCTGGGGAGCAGGAACCAGAGCCCTCTCCACACATGAGGTGA

8. M8H5GIEGR-Manatee (SEQ ID NO: 66)
ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCAGCCCAGTGAAAACACCACTCAACGACATCCAAGTGCAGGA-
20 GAACTTTGACCTCCCTCGGATCTACGGGAATGGTTCAACATAGCCATTGG-
CTCCACCTGCCAATGGCTGAAGAGGTTGAAGGCCGGGCGGAC-
CATGAGCACCTGGTCCTGGGAGAGGGAGCTACAGACACAGA-
GATCAGCACAACCAGCACTCGTTGGCGGAAAGGCTTCTGTGAGGA-
GATCTCTGGGGCATATGAGAAAACAGACACAGCTGGGAAGTTCCTTTATCACG-
25 GATCCAAATGGAATGTAACCTTGGAGTCCTATGTGGTCCACACCAACTAT-
GATGAGTACGCCATTTTTCTGACCAAGAAATTCAGCCGCTATGGACTCAC-
CATTACTGCTAAGCTCTATGGGCGGCAGCCTCAGGTGAGGGAGAGCCTCCTG-
GAGGAGTTCAGGGAATTTGCCCTGGGTGTGGGCATCCCTGAG-
GATTCATCTTCACCACGGCCGACAAAGGTGAGTGTGTCCCTGGAGAGCAG-
30 GAGCCAGAACCCACCGCAGCCCTGAGATGA

9. M8H5GIEGR-Plaice (SEQ ID NO: 67)
35 ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCCTCCCTGT-
GCTCCCTGAACCTCTTTACCCGACACAGGAGAACTTTGATCTGACCCGGTTTGTG
GGGACATGG-
CACGATGTTGCCTTGACGAGCAGCTGCCCCCATATGCAGCGTAACAGGGCG-
40 GATGCAGCCATTGGTAACTGGTTCTGGAGAAAGACACTGGAAACAAACTCAAGG
TGACACGAAGTAGACTCAGACATGGAACATGTGTGGAGATGTCTGGA-
GAATATGAGT-
TAACCAGCACACCAGGACGAATCTTCTACCATATTGACAGGTGGGATGCAGACGT
GGACGCCTACGTGGTTCACACCAACTACGACGAGTACGCAATTATAA-
45 TAATGAGCAAACAGAAAACATCGGGGGAGAACAGCACCTCACTCAAGCTGTACAG
TCGGACGATGTCTGTGAGAGACACTGTGCTGGATGACTTCAAACTCTGGTCA-
GACATCAGGGAATGAGTGACGACACCATTATCATCAAGCAGAACAAAGGTGACTG

TATTCCTGGAGAGCAGGTGGAAGAAGCACCATCTCAGCCA-
GAGCCCAAGCGGTGA

10. M8H5GIEGR-Orangutan (SEQ ID NO: 68)

5 ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGCCCTGTGCCGACGCCGCCGACAACATCCAAGT-
GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
GAGCACCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA-
10 GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACATCTG-
GAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCGTCATGGACCCAC-
15 CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAAACCCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAACAGGAACCAGAGCCCATCT-

20 11. M8H5GIEGR-Human P35K (SEQ ID NO:69)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGCCCTGTGCCAACGCCGCCGACAACATCCAAGT-
GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
CCATCGGTTCCACCTGCAAAATGGCTGAAGAAGATCATGGACAGGATGACAGT-
25 GAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA-
GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
GAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
30 CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
35 TAATCCCGAGATGA

12. M8H5GIEGR-Human M41K (SEQ ID NO: 70)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGCCCTGT-
40 GCCAACGCCGCCGACAACATCCAAGTGCAGGAAAACCTTCAATATCTCTCGGATC
TATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
CTGAAGAAGATCAAA-
GACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA
GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
45 GAGCTTATGAGAAAACAGATACTGATGGGAAGTTTCTCTATCACAAATCCAAATGG
AACATAACCATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCACCATTACTGCCAAG

CTCTACGGGCGGGCGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGGCTGACCG
AGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCTTAATCCCGA-
GATGA

5

13. M8H5GIEGR-Human R66H (SEQ ID NO: 71)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGCCCTGTGCCAACGCCGCCCGACAACATCCAAGT-
GCAGGAAACTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
10 CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
GAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA-
GATCAGCATGACCAGCACTCATTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
GAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
15 CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
20 CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

14. M8H5GIEGR-Human T75K (SEQ ID NO: 72)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
25 CCGCGGCCCTGTGCCAACGCCGCCCGACAACATCCAAGT-
GCAGGAAACTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
GAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA-
GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGAAATCTG-
30 GAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGGCGGG-
35 CGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

40 15. M8H5GIEGR-Human T75Y (SEQ ID NO: 73)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGCCCTGT-
GCCAACGCCGCCCGACAACATCCAAGTGCAGGAAACTTCAATATCTCTCGGATC
TATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
45 CTGAAGAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGGC
GCTACAGAGGCGGAGATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGT-
GAG-
GAGTATTCTGGAGCTTATGAGAAAACAGATACTGATGGGAAGTTTCTCTATCACAA

ATCCAAATGGAACATAACCATGGAGTCCTATGTGGTCCACACCAACTATGATGAG-
TATGC-

CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCACCATTACTGCCAAG
CTCTACGGGCGGGCGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
5 GGTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGGCTGACCG
AGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCTTAATCCCGA-
GATGA

16. M8H5GIEGR-Human M99K (SEQ ID NO: 74)

10 ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGCCCTGTGCCAACGCCGCCCGACAACATCCAAGT-
GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
GAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA-
15 GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
GAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAACCAAAGAGTCCTATGT-
GGTCCACACCAACTATGATGAGTATGCCATTTTCCTGACCAAGAAATTCAGCCGC-
CATCATGGACCCACCATTACTGCCAAGCTCTACGGGCGGG-
20 CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

25

17. M8H5GIEGR-Human S101Y (SEQ ID NO: 75)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGCCCTGTGCCAACGCCGCCCGACAACATCCAAGT-
GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
30 CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
GAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA-
GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
GAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAACCATGGAGTATTATGT-
35 GGTCCACACCAACTATGATGAGTATGCCATTTTCCTGACCAAGAAATTCAGCCGC-
CATCATGGACCCACCATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
40 TAATCCCGAGATGA

18. M8H5GIEGR-Human K69.92.118.130R (SEQ ID NO: 76)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
45 CCGCGGCCCTGT-
GCCAACGCCGCCCGACAACATCCAAGTGCAGGAAAACCTTCAATATCTCTCGGATC
TATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
CTGAAGAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGGC

GCTACAGAGGCGGAGATCAGCATGACCAGCACTCGTTGGCGGCGT-
GGTGTCTGTGAGGA-
GACGTCTGGAGCTTATGAGAAAACAGATACTGATGGGAAGTTTCTCTATCACCGTT
CCAAATGGAACATAACCATGGAGTCCTATGTGGTCCACACCAACTATGATGAG-
5 TATGC-
CATTTTCCTGACCAAGCGTTTCAGCCGCCATCATGGACCCACCATTACTGCCCGT
CTCTACGGGCGGGCGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGGCTGACCG
AGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCTTAATCCCGA-
10 GATGA

19. M8H5GIEGR-Coelacanth (SEQ ID NO: 77)
ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGAAGTCCCCTTCGAGATGAAGACATCCAAGTGCAGGA-
15 GAACTTTGACCTTCCCAG-
GATTTATGGAAAATGGTACGAAATTGCAATCGCTTCGACCTGTCCCTGGGTGAAG
AATCACAAGGATAAGATGTTTCATGGGAACTATGGTGCTACAAGAGGGAGAGCA-
GAGTGACCGGATCAGTACCACCTCCACCCGAATCAGG-
GATGGAACCTGCTCACAGATCACTGGATATTACACGT-
20 TAACCACAACACCTGGGAAGTTCGCTTATCACAATTCTAAATGGAACCTTGGATGTC
AACAGTTATGTTGTTCACTAACTATGACGAATACTCGATTGT-
GAT-
GATGCAGAAATACAAAAGCTCTAACTCTACCACTACAGTCCGACTCTATGGAAGAA
CTCAAGAGCTACGAGACAGCTTGCATGCCGAGTTCAAAAAGTTTGCTCTG-
25 GATCAGGGAATAGATGAGGACTCCATTTACATTCTGCCAAAAGAGATGAATGT-
GTACCTGGTGAACCTAAAGCAGAATCTCTCATGGCACGTTGA

21. M8H5GIEGR-Human L89T (SEQ ID NO: 78)
ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
30 CCGCGGCCCTGTGCCAACGCCGCCGACAACATCCAAGT-
GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
GAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA-
GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
35 GAGCTTATGAGAAAACAGATACTGATGGGAAGTT-
TACCTATCACAAATCCAATGGAACATAACCATGGAGTCCTATGT-
GGTCCACACCAACTATGATGAGTATGCCATTTTCCTGACCAAGAAATTCAGCCGC-
CATCATGGACCCACCATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
40 GGTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

45 22. M8H5GIEGR-Human N1796D (SEQ ID NO: 79)
ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGCCCTGT-
GCCAACGCCGCCGACAACATCCAAGTGCAGGAAAACCTTCGATATCTCTCGGATC

TATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
CTGAAGAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGGC
GCTACAGAGGCGGAGATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGT-
GAGGA-
5 GACGTCTGGAGCTTATGAGAAAACAGATACTGATGGGAAGTTTCTCTATCACAAAT
CCAAATGGGATATAACCATGGAGTCCTATGTGGTCCACACCAACTATGATGAG-
TATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCACCATTACTGCCAAG
CTCTACGGGCGGGCGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
10 GGTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGGCTGACCG
AGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCTTAATCCCGA-
GATGA

23. M8H5GIEGR-Human T45K (SEQ ID NO: 80)

15 ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGGCCCTGTGCCAACGCCGCCCGACAACATCCAAGT-
GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGAAAGT-
GAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA-
20 GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
GAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
25 CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

