



US 20250257106A1

(19) **United States**
(12) **Patent Application Publication** (10) **Pub. No.: US 2025/0257106 A1**
MA et al. (43) **Pub. Date: Aug. 14, 2025**

(54) **COMPLEMENT-INHIBITING HYBRID PROTEIN**

Publication Classification

(71) Applicant: **LongBio Pharma (Suzhou) Co., Ltd.**, Shanghai (CN)

(51) **Int. Cl.**
C07K 14/47 (2006.01)
A61K 38/00 (2006.01)
C07K 14/705 (2006.01)
C12P 21/02 (2006.01)
(52) **U.S. Cl.**
CPC *C07K 14/472* (2013.01); *C07K 14/70596* (2013.01); *C12P 21/02* (2013.01); *A61K 38/00* (2013.01); *C07K 2319/00* (2013.01)

(72) Inventors: **Haili MA**, Shanghai (CN); **Heng LIU**, Shanghai (CN)

(21) Appl. No.: **18/856,041**

(22) PCT Filed: **Apr. 6, 2023**

(86) PCT No.: **PCT/CN2023/086593**

§ 371 (c