30

24. M8H5GIEGR-Human A135E (SEQ ID NO: 81)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
CCGCGGGCCCTGTGCCAACGCCGCCCGACAACATCCAAGT-
GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
35 CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
GAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA-
GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
GAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
40 CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGG-
CGGGAACCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
45 CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

25. M8H5GIEGR-Human V170S (SEQ ID NO: 82)

ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
 CCGCGGCCCTGTGCCAACGCCGCCGACAACATCCAAGT-
 GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
 CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
 5 GAGCACGCTGGTGTGCTGGGAGAGGGGCGCTACAGAGGCGGA-
 GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
 GAGCTTATGAGAAAACAGATACT-
 GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
 CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
 10 CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
 CATTACTGCCAAGCTCTACGGGCGGG-
 CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
 GGTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
 CTGACCGAGGTGAATGTTCTCCTGGGGAGCAGGAACCAGAGCCCATCT-

15 26. M8H5GIEGR-Human V148D (SEQ ID NO: 83)
 ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
 CCGCGGCCCTGTGCCAACGCCGCCGACAACATCCAAGT-
 GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
 20 CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
 GAGCACGCTGGTGTGCTGGGAGAGGGGCGCTACAGAGGCGGA-
 GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
 GAGCTTATGAGAAAACAGATACT-
 GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
 25 CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
 CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
 CATTACTGCCAAGCTCTACGGGCGGG-
 CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGA-
 GATGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
 30 CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
 TAATCCCGAGATGA

27. M8H5GIEGR-Human G172Q (SEQ ID NO: 84)
 ATGCATCACCATCACCATCACCATCACGGTGGAGGAGGGGGTATCGAGGG-
 35 CCGCGGCCCTGTGCCAACGCCGCCGACAACATCCAAGT-
 GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
 CCATCGGTTCCACCTGCCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
 GAGCACGCTGGTGTGCTGGGAGAGGGGCGCTACAGAGGCGGA-
 GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
 40 GAGCTTATGAGAAAACAGATACT-
 GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
 CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
 CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
 CATTACTGCCAAGCTCTACGGGCGGG-
 45 CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
 GGTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
 CTGACCGAGGTGAATGTGTCCCTCAGGAGCAGGAACCAGAGCCCATCT-

33. M8H4DK-Human M41K+R66H (SEQ ID NO: 85)

- ATGCATCACCATCACCATCACCATCACGATGACGATGACAAGGGCCCTGT-
GCCAACGCCGCCCCGACAACATCCAAGTGCAGGAAAACCTTCAATATCTCTCG-
GATCTATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCCTGG-
CTGAAGAAGATCAAAGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGG-
5 CGCTACAGAGGGCGGAGATCAGCATGACCAGCACTCATTGG-
CGGAAAGGTGTCTGTGAGGAGACGTCTGGAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
10 CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA
- 15 34. M8H4DK-Human M41K+N1796D (SEQ ID NO: 86)
ATGCATCACCATCACCATCACCATCACGATGACGATGACAAGGGCCCTGT-
GCCAACGCCGCCCCGACAACATCCAAGTGCAGGAAAACCTTCGATATCTCTCG-
GATCTATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCCTGG-
CTGAAGAAGATCAAAGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGG-
20 CGCTACAGAGGGCGGAGATCAGCATGACCAGCACTCGTTGG-
CGGAAAGGTGTCTGTGAGGAGACGTCTGGAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGGATATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
25 CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA
- 30 35. M8H4DK-Human R66H+N1796D (SEQ ID NO: 87)
ATGCATCACCATCACCATCACCATCACGATGACGATGACAAGGGCCCTGT-
GCCAACGCCGCCCCGACAACATCCAAGTGCAGGAAAACCTTCGATATCTCTCG-
GATCTATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCCTGG-
CTGAAGAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGG-
35 CGCTACAGAGGGCGGAGATCAGCATGACCAGCACTCATTGG-
CGGAAAGGTGTCTGTGAGGAGACGTCTGGAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGGATATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
40 CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA
- 45 36. M8H4DK-Human M41K+R66H+N1796D (SEQ ID NO: 88)
ATGCATCACCATCACCATCACCATCACGATGACGATGACAAGGGCCCTGT-
GCCAACGCCGCCCCGACAACATCCAAGTGCAGGAAAACCTTCGATATCTCTCGGATC

TATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
CTGAAGAAGATCAAA-
GACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGGCGCTACAGAGGCGGA
GATCAGCATGACCAGCACTCATTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
5 GAGCTTATGAGAAAACAGATACTGATGGGAAGTTTCTCTATCACAAATCCAAATGG
GATATAACCATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTGAGCCGCCATCATGGACCCACCATTACTGCCAAG
CTCTACGGGCGGGCGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGGCTGACCG
10 AGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCTTAATCCCGA-
GATGA

37. M8H4DK-Human M41K (SEQ ID NO: 89)
ATGCATCACCATCACCATCACCATCACGATGACGATGACAAGGGCCCTGT-
15 GCCAACGCCGCCCGACAACATCCAAGTGCAGGAAAACCTTCAATATCTCTCG-
GATCTATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
CTGAAGAAGATCAAAAGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGG-
CGCTACAGAGGCGGAGATCAGCATGACCAGCACTCGTTGG-
CGGAAAGGTGTCTGTGAGGAGACGTCTGGAGCTTATGAGAAAACAGATACT-
20 GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTGAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
25 GGTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

38. M8H4DK-Human R66H (SEQ ID NO: 90)
ATGCATCACCATCACCATCACCATCACGATGACGATGACAAGGGCCCTGT-
30 GCCAACGCCGCCCGACAACATCCAAGTGCAGGAAAACCTTCAATATCTCTCG-
GATCTATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
CTGAAGAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGG-
CGCTACAGAGGCGGAGATCAGCATGACCAGCACTCATTGG-
CGGAAAGGTGTCTGTGAGGAGACGTCTGGAGCTTATGAGAAAACAGATACT-
35 GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTGAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
40 GGTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

39. M8H4DK-Human N1796D (SEQ ID NO: 91)
45 ATGCATCACCATCACCATCACCATCACGATGACGATGACAAGGGCCCTGT-
GCCAACGCCGCCCGACAACATCCAAGTGCAGGAAAACCTTCGATATCTCTCGGATC
TATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
CTGAAGAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGGC
GCTACAGAGGCGGAGATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGT-
50 GAGGA-
GACGTCTGGAGCTTATGAGAAAACAGATACTGATGGGAAGTTTCTCTATCACAAAT

CCAAATGGGATATAACCATGGAGTCCTATGTGGTCCACACCAACTATGATGAG-
TATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCACCATTACTGCCAAG
CTCTACGGGCGGGCGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
5 GGTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGGCTGACCG
AGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCTTAATCCCGA-
GATGA

40. M8H-Human wt (SEQ ID NO: 92)

10 ATGCATCACCATCACCATCACCATCACGGCCCTGT-
GCCAACGCCGCCCCGACAACATCCAAGTGCAGGAAAACCTTCAATATCTCTCG-
GATCTATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
CTGAAGAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGG-
CGCTACAGAGGGCGGAGATCAGCATGACCAGCACTCGTTGG-
15 CGGAAAGGTGTCTGTGAGGAGACGTCTGGAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGGCGGG-
20 CGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

41. M8H-Human R66H+N1796D (SEQ ID NO: 93)

25 ATGCATCACCATCACCATCACCATCACGGCCCTGT-
GCCAACGCCGCCCCGACAACATCCAAGTGCAGGAAAACCTTCGATATCTCTCG-
GATCTATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGG-
CTGAAGAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGG-
CGCTACAGAGGGCGGAGATCAGCATGACCAGCACTCATTGG-
30 CGGAAAGGTGTCTGTGAGGAGACGTCTGGAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGGATATAAC-
CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGGCGGG-
35 CGCCGCAGCTGAGGGAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

42. untagged-Human R66H+N1796D (SEQ ID NO: 94)

40 ATGGGCCCTGTGCCAACGCCGCCCCGACAACATCCAAGT-
GCAGGAAAACCTTCGATATCTCTCG-
GATCTATGGGAAGTGGTACAACCTGGCCATCGGTTCCACCTGCCCCTGGCTGAA
GAAGATCATGGACAGGATGACAGTGAGCACGCTGGTGCTGGGAGAGGG-
CGCTACAGAGGCG-
45 GAGATCAGCATGACCAGCACTCATTGGCGGAAAGGTGTCTGTGAGGAGACGTCT
GGAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGGATA-
TAACCATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGCCATTTTCCTG

ACCAAGAAATTCAGCCGCCATCATGGACCCACCATTACTGCCAAGCTCTACGGG-
CGGG-
CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGTGGTTGCCAGGGTG
TGGGCATCCCTGAGGACTCCATCTTCACCATGGCTGACCGAGGTGAATGT-
5 GTCCCTGGGGAGCAGGAACCAGAGCCCATCTTAATCCCGAGATGA

61. untagged-Human wt (SEQ ID NO: 95)
ATGGGCCCTGTGCCAACGCCGCCGACAACATCCAAGT-
GCAGGAAAACCTTCAATATCTCTCGGATCTATGGGAAGTGGTACAACCTGG-
10 CCATCGGTTCCACCTGCCCTGGCTGAAGAAGATCATGGACAGGATGACAGT-
GAGCACGCTGGTGTCTGGGAGAGGGCGCTACAGAGGCGGA-
GATCAGCATGACCAGCACTCGTTGGCGGAAAGGTGTCTGTGAGGAGACGTCTG-
GAGCTTATGAGAAAACAGATACT-
GATGGGAAGTTTCTCTATCACAAATCCAAATGGAACATAAC-
15 CATGGAGTCCTATGTGGTCCACACCAACTATGATGAGTATGC-
CATTTTCCTGACCAAGAAATTCAGCCGCCATCATGGACCCAC-
CATTACTGCCAAGCTCTACGGGCGGG-
CGCCGCAGCTGAGGGAAACTCTCCTGCAGGACTTCAGAGT-
GGTTGCCAGGGTGTGGGCATCCCTGAGGACTCCATCTTCACCATGG-
20 CTGACCGAGGTGAATGTGTCCCTGGGGAGCAGGAACCAGAGCCCATCT-
TAATCCCGAGATGA

Rationale of A1M variant constructions

Human wt-A1M, 11 A1M-homologues from various species, and 15 A1M-variants with
25 point mutations were constructed, expressed, purified and analysed in phase I, i.e. a
total of 27 A1M-variants (Table 2). The 11 A1M-homologues were selected as follows.
A1M is well conserved between species. A1M sequences of different species were
searched for in data bases (www.ncbi.nlm.nih.gov and www.uniprot.org). AMBP se-
quences from 67 different species were found. The sequences were investigated for
30 presence of A1M-specific functional groups (K69, K92, K118, K130, Y22; Y132, H122
and H123), lipocalin motifs (SCR1, 2, 3) (see refs in Introduction) and predicted carbo-
hydrate binding sites (Escribano et al., 1990). Five were classified as non-A1M and dis-
missed because they lacked cystein 34, three were dismissed because they lacked
cystein 169, four because they were incomplete, and two because they had long, ques-
35 tionable, inserts. The amino acid sequences of the remaining 53 homologues were
aligned and their suggested 3D structure were modelled by projecting their amino acid
sequences on the crystal structure of human A1M (Meining and Skerra, 2012) to iden-
tify the location of individual amino acid side chains in the lipocalin loops or on the in-
side or outside of the pocket. The result of this was used for construction of point muta-
40 tions affecting A1M-function as described below. The rationale for selection of the final
11 homologues (see Table 2) was a combination of wide-spread evolutionary represen-
tation, methodological feasibility, potentially increased environmental oxidative stress in

the living habitat of the species, presence or absence of A1M-specific functional groups, and lack of glycosylation. The 15 A1M-variants with point mutations were selected as follows. Eight of the variants were mutated in strategic amino acids for functional properties and seven variants in exposed positions to improve stability/solubility (Table 2). The eight functional mutations were selected as they 1) occurred in other species in positions in the 3D-structural model (see above) that could potentially influence the function of A1M and/or 2) theoretically would lower the pKa of cysteine 34 and/or 3) theoretically would provide an extra radical trapping site. Predicted data for all variants was calculated using the http://www.alphalyse.com/gpmaw_lite.html tool (Table 3).

P017240DK1

Table 2. Variants expressed and analysed in A1M variants project phase I.

No*	Name	Systematic name	Protein sequence	Gene Identifier	Aim of mutation	Rationale	N-terminal tag
60	human wt	<i>Homo sapiens</i>	NP_001624	gi:4502067			M8H4DK
1	mouse	<i>Mus musculus</i>	NP_031469	gi:311703	Species	only rodent, lab animal, we have antibodies	M8H5GIEGR
2	naked mole-rat	<i>Heterocephalus glaber</i>	XP_004885260	gi:512843527	Species	long-lived, low tumor frequency suggests good endogenous antioxidant	M8H5GIEGR
3	frog	<i>Xenopus laevis</i>	NP_001080820	gi:147905774	Species	no predicted carbohydrate suggests high solubility of polypeptide, 1 extra Cys, high pI	M8H5GIEGR
4	chicken	<i>Gallus gallus</i>	NP_001264627	gi:478732996	Species	only bird	M8H5GIEGR
5	rabbit	<i>Oryctolagus cuniculus</i>	XP_008271622	gi:655878696	Species	lab animal, we have antibodies	M8H5GIEGR
6	squirrel monkey	<i>Saimiri boliviensis</i>	XP_003925255 ₁	gi:403266141	Species	two primates, less substitutions compared to human, easier readout	M8H5GIEGR
7	walrus	<i>Odobenus rosmarus</i>	XP_004397010	gi:472354701	Species	uneven oxygenation due to diving	M8H5GIEGR
8	manatee	<i>Trichechus manatus</i>	XP_004372191	gi:047136326 ₇	Species	uneven oxygenation due to diving	M8H5GIEGR
9	plaice	<i>Pleuronectes platessa</i>	2021284A	gi:1091607	Species	uneven oxygenation due to diving	M8H5GIEGR
10	orangutan	<i>Pongo abelii</i>	NP_001127069	gi:197102775	Species	one substitution - H122R - in an interesting position: reductase and heme-binding	M8H5GIEGR
11	human P35K				Functional	a) occurs in birds and frogs, b) may lower pKa of C34, c) may be radical trapping site	M8H5GIEGR

P017240DK1

12	human M41K				Functional	a) occurs in many species, b) may lower pKa of C34, c) may be radical trapping site	M8H5GIEGR
13	human R66H				Functional	occurs in many species, c.f. rodents and rabbits	M8H5GIEGR
14	human T75K				Functional	a) may lower pKa of C34, b) may be radical trapping site	M8H5GIEGR
15	human T75Y				Functional	may be radical trapping site	M8H5GIEGR
16	human M99K				Functional	a) may lower pKa of C34, b) may be radical trapping site	M8H5GIEGR
17	human S101Y				Functional	may be radical trapping site	M8H5GIEGR
18	human K69,92,118,130R				Functional	may be radical trapping sites	M8H5GIEGR
19	coelacanth	<i>Latimeria chalumnae</i>	XP_005994505.1	gi:556972695	Species	most primitive A1M-sequence identified	M8H5GIEGR
21	human L89T				Solubility	exposed position, increased hydrophilicity, occurs in many species	M8H5GIEGR
22	human N17,96D				Solubility	carbohydrate sites, increased hydrophilicity	M8H5GIEGR
23	human T45K				Solubility	exposed position, increased charge, occurs in many species	M8H5GIEGR
24	human A135E				Solubility	exposed position, increased charge, occurs in many species	M8H5GIEGR
25	human V170S				Solubility	exposed position, increased hydrophilicity, occurs in many species	M8H5GIEGR
26	human V148D				Solubility	exposed position, increased hydrophilicity, occurs in many species	M8H5GIEGR

P017240DK1

27	human G172Q				Solubility	exposed position, increased hydrophobicity, occurs in birds	M8H5GIEGR
----	-------------	--	--	--	------------	---	-----------

- Numbering corresponds to the sequences mentioned above

Table 3. Predicted data for phase I variants

#	Variant	# amino acids	mw (Da)	pI	Net charge (D+E)-(R+K)	ext. Co-eff 280nm (cm ⁻¹)	Hy-drophob . index
60	human wt	197	22561	6.4	-5	1.47	-0.64
1	mouse	200	22732	6.6	-4	1.46	-0.60
2	naked mole rat	201	22741	9.2	2	1.40	-0.73
3	frog	197	22660	8.0	0	1.22	-0.79
4	chicken	197	22302	6.5	-3	1.40	-0.76
5	rabbit	201	22911	6.9	-2	1.45	-0.61
6	Sq. monkey	201	23066	7.3	-1	1.68	-0.69
7	walrus	201	22837	6.6	-4	1.39	-0.63
8	manatee	200	22469	6.6	-3	1.47	-0.54
9	plaice	202	22896	6.6	-4	0.89	-0.71
10	orangutan	201	22733	7.2	-1	1.46	-0.58
11	human P35K	201	22745	7.2	-1	1.46	-0.58
12	human M41K	201	22711	7.2	-1	1.46	-0.60
13	human R66H	201	22695	6.7	-3	1.46	-0.57
14	human T75K	201	22741	7.2	-1	1.46	-0.59
15	human T75Y	201	22776	6.9	-2	1.51	-0.58
16	human M99K	201	22711	7.2	-1	1.46	-0.60
17	human S101Y	201	22790	6.9	-2	1.51	-0.57
18	human K69.92.118.130R	201	22826	6.9	-2	1.45	-0.58
19	coelacanth	198	22741	6.7	-3	1.38	-0.78
21	human L89T	201	22702	6.9	-2	1.46	-0.59
22	human N17.96D	201	22716	6.5	-4	1.46	-0.57
23	human T45K	201	22741	7.2	-1	1.46	-0.59
24	human A135E	201	22772	6.7	-3	1.45	-0.60
25	human V170S	201	22702	6.9	-2	1.46	-0.60
26	human V148D	201	22730	6.7	-3	1.46	-0.61
27	human G172Q	201	22785	6.9	-2	1.45	-0.59

5 Phase I

The variants were expressed in parallel shake-flasks. There was a variation in expression levels and for a few A1M-variants expression had to be repeated several times to obtain reasonable protein amounts. Good expression and relatively large amounts after purification (>10mg/L culture) were obtained for mouse, squirrel monkey, walrus,

10 orangutan, M41K, R66H, T75Y, M99K, N17,96D, T45K and V148D. The expression

- and amount of purified protein was similar to previous yields for human wt A1M (12mg/L) expressed and purified under similar conditions. Variants with lower expression levels resulting in lower yields (1-10mg/L) were naked mole rat, frog, rabbit, manatee, P35K, T75K, S101Y, L89T, V170S and G172Q. Chicken, plaice and coelacanth
- 5 A1M displayed suboptimal expression levels resulting in a purification yield of <1mg/L culture. The quadruple human mutant K69,92,118,130R expressed well, but was impossible to refold successfully. Also T75K-A1M showed increased precipitation tendencies during purification.
- 10 All variants were purified to >99% purity according to SDS-PAGE except chicken and coelacanth A1M, which were purified to 95% purity. According to SDS-PAGE all variants contained a maximum of 0.5% covalent dimers with no major differences between the variants. Circular dichroism revealed about 2% alpha helix and 40% β -sheet structure for all variants without major differences, the expected values for A1M (Kwasek et al., 2007). After purification, all variants were analysed as summarized in Table 4.
- 15 Firstly, stability and solubility was analysed. The aggregation tendency and thermostability of freshly thawed, unstressed 0.1mM protein solutions were analysed by dynamic light scattering (DLS), PAGE (aggregation) and differential scanning fluorimetry (DSF) (thermostability). In addition, the tendency to aggregate in response to shearing forces
- 20 or concentration to 1mM was investigated. Human wt A1M performed relatively well in these assays (Table 4), but some variants performed even better in the stability/solubility properties: chicken A1M, manatee A1M and N17,96D-A1M. The chicken and manatee A1M differed from human wt A1M in too many positions to enable assumptions on individual amino acids beneficial for human A1M and was not used in further studies.
- 25 When functional properties were investigated, M41K and R66H were found to have the same (R66H) or improved (M41K) functions. As these variants also had higher thermostability and only slightly higher tendency to aggregate than most other variants, they were selected for further investigation. Hence, N17,96D-A1M, M41K-A1M and R66H-A1M were continued to phase II of the project.

P017240DK1

Table 4. Solubility, Stability and functional properties of phase I variants.

		Stability and solubility properties					Functional Properties					
		Aggregation at 0.1mM		Thermo-stability	Seraing stability at 0.1mM		Ag-gre-gation at 1mM		Free Thiols	Binding of heme	Reduction of heme	Reduction of ABTS
		PAGE	DLS	DSF	DLS		PAGE					
		Aggre-gates (%)	Aver-age radius (nm)	Meas-ure-ments used (# of 6)	T _m	Average radius (nm)	Meas-ure-ments used (# of 6)	Ag-gre-gates (%)	nmol L-cystein/nmol A1M	μM bound /μM added	Peak λ (380-420nm)	Average Max Net AUC
No	Name											
60	human, wt	4.8	3.2	6	46.6	3.8	4	22.3	0.6	0.81	405	-1.84 to -2.04
1	mouse	45.8	20	6	43	32.8	6	93.7	0.3	0.71	393	-1.43
2	naked mole rat	54.8	2.7	6	51.6	3	6	80.1	0.1	0.69	397	-1.42
3	frog	9.9	2.8	5	NM	3	1	24.1	0.1	0.81	410	-1.33
4	chicken	2.4	3.3	5	67.5	3.4	6	3.5	0.1	0.66	391	-1.98
5	rabbit	50.2	33	6	48.7	43.4	6	28.2	0.1	0.65	395	-1.50
6	sq. monkey	20.2	3.2	6	49.9	3	6	38.3	0.3	0.72	407	-1.98
7	walrus	4.0	3.8	3	NM	3.9	1	52.3	0.3	0.81	398	-1.33
8	manatee	1.7	3.5	6	53.2	3.6	5	2.0	0.1	0.80	391	-1.81
9	plaice	6.6	3.1	6	NM	3.1	6	30.0	0.0	0.73	388	-1.25
10	orangutan	18.3	3.1	6	45.9	3	4	29.7	0.6	0.77	401	-2.06

P017240DK1

11	human P35K	94.5	29.2	6	44.6	40.9	6	86.7	0.1	0.81	396	-1.29
12	human M41K	20.5	3.2	3	47.9	3.2	5	27.5	0.4	0.92	414	-1.87
13	human R66H	13.6	3.1	4	47.8	19	6	37.0	0.7	0.79	405	-2.19
14	human T75K	60.0	42.3	6	43.3	82	6	96.3	0.3	0.76	393	-1.28
15	human T75Y	36.1	35.7	6	49.8	32.4	6	45.7	0.4	0.87	392	-2.08
16	human M99K	37.7	65.8	6	45.9	56.4	6	44.8	0.4	0.69	397	-1.71
17	human S101Y	36.8	2.7	6	46.1	3	3	70.7	0.4	0.79	403	-1.63
Functional Properties												
Stability and solubility properties												
				Aggregation at 0.1mM				Thermo-stability		Sheraing stability at 0.1mM		Ag-gre-gation at 1mM
				PAGE	DLS	DLS	Meas-ure-ments used (# of 6)	DSF	T _m	Average radius (nm)	Meas-ure-ments used (# of 6)	
				Aggre-gates (%)	Aver-age radius (nm)	Meas-ure-ments used (# of 6)						
N o												
60	human, wt	4.8	3.2	6	46.6	3.8	4	22.3	0.6	0.81	405	-1.84 to -2.04
19	coelacanth	4.1	3.7	6	NM	3.8	6	38.2	0.2	0.68	396	-2.00
21	human L89T	42.1	2.9	1	46.1	25.7	6	51.6	0.5	0.96	403	-1.69
22	human N17.96D	2.9	3.1	6	49.6	3.7	4	9.9	0.6	0.80	409	-2.48
23	human T45K	91.2	48.5	6	42.8	86.9	6	89.9	0.3	0.94	394	-1.29

P017240DK1

24	human A135E	54.6	23.9	6	46.1	17.5	6	90.7	0.3	0.70	400	-1.47
25	human V170S	40.1	3	6	46	23.7	6	58.8	0.4	0.90	399	-2.04
26	human V148D	35.8	3	6	46	3	2	44.9	0.6	0.86	402	-2.16
27	human G172Q	93.4	36.1	6	44.1	34.2	6	89.5	0.2	0.81	396	-1.95

Properties regarded to be better for variant compared to human wt A1M are marked in dark blue with white, bold text. Properties regarded to be inferior of variant compared to human wt A1M are marked in red. All other properties are regarded as similar to human wt A1M.

In PAGE, variants with a significantly higher percentage of large aggregates are judged as inferior to human wt A1M.

In DLS, an average radius >> 3nm is regarded as indication of aggregates as well as difficulties in recording of measurements (seen as lower number of measurements used for calculations).

In DSF a $T_m > 3$ S.D of human wt A1M is regarded as higher or lower.

Free thiols of human wt A1M is typically in the range 0.6-0.8 anything significantly below this is regarded as inferior.

In heme binding assay, human wt A1M binds to $\approx 80\%$ in this assay. Binding of other variants below 75% and above 85% is regarded as different.

10 In heme reduction assay a Soret band peak maximum <400nm is regarded as inferior of human wt A1M and a shift ≥ 409 nm is regarded as better than human wt A1M.

In ABTS assay, a T-test p-value of <0.2 is regarded to indicate an inferior ABTS reduction capacity.

Phase II

The aim of phase II of the project was to combine the beneficial mutations (M41K, R66H and N17,96D) to investigate if even more stable variants with similar or improved function could be constructed. In addition, the same N-terminal tag (M8H4DK) which is used for human wt A1M was introduced (Table 5). Thus, the new combined M8H4DK-variants were compared to M8H4DK-human wt.

Table 5. Variants expressed in phase II with predicted data

#	Variant	# amino acids	mw (Da)	pI	Net charge (D+E)-(R+K)	ext. co-eff 280nm (cm ⁻¹)	hydrophobic index
12	M8H5GIEGR- M41K	201	22711	7.2	-1	1.46	-0.60
13	M8H5GIEGR- R66H	201	22695	6.7	-3	1.46	-0.57
22	M8H5GIEGR- N17,96D	201	22716	6.5	-4	1.46	-0.57
60	M8H4DK-wt	197	22561	6.4	-5	1.47	-0.64
33	M8H4DK-M41K+R66H	197	22539	6.4	-5	1.47	-0.67
34	M8H4DK-M41K+N17,96D	197	22560	6.2	-6	1.47	-0.67
35	M8H4DK-R66H+N17,96D	197	22544	6.1	-8	1.47	-0.64
36	M8H4DK-M41K-R66H+N1796D	197	22541	6.2	-7	1.47	-0.67

All four new variants (M8H4DK-M41K+R66H, M8H4DK-M41K+N17,96D, M8H4DK-R66H+N17,96D and M8H4DK-M41K+R66H+N17,96D) expressed very well and yields of 29, 13, 37 and 15 mg pure protein/L were achieved. The four new variants were purified to >99% purity according to SDS-PAGE and showed the same or lower percentage of covalent dimers except M8H4DK-M41K+R66H that showed approximately 2% covalent dimers.

Solubilization/stability properties were analysed (Table 6a). Data revealed that the combination of M41K and R66H was suboptimal for thermostability. The N17,96D mutation, on the other hand, was generally beneficial for solubilization and stability, and the combination with R66H (i.e. the M8H4DK-R66H+N17,96D-A1M) yielded even better solubilization and stability.

Functional properties were analysed (Table 6b). Data showed that the M41K mutation had increased heme reduction capacity, but decreased heme binding. When the M41K mutation was combined with N17,96D, however, it showed less potential to rescue
5 cells from heme-induced cell death. The R66H+N17,96D combination yielded similar results compared to wt-A1M in all functional aspects investigated. Therefore, the M8H4DK-tagged R66H+N17,96D A1M is a promising candidate after phase II, while the M41K-mutation is less useful.

P017240DK1

90

Table 6a. Solubility, Stability and functional properties of phase II variants.

		Stability and solubility properties									
		Aggregation at 0.1mM				Thermo-stability		Sheraing stability at 0.1mM		Aggregation at 1mM	
	No	Name	PAGE	DLS		SEC-FPLC	DSF	DLS		PAGE	SEC-FPLC
				Average radius (nm)	Measurements used (# of 6)			Average radius (nm)	Measurements used (# of 6)		
	60	M8H4DK-wt	1-2	3.0	6	87-94	46.6	3.0	2	8	80-87
	12	M8H5GIEGR-M41K	34	6.0	3	ND	51.8	4.1	2	33	ND
	13	M8H5GIEGR-R66H	25	3.2	6	ND	47.3	10.3	1	51	ND
	22	M8H5GIEGR-N1796D	1	3.0	4	94	49.9	6.5	2	4.2	82
	33	M8H4DK-M41K+R66H	19	3.4	1	ND	34.3	13.7	5	42.9	ND
	34	M8H4DK-M41K+N1796D	1	3.4	6	96	52.7	3.2	5	1.0	77
	35	M8H4DK-R66H+N1796D	0.4	3.3	6	93	51.6	3.1	6	0.4	89
	36	M8H4DK-M41K+R66H+N1796D	0.8	3.3	5	82	37.4	3.6	6	1.1	63

Properties regarded to be better for variant compared to human wt A1M are marked in dark blue with white, bold text. Properties regarded to be inferior of variant compared to human wt A1M are marked in red. All other properties are regarded as similar to human wt A1M.

In PAGE, variants with a significantly higher percentage of large aggregates are judged as inferior to human wt A1M.

In DLS an average radius >> 3nm is regarded as indication of aggregates as well as difficulties in recording of measurements (seen as lower number of measurements used for calculations).

In DSF a T_m >3 S.D of human wt A1M is regarded as higher or lower.

P017240DK1

Table 6b. Functional properties of phase II variants.

		Functional Properties			
		Binding to heme-agarose	Red-shift of heme-binding peak	Reduction of ABTS	Rescue from heme induced cell death
		μM bound / μM added	Peak λ (nm)	Average Max Net AUC	Ocular observation of graph in two separate experiments
No	Name				
60	M8H4DK-wt	0.85	415	-2.1	
12	M8H5GIEGR-M41K	0.81	408	-1.2	Similar as wt
13	M8H5GIEGR-R66H	0.84	ND	-1.5	Similar as wt
22	M8H5GIEGR-N1796D	0.82	415	-2.2	Similar as wt
33	M8H4DK-M41K+R66H	0.68	417	-1.8	Similar as wt
34	M8H4DK-M41K+N1796D	0.74	419	-2.2	Weaker than wt
35	M8H4DK-R66H+N1796D	1.07	413	-2.1	Similar as wt
36	M8H4DK-M41K+R66H+N1796D	1.13	412	-2.0	Similar as wt

Human wt A1M binds to $\approx 80\%$ in the heme-agarose binding assay. Binding of other variants below 75% and above 85% is regarded as different.

In heme-binding absorbance assay, a redshift to $<400\text{nm}$ is regarded as inferior of human wt A1M and a shift $\geq 417\text{nm}$ is regarded as better than human wt A1M.

In ABTS assay, a T-test p-value of <0.2 is regarded to indicate an inferior ABTS reduction capacity.

Phase III

The first aim of phase III was to compare M8H4DK-tagged R66H+N17,96D-A1M with wt-A1M and promising M8H4DK-tagged single mutations of phase I and II, when expressed in parallel under identical conditions. The second aim was to investigate the influence of the tag on stability, solubility and function. The N-terminal tag of human wt A1M (M8H4DK) consist of 8 histidines tag followed by an enterokinase cleavage site. The cleavage site enables cleavage of the histidine tag after expression and purification. Therefore, wt-A1M and R66H+N17,96D-A1M were expressed with N-terminal M8H4DK- and M8H-tags and without a tag. The third aim was to compare in more detail the properties of M8H4DK-wt A1M (Wt-A1M) and M8H4DK-R66H+N17,96D-A1M (35-A1M). All expressed variants of phase III with predictive data are shown in Table 7.

Table 7. Variants expressed in phase III with predicted data

#	Variant	# amino acids	mw (Da)	pI	Net charge (D+E)-(R+K)	ext. co-eff 280nm (cm ⁻¹)	hydrophobic index
60	M8H4DK-wt	197	22561	6.4	-5	1.47	-0.64
35	M8H4DK-R66H+N1796D	197	22544	6.1	-8	1.47	-0.64
37	M8H4DK-M41K	197	22558	6.5	-4	1.47	-0.67
38	M8H4DK-R66H	197	22542	6.3	-8	1.47	-0.64
39	M8H4DK-N17.96D	197	22563	6.1	-7	1.47	-0.64
40	M8H-wt	192	21973	6.9	-2	1.51	-0.57
41	M8H-R66H+N1796D	192	21956	6.4	-5	1.51	-0.56
42	M-R66H+N1796D	184	20859	5.6	-5	1.59	-0.45
61	M-wt	184	20876	6.3	-2	1.59	-0.45

The parallel expression and purification allowed more accurate comparison of expression levels and purification yields. It was found that all A1M variants expressing bacteria grow with the same rate (not shown) and the same relatively good yield (Figure 2). The only difference was a slower expression of the untagged variants (#60 and #42) compared to the others. From the 2x750ml expressions it was possible to purify 20-60mg of each variant (Figure 3a). All histidine-tagged variants were purified to >99%

purity, while the untagged variants (#61 and #42) showed a somewhat lower purity (around 90%) (Figure 3b). All displayed a low amounts of covalent dimers. PAGE analysis showed less than 1-2% large aggregates of all variants except M8H-wt (40) and untagged wt-A1M (#61) (Figure 3c). Thus, all single and multiple M8H4DK-tagged, M8H-tagged and untagged A1M variants could be produced.

The aggregation and thermostability of 100µM solutions were analysed by reversed-phase HPLC, dynamic light scattering and differential scanning fluorimetry (Table 8). The M8H4DK-tagged variants were eluted in RP-HPLC at approximately 12 min, with 87-90% of the protein in the main peak. The only exception was M8H4DK-M41K that appeared at a retention time of 9.8 minutes and only 80% of the protein in the main peak. The data once more indicate that the M41K mutation yields unexpected protein properties of A1M, and is suboptimal. Changing the tag to M8H or completely removing the tag, resulted in increased retention times to 12.2 and 12.9 min, respectively, indicating more hydrophobic molecules, as expected. Less protein appeared in the main peak for M8H-wt and both untagged variants, possibly reflecting the slightly lower purity seen in SDS-PAGE of these variants. DLS-data indicated an average radius around 3 of all variants, and the possibility to collect data from all 6 recordings. All variants had the same or higher thermostability as M8H4DK-wt A1M. Higher stability was seen for all variants carrying the N17,96D mutation.

Table 8. Aggregation and thermostability of non-stressed phase III variants.

#	variant	RP-HPLC		DLS		DSF
		RT [min]	Main peak (% of total)	Average radius (nm)	Measurements used (# of 6)	T _m
60	M8H4DK-wt	12.1	87	2.9	6	46.6
35	M8H4DK-R66H+N1796D	12.1	89	3.2	6	49.8
37	M8H4DK-M41K	9.8	80	3.3	6	46.8
38	M8H4DK-R66H	12.1	89	2.9	6	46.4
39	M8H4DK-N1796D	12.1	90	3.0	6	49.7
40	M8H-wt	12.2	78	3.2	6	47.1
41	M8H-R66H+N1796D	12.2	88	3.2	6	50.2
61	M-wt	12.9	81	3.8	6	46.3
42	M-R66H+N1796D	12.9	82	3.4	6	50.2

The variants of phase III were exposed to shearing forces by pipetting 80 times in a narrow pipett tip. The shearing-stress induced aggregation was the analysed with DLS. In general, all variants with the N17,96D mutation showed high tolerance towards shearing stress (Table 9), with the exception of M8H-tagged R66H+N1796D variant (#41). Also, the tolerance towards stress induced by concentration to 1mM in different buffers was investigated. All variants were concentrated to 1mM in 20mM Tris-HCl + 0.125M NaCl pH 8.0 or 7.4, and were immediately analysed by PAGE (Figure 4a, Table 10). The 0.1 mM and 1 mM samples were also analysed with PAGE after a single freeze-thaw cycle in the Tris-buffers or PBS (Figure 4b, Table 10). Among the M8H4DK-tagged variants all the N17,96D-carrying mutations displayed an increased tolerance to concentration, freeze-thawing, decreased pH and buffer-change (to PBS) compared to wt A1M. The addition of the R66H mutation improved the stability and solubility even further. The R66H mutation alone had approximately the same tolerance as wt A1M, while the M41K mutation had significantly lower tolerance. Shortening or removal of the N-terminal tag had a negative effect on the tolerance, which was more pronounced for wt A1M compared to corresponding R66H+N17,96D variants.

Table 9. Shearing stability of phase III variants.

#	variant	Average radius (nm)	Measurements used (#of 6)
60	M8H4DK-wt	2.9	3
35	M8H4DK-R66H+N1796D	3.0	6
37	M8H4DK-M41K		0
38	M8H4DK-R66H		0
39	M8H4DK-N1796D	3.1	6
40	M8H-wt	3.1	1
41	M8H-R66H+N1796D		0
61	M-wt	3.9	5
42	M-R66H+N1796D	3.4	6

P017240DK1

Table 10. Concentration and freeze-thaw stability of phase III variants in different buffers.

#	variant	Large aggregates (%)						
		0.1mM pH 8.0	1mM pH 8.0	1mM pH 8.0, 1x freeze-thaw	0.1mM pH 7.4	1mM pH 7.4	1mM pH 7.4, 1x freeze-thaw	1mM PBS pH 7.4, 1x freeze-thaw
60	M8H4DK-wt	0.3	1.1	2.0	0.4	1.2	3.1	13.2
35	M8H4DK-R66H+N1796D	0.2	0.3	0.1	0.4	0.2	0.3	1.3
37	M8H4DK-M41K	1.3	4.1	4.8	0.8	7.9	11.9	15.6
38	M8H4DK-R66H	0.6	0.5	1.0	0.2	0.7	2.4	16.0
39	M8H4DK-N1796D	0.7	1.1	0.1	0.2	0.1	0.7	2.5
40	M8H-wt	0.8	7.2	3.6	8.4	10.2	11.4	14.0
41	M8H-R66H+N1796D	0.3	0.4	0.4	0.3	0.4	1.2	7.6
61	M-wt	0.4	2.0	2.3	9.1	1.6	3.9	17.9
42	M-R66H+N1796D	0.4	0.3	0.5	0.4	0.5	0.5	1.6

The functional activities of all phase III variants were investigated in an extended program to cover all possible aspects of known A1M functions. Non-heme related reduction capacity was investigated on the synthetic ABTS radical and in the commercially available oxygen radical antioxidant capacity (ORAC) assay (Figure 5). M8H4DK-tagged N17,96D-A1M and R66H+N17,96D-A1M showed a similar reduction capacity as wt A1M, while M8H4DK-tagged M41K-A1M and R66H-A1M showed a slightly lower ABTS reduction capacity (Figure 5A). Shortening or removal of the tag had little effect, but the R66H-N17,96D variants show a tendency to higher reduction capacity than wt A1M when M8H-tagged and untagged. The performance of M8H4DK-tagged wt A1M in the ORAC assay was set to 100% and the other variants were compared in relation to this. All M8H4DK- variants showed the same results in ORAC as M8H4DK-wt except M8H4DK-R66H+N17,96D which had significantly higher capacity (Figure 5b). Also the M8H-tagged and untagged R66H+N17,96D-A1M showed a slightly higher ORAC capacity than wt-A1M. The reduction capacity was also investigated in a cytochrome c reduction assay (Figure 6). Most variants showed a slightly lower reduction capacity at the lower concentrations, compared to wt-A1M, and for N17,96D and R66H+N17,96D this was significant (at 0.3-0.6 μ M) (Figure 6A). The importance of this is not understood. Shortening and removal of the tag had no influence of this property.

Heme-binding was investigated in a free heme incorporation assay and in a heme-agarose binding assay (Figure 6). The incorporation and coordination of free heme is seen as the appearance of an absorbance peak around 410-415 nm, the so-called Soret band (Karnaukhova et al., 2014; Rutardottir et al, 2016). Heme normally has an absorbance peak around 380-390nm and this can be seen when it is dissolved alone or mixed and incubated with a non-binding control protein like ovalbumin. However, when heme is mixed with wt A1M, an absorbance peak is formed between 410-415nm, yielding a so-called red-shift (i.e. a shift of the peak towards higher wave-lengths), and an increased Abs(413/386) ratio. It was shown that the degree of red-shift and the Abs(413/386) ratio are related to the heme-binding function of A1M (Rutardottir et al, 2016). These properties were therefore compared for all variants in phase III (Figure 6b). Among the M8H4DK-tagged variants all showed about the same red-shift as wt A1M except the M41K variants which had a stronger shift. Interestingly, shortening the tag to M8H increased the red-shift, while removal of the tag decreased the red-shift. Clearly, the tag had a strong influence of this property. The binding of heme to A1M was investigated as binding of a dilution series of A1M to heme-agarose in comparison

to binding to a control agarose. Typically, a control protein like ovalbumin shows no binding to heme-agarose above the background binding to the control agarose (Figure 6C). Typically, 70-80% of added M8H4DK-wt A1M (35-A1M) was bound to the heme agarose, whereas no-binding was seen to control agarose. The binding was compared
5 for all variants in phase III. All showed similar degree of binding to the heme-agarose (66-72%), while ovalbumin showed no binding.

Heme binding of M8H4DK-wt A1M (wt-A1M) and M8H4DK-35-A1M (35-A1M) was further analysed (Figure 11) using a combination of migration shift and fluorescence analysis (Karnaukhova et al. 2014). As a result of heme-incorporation, the migration of wt-A1M and 35-A1M in native PAGE analysis was slower at heme:protein ratio < 1, and faster at heme:protein ratio > 1, showing the same dependence upon heme-concentration (Fig. 11A and B). At high heme concentrations, both variants showed a tendency towards oligomerization, supporting previous findings for wt-A1M (Karnaukhova et al.
10 2014). Likewise, heme-incorporation induced quenching of tryptophan fluorescence in both wt-A1M and 35-A1M, with similar kinetics (Fig. 11B). The coordination of heme in A1M was previously shown to induce formation of a UV-absorbance peak at 415 nm (Ruttarsdottir et al., 2016; karnaukhova et al., 2014). Similar UV-absorbance patterns of wt-A1M- and 35-A1M/heme complexes were seen (Fig. 11C), with only a small red-
15 20 shift of the 35-A1M peak (415→417 nm).

The rate of reduction of ABTS-radicals of wt-A1M and 35-A1M was investigated [10] and was similar for wt-A1M and 35-A1M using unstressed, freshly prepared proteins (Fig. 12A). After storage for 7 days, wt-A1M displayed a slower reduction rate after
25 concentration to 1 mM and at 22°C (Fig. 12B), suggesting a loss of protein activity in addition to the aggregation shown at these conditions (Table 11). The cytochrome c-reduction (Allhorn et al., 2005) was slightly slower for 35-A1M (Fig. 12C), whereas the antioxidation capacity measured by the ORAC assay was somewhat higher for 35-A1M (Fig. 12D).

Table 11. Physicochemical properties of M8H4DK-wt A1M (wt-A1M) and M8H4DK-35-A1M (35-A1M).

	Conditions ¹	Concentration	Method	Parameter	wt-A1M	35-A1M
5	<i>Unstressed²</i>					
		0.1 mM	HPLC	Elution time (min)	12.1	12.1
				Monomer (%)	87	89
			DLS	Radius (nm)	2.9	3.2
			PAGE	Aggregates (%)	1.5	0.4
10			SEC	Monomer (%)	87	93
<hr/>						
	<i>Stressed</i>					
	Heat	0.1 mM	DSF	T _m (°C)	46.6	51.6
	High conc	1 mM	PAGE	Aggregates (%)	1.1	0.4
15	pH 7.4 ³	1 mM	PAGE	Aggregates (%)	1.2	0.4
	Freeze-thaw ²	1 mM	PAGE	Aggregates (%)	2.0	0.4
	Freeze-thaw/PBS	1 mM	PAGE	Aggregates (%)	13.2	7.6
<hr/>						
	<i>Prolonged storage⁴</i>					
20	4°C – 1 day	1 mM	SEC	Recovery ⁵ (%)	100	100
				Aggregates (%)	17.9	3.9
	4°C – 7 days	1 mM	SEC	Recovery (%)	100	100
				Aggregates (%)	12.1	3.7
	RT – 7 days	1 mM	SEC	Recovery (%)	44	100
25				Aggregates (%)	0.2	0.1
	37°C – 1.5 h	1 mM	SEC	Recovery (%)	67	100
				Aggregates (%)	53	28
	37°C – 4.5 h	1 mM	SEC	Recovery (%)	0	44
				Aggregates (%)	ND ⁶	15

30

Footnotes:

1. Room temperature unless otherwise stated.

2. 20 mM Tris-HCl, 0.125 M NaCl, pH 8.0

3. 20 mM Tris-HCl, 0.125 M NaCl, pH 7.4

35 4. PBS

5. Calculated after centrifugation 10,000xg, from total peak area compared to starting material.

6. Not determined

- 5 A biologic assay where the ability of the A1M variants to rescue K562 cells from heme-induced cell death was investigated (Figure 7). The degree of cell death, according to lactate dehydrogenase release, was set to 100% for cells incubated with heme without A1M. All M8H4DK-tagged variants, including wt-A1M, showed almost 100% rescue at 10 μ M (Figure 7), with no significant differences between the variants. The M8H-tagged
- 10 variants showed almost the same rescuing potential, while a slightly lower potential was observed for the untagged variants.

Cell protection capacity of M8H4DK-wt A1M (35-A1M) and M8H4DK-35-A1M (wt-A1M)

- The cell protection properties of the two A1M-variants were tested in K562 cells and a
- 15 human kidney proximal tubule epithelial cell line (HK-2), exposed to free heme and free iron. Fig. 12 shows that both A1M-variants completely inhibited the cell-death, measured by extracellular release of the cytosolic marker LDH, of K562-cells exposed to 0.1 mM heme. The dose-response curves of wt-A1M and 35-A1M overlaps almost completely, suggesting similar cell-protection capacities.

20

- Protection of HK-2 cells are shown in Fig. 14. First, cell damage was induced by the Fenton reaction, a mixture of free iron, ascorbate and hydrogen peroxide which generates hydroxyl radicals. Cell viability, measured by the WST-1 assay, was restored by both A1M-variants, following overlapping dose-response curves (Fig. 14A). The upregulation of heme oxygenase-1 (HO-1), a well-documented biomarker of oxidative stress-
- 25 response (Alam et al., 1999), was significantly suppressed by wt-A1M and 35-A1M to similar degrees (Fig. 14B). Cell damage of HK-2 cells was also induced by incubation with heme, and could be inhibited by both A1M-variants (Fig. 14C and D). Again, cell viability measured by the WST-1 assay was restored by both proteins to a similar de-
- 30 gree, using two heme concentrations, 10 and 30 μ M (Fig. 14C). The upregulation of HO-1 and another cellular stress response gene, Hsp70, was inhibited to a similar degree by both A1M-variants (Fig. 14D).

In vivo distribution of M8H4DK-wt A1M (wt-A1M) and M8H4DK-35-1M (35-A1M) in

- 35 mice

Intravenously injected A1M was previously shown to be rapidly cleared from blood and predominantly localized to kidneys and liver in rats and mice (Larsson et al., 2001; Ahlstedt et al., 2015). We compared the clearance rates in blood and distribution in organs of wt-A1M and 35-A1M (Fig. 15). Similar turnover rates were seen during the first 60 min, whereas 35-A1M was cleared more rapidly after 1h. The distribution of A1M in the investigated organs after 10 and 30 min did not show any significant differences between the two proteins. Both wt-A1M and 35-A1M were found predominantly in kidneys, and smaller amounts were seen in heart, liver, lung, skin and spleen, while negligible amounts were found in brain.

In vivo protection of kidneys by M8H4DK-wt A1M (wt-A1M) and M8H4DK-35-A1M (35-A1M)

Previously, A1M has shown *in vivo* therapeutic effects in animal models where heme- and oxidative stress-related kidney injuries are induced (Wester-Rosenlöf et al., 2014; Nääv et al., 2015; Sverrisson et al., 2014). Here, we investigated the *in vivo* protective effects of the two A1M-variants in a glycerol-injection rhabdomyolysis mouse model where acute kidney injuries (AKI) develops as a result of muscle rupture with release of myoglobin, free heme, radicals and other tissue components. The glycerol-injection resulted in a massive upregulation of the HO-1 and Hsp70 genes, two biomarkers of cellular stress (Fig. 16A and B), and most of the upregulation was inhibited by simultaneous injection of wt-A1M or 35-A1M. No significant difference between the two A1M-variants were seen at the applied doses (7 mg/kg animal weight).

In summary, the investigations in phase III showed that the N17,96D-containing A1M-mutations significantly added stability to the A1M molecule, which was further improved when the R66H mutation was added. Furthermore, M8H4DK-tagged A1M variants were more stable than variants with a shorter tag or no tag. Therefore, the tag does not only serve as a purification tool but also provide A1M with increased stability and stability and does not interfere with function. The M8H4DK-R66H+N1796D (35-A1M) molecule show the same or better functional properties, possibly with the exception of the cytochrome c reduction. In conclusion, stability/solubility and functional studies suggest that M8H4DK-R66H+N17,96D (35-A1M) has improved molecular properties compared to M8H4DK-wt A1M (wt-A1M), and both Wt-A1M and 35-A1M show *in vivo* protective effects on kidneys using a rhabdomyolysis glycerol-injection model of acute kidney injury.

Further stability studies on M8H4DK-R66H+N1796D vs M8H4DK-wt.

To further compare M8H4DK-R66H+N17,96D-A1M and wt-A1M, aggregation and function were tested with SEC-FPLC and the ABTS reduction assay after freeze-thaw cycles and storage at +4°C and room temperature (Figure 8). Samples (100µM) of wt-A1M and R66H+N17,96D-A1M both tolerated 5x freeze-thaw cycles very well. As previously shown wt-A1M showed more aggregation after concentration to 1mM than R66H+N17,96D-A1M, and similar results were seen after storage for 1 week compared to overnight. The increased aggregation of wt-A1M did not affect the ABTS reduction activity of 2µM-samples. Ocular inspections show completely clear solutions of both variants after freeze-thaw, concentration and storage for 1 week at +4°C. To stress the A1M solutions further, 100µM and 1mM solutions were stored in room temperature for 1 week. Ocular inspection of the 100µM solutions showed a slight cloudiness of wt-A1M, but the R66H+N17,96D-A1M solution remained clear. SEC-FPLC showed a decreased area under the curve of wt-A1M to 80% of the starting material, indicating loss of wt-A1M, although no increase in large aggregates was seen. The loss was most likely caused by precipitation of protein and removal by the filtration before the FPLC-run. For R66H+N17,96D-A1M, some increased aggregation could be seen, but no loss of protein after the storage at room-temperature for one week at 100 µM. In the ABTS assay, R66H+N17,96D-A1M still showed full activity, while a small decrease was seen for wt-A1M. Ocular inspection of 1mM solutions stored for 1 week showed cloudiness of both variants, but the wt-A1M solution appeared thicker. When the solutions were diluted 10 times for loading onto the SEC, the precipitate of the R66H+N17,96D-A1M, but not wt-A1M, was resolved. In the SEC, very little aggregates were seen, but only 44% of the wt-A1M starting material remained. In contrast, 100% of R66H+N17,96D-A1M remained. In the ABTS assay wt-A1M had seriously decreased activity while the activity of R66H+N17,96D-A1M stayed the same. This study shows that M8H4DK-R66H+N17,96D-A1M can tolerate storage at high concentrations at room temperature, whereas wt-A1M cannot.

The stability of 1mM solutions at 37°C was investigated. The samples were incubated for 1.5, 2.5 and 4.5 h and then ocularly inspected, analysed by SEC-FPLC and the ABTS reduction assay (Figure 9). After 1.5 h incubation, both solutions were still clear when ocularly inspected, but the SEC-FPLC revealed more aggregates in the wt-A1M solution (53%) compared to the R66H+N17,96D-A1M solution (28%). Only 67% of the starting material was eluted on the SEC of wt-A1M, while 100% was eluted of R66H+N17,96D-A1M, estimated by the area under the curve. Both variants showed full

activity in the ABTS assay. After 2.5 h of incubation the wt-A1M solution had become cloudy with precipitates impossible to resolve. 40% of the starting material was found on the SEC and the ABTS activity was significantly decreased. The R66H+N17,96D-A1M solution was still clear and showed full activity in the ABTS assay, but large aggregates had increased from 28 to 38% in the SEC-FPLC. After 4.5 h, the wt-A1M had developed into a thick substance, impossible to pipett. Therefore, it was not further analysed. The R66H+N17,96D-A1M had become cloudy, but was still possible to pipette. Some precipitation remained after dilution. The amount of protein eluted on the SEC was 44% of the starting material and a decreased activity in the ABTS assay was seen.

The study suggest that following incubation at +37°C, A1M first forms resolvable aggregates, then it forms irreversible aggregates and loss of activity in the ABTS assay. Complete irreversible precipitation of wt-A1M was seen after 2.5 hours of incubation, and partial irreversible precipitation was seen only after 4.5 hours for M8H4DK-R66H+N17,96D-A1M. Hence, M8H4DK-R66H+N17,96D-A1M is more resistant to storage at +37°C than M8H4DK-wt-A1M.

References

- Ahlstedt J, Tran TA, Strand F, Holmqvist B, Strand S-E, Gram M, and Åkerström B. Biodistribution and pharmacokinetics of recombinant α_1 -microglobulin and its potential use in radioprotection of kidneys. *Am J Nucl Med Mol Imaging* 5(4): 333-347, 2015.
- Alam J, Stewart D, Touchard C, Boinapally S, Choi AM, and Cook JL. Nrf2, a Cap'n'Collar transcription factor, regulates induction of the heme-oxygenase gene. *J Biol Chem* 274: 26071-26078, 1999.
- Allhorn M, Berggård T, Nordberg J, Olsson ML, Åkerström B. Processing of the lipocalin α_1 -microglobulin by hemoglobin induces heme-binding and heme-degradation properties. *Blood* 99; 2002: 1894-1901.
- Allhorn M, Klapyta A, Åkerström B. Redox properties of the lipocalin α_1 -microglobulin : reduction of cytochrome c, hemoglobin, and free iron. *Free Radic Biol Med* 38; 2005: 557-567.
- Anderson UD, Rutardottir S, Centlow M, Olsson MG, Kristensen, K, Isberg PE, Thilaganathan B, Åkerström B, Hansson SR. Fetal hemoglobin and α_1 -microglobulin as first and

early second trimester predictive biomarkers of preeclampsia. *Am J Obst Gynecol* 204; 2011: 520.e1-520.e5

- 5 Berggård T, Cohen A, Persson P, Lindqvist A, Cedervall T, Silow M, Thøgersen IB, Jönsson J-Å, Enghild JJ, Åkerström B. α_1 -Microglobulin chromophores are located to three lysine residues semiburied in the lipocalin pocket and associated with a novel lipophilic compound. *Protein Sci.* 8; 1999: 2611-2620.

- 10 Berggård T, Thelin N, Falkenberg C, Enghild JJ, Åkerström B. Prothrombin, albumin and immunoglobulin A form covalent complexes with α_1 -microglobulin in human plasma. *Eur J Biochem* 245; 1997: 676-683.

- 15 Bratt T, Olsson H, Sjöberg EM, Jergil B, Åkerström B. Cleavage of the α_1 -microglobulin-bikunin precursor is localized to the Golgi apparatus of rat liver cells. *Biochim Biophys Acta* 1157; 1993: 147-154.

- 20 Centlow M, Carninci P, Nemeth K, Mezey E, Brownstein M, Hansson SR. Placental expression profiling in preeclampsia: local overproduction of hemoglobin may drive pathological changes. *Fertility and sterility* 90; 2008: 1834-1843.

Ekström B., Berggård I. Human α_1 -microglobulin: Purification procedure, chemical and physicochemical properties. *J Biol Chem* 252(22); 1977: 8048-8057.

- 25 Escribano J, Lopez-Otin C, Hjerpe A, Grubb A, Mendez E. Location and characterization of the three carbohydrate prosthetic groups of human protein HC. *FEBS Lett* 266(1-2); 1990: 167-170.

Flower DR. The lipocalin family. *Biochem J* 318; 1996: 1-14.

- 30 Gram M, Anderson UD, Johansson ME, Edström-Hägerwall A, Larsson I, Jälmby M, Hansson SR, and Åkerström B. The human endogenous protection system against cell-free hemoglobin and heme is overwhelmed in preeclampsia and provides potential biomarkers and clinical indicators. *PloS One* 10(9): e0138111, 2015.

Hansson SR, Gram M and Åkerström B. Fetal hemoglobin in preeclampsia: a new etiological tool for prediction/diagnosis and a potential target for therapy. *Curr Opin Obstet Gynecol.* 25(6); 2013: 448-455.

- 5 Karnaukhova E, Rutardottir S, Majabi M, Wester Rosenlöf L, Alayash AI, Åkerström B. Characterization of heme binding to recombinant α_1 -microglobulin. *Front Physiol* 5; 2014: 465. doi: 10.3389/fphys.2014.00465.

- 10 Kaumeyer JF, Polazzi JO, Kotick MP. The mRNA for a proteinase inhibitor related to the HI-30 domain of inter-alpha-trypsin inhibitor also encodes α_1 -microglobulin (protein HC). *Nucleic Acids Res* 14(20); 1986: 7839-7850.

- Kwasek A, Osmark P, Allhorn M, Lindqvist A, Åkerström B, Wasylewski Z. Production of recombinant human α_1 -microglobulin and mutated forms involved in chromophore
15 formation. *Prot. Expr. Purif.* 53; 2007: 145-152.

Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227; 1970: 680-685.

- 20 Larsson J, Wingårdh K, Berggård T, Davies JR, Lögdberg L, Strand S-E, Åkerström B. Distribution of ^{125}I -labelled α_1 -microglobulin in rats after intravenous injection. *J Lab Clin Med* 137: 165-175, 2001

- Lindqvist A, Bratt T, Altieri M, Kastern W, Åkerström B. Rat α_1 -microglobulin: co-expression in liver with the light chain of inter-alpha-trypsin inhibitor. *Biochim Biophys Acta* 1130; 1992: 63-67.
25

- May K, Rosenlöf L, Olsson MG, Centlow M, Mörgelin M, Larsson I, Cederlund M, Rutardottir S, Schneider H, Siegmund W, Åkerström B, Hansson SR. Perfusion of human
30 placenta with haemoglobin introduces preeclampsia-like injuries that are prevented by α_1 -microglobulin. *Placenta* 32(4); 2011: 323-332.

Meining W, Skerra A. The crystal structure of human α_1 -microglobulin reveals a potential haem-binding site. *Biochem J.* 445(2); 2012: 175-82.

- Nääv Å, Erlandsson L, Axelsson J, Larsson I, Johansson M, Wester Rosenlöf L, Mörgelin M, Casslén V, Gram M, Åkerström B, Hansson SR. A1M ameliorates preeclampsia-like symptoms in placenta and kidney induced by cell-free fetal hemoglobin in rabbit. *PloS One*, 10(5); 2015: e0125499.
- 5 Olsson MG, Allhorn M, Larsson J, Cederlund M, Lundqvist, K, Schmidtchen A, Sørensen OE, Mörgelin M, Åkerström B. Up-regulation of A1M/ α_1 -microglobulin in skin by heme and reactive oxygen species gives protection from oxidative damage. *PLoS One* 6(11); 2011: e27505.
- 10 Olsson MG, Centlow M, Rutardottir S, Stenfors I, Larsson J, Hosseini-Maaf B, Olsson ML, Hansson SR, Åkerström B. Increased levels of free hemoglobin, oxidation markers, and the antioxidative heme scavenger α_1 -microglobulin in preeclampsia. *Free Rad. Biol. Med.* 48; 2010: 284-291.
- 15 Olsson MG, Olofsson T, Tapper H, Åkerström B. The lipocalin α_1 -microglobulin protects erythroid K562 cells against oxidative damage induced by heme and reactive oxygen species. *Free Rad Res.* 42; 2008: 725-736.
- 20 Olsson MG, Rosenlöf LW, Kotarsky H, Olofsson T, Leanderson T, Mörgelin M, Fellman V, Åkerström B. The radical-binding lipocalin A1M binds to a Complex I subunit and protects mitochondrial structure and function. *Antiox Redox Signal.* 18(16); 2013: 2017-2028.
- 25 Rutardottir S, Karnaukhova E, Nantasenamat C, Songtawee N, Prachayasittikul V, Rajabi M, Wester Rosenlöf L, Alayash AI, Åkerström B. Studies of two heme binding sites in A1M/ α_1 -microglobulin using site-directed mutagenesis and molecular simulation. *Biochim Biophys Acta* 1864(1); 2016: 29-41.
- 30 Sala A, Campagnoli M, Perani E, Romano A, Labó S, Monzani E, Minchiotti L, Galliano M. Human α_1 -microglobulin is covalently bound to kynurenine-derived chromophores. *J Biol Chem* 279(49); 2004: 51033-51044.

Sverrisson K, Axelsson J, Rippe A, Gram M, Åkerström B, Hansson SR. Rippe B. Extracellular fetal hemoglobin (HbF) induces increases in glomerular permeability. Inhibition with α_1 -microglobulin (A1M) and Tempol. *Am J Physiol Renal Physiol* 306(4); 2014: F442-448.

5

Wester-Rosenlöf L, Casslén V, Axelsson J, Edström-Hägerwall A, Gram M, Holmquist M, Johansson ME, Larsson I, Ley D, Marsal K, Mörgelin M, Rippe B, Rutardottir S, Shohani B, Åkerström B, Hansson SR. A1M/ α_1 -microglobulin protects from heme-induced placental and renal damage in a pregnant sheep model of preeclampsia. *PLoS One* 9(1); 2014: e86353.

10

Åkerström B, Borregaard N, Flower D, Salier JP. Lipocalins, an introduction. In *Lipocalins*. Eds. Åkerström B, Borregaard N, Flower D, Salier JP, Landes Bioscience, Georgetown TX.

15

Åkerström B, Gram M. A1M, an extravascular tissue cleaning and house-keeping protein. *Free Rad Biol Med* 74; 2014: 274-282.

20

Åkerström B, Maghzal GJ, Winterbourn CC, Kettle AJ. The lipocalin α_1 -microglobulin has radical scavenging activity. *J Biol Chem* 282; 2007: 31493-31503.

Åkerström B, Bratt T, Enghild JJ. Formation of the α_1 -microglobulin chromophore in mammalian and insect cells: a novel post-translational mechanism? *FEBS Lett* 362, 50-54, 1995.

25

Claims

1. An alpha-1-microglobulin derived protein selected from the amino acid sequence consisting of formula I:

X¹-X²-X³-X⁴-X⁵-X⁶-X⁷-X⁸-X⁹-X¹⁰-X¹¹-X¹²-X¹³-X¹⁴-GPVPTPPDN IQVQENF-X¹⁵-IS
 RIYGKWYNLA IGSTCPWLKK I-X¹⁶-DRMTVSTL VLGEGATEAE ISMTST-X¹⁷-WRK
 GVCEETSGAY EKTDTDGKFL YHKS KW-X¹⁸-ITM ESYVVHTNYD EYAIFLTKKF
 SRHHGPTITA KLYGRAPQLR ETLLQDFRVV AQGVGIPEDS IFTMADRGE C
 VPGEQEPEPI LIPR (SEQ ID NO: 10)

and the amino acid sequence consisting of formula II:

X¹-X²-X³-X⁴-X⁵-X⁶-X⁷-X⁸-X⁹-X¹⁰-X¹¹-X¹²-X¹³-X¹⁴-GPVPTPPDN IQVQENF-X¹⁵-IS
 RIYGKWYNLA IGSTCPWLKK I-X¹⁶-DRMTVSTL VLGEGATEAE ISMTST-X¹⁷-WRK
 GVCEETSGAY EKTDTDGKFL YHKS KW-X¹⁸-ITM ESYVVHTNYD EYAIFLTKKF
 SRHHGPTITA KLYGRAPQLR ETLLQDFRVV AQGVGIPEDS IFTMADRGE C
 VPGEQEPEPI (SEQ ID NO: 17),

wherein at least one X is present;

X¹ is absent or represents Met or N-formyl Met;

X² is absent or represents His;

X³ is absent or represents His;

X⁴ is absent or represents His;

X⁵ is absent or represents His;

X⁶ is absent or represents His;

X⁷ is absent or represents His;

X⁸ is absent or represents His;

X⁹ is absent or represents His;

X¹⁰ is absent or selected from Asp, Glu, Lys, or Arg;

X¹¹ is absent or selected from Asp, Glu, Lys, or Arg;

X¹² is absent or selected from Asp, Glu, Lys, or Arg;

X¹³ is absent or selected from Asp, Glu, Lys, or Arg;

X¹⁴ is absent or selected from Asp, Glu, Lys, or Arg;

X¹⁵ represents Asp;

X¹⁶ represents Met;

X¹⁷ represents Arg or His or Lys;

X¹⁸ represents Asp;

or a pharmaceutically acceptable salt thereof, wherein the alpha-1-microglobulin derived protein has improved thermostability compared to human wild-type alpha-1-microglobulin

2. The alpha-1-microglobulin derived protein according to claim 1, wherein

X¹ is Met or N-formyl Met,

X², X³, X⁴, X⁵, X⁶, X⁷ are His,

X⁸ and X⁹ are absent,

X¹⁰, X¹¹, X¹² and X¹³ are Asp,

X¹⁴ is Lys,

and X¹⁵-X¹⁸ are as defined above.

3. The alpha-1-microglobulin derived protein according to claim 1, wherein

X¹ is Met or N-formyl Met,

X², X³, X⁴, X⁵, X⁶, X⁷, X⁸ and X⁹ are His,

X¹⁰, X¹¹, X¹² and X¹³ are Asp,

X¹⁴ is Lys,

and X¹⁵-X¹⁸ are as defined above.

4. The alpha-1-microglobulin derived protein according to claim 1, wherein

X¹ is Met,

X², X³, X⁴, X⁵, X⁶, X⁷, X⁸, X¹⁰, X¹¹, X¹², X¹³, and X¹⁴, are absent,

X¹⁵ is Asp,

X¹⁶ is Met,

X¹⁷ is His,

and X¹⁸ is Asp.

5. The alpha-1-microglobulin according to claim 4 having an amino acid sequence corresponding to SEQ ID NO: 3 or SEQ ID NO: 18.

6. The alpha-1-microglobulin derived protein according to claim 1, wherein

X¹ is Met or N-formyl Met,

X², X³, X⁴, X⁵, X⁶, X⁷ are His,

X⁸ and X⁹ are absent,

X¹⁰, X¹¹, X¹² and X¹³ are Asp,
 X¹⁴ is Lys,
 X¹⁵ is Asp,
 X¹⁶ is Met,
 X¹⁷ is His,
 and X¹⁸ is Asp.

7. The alpha-1-microglobulin derived protein according to claim 6 having an amino acid sequence corresponding to SEQ ID NO: 5 or SEQ ID NO: 20.
8. The alpha-1-microglobulin derived protein according to claim 1, wherein
 X¹ is Met or N-formyl Met,
 X², X³, X⁴, X⁵, X⁶, X⁷, X⁸ and X⁹ are His,
 X¹⁰, X¹¹, X¹² and X¹³ are Asp,
 X¹⁴ is Lys,
 X¹⁵ is Asp,
 X¹⁶ is Met,
 X¹⁷ is His,
 and X¹⁸ is Asp.
9. The alpha-1-microglobulin derived protein according to claim 8 having an amino acid sequence corresponding to SEQ ID NO: 7 or SEQ ID NO: 22.
10. The alpha-1-microglobulin derived protein according to any of the preceding claims, wherein said protein is an antioxidant.
11. The alpha-1-microglobulin derived protein according to any of the preceding claims, wherein said protein binds heme.
12. A pharmaceutical composition comprising an alpha-1-microglobulin derived protein according to any one of claims 1-17 and one or more pharmaceutically acceptable excipients.
13. An alpha-1-microglobulin derived protein according to any of the claims 1-17 for use in medicine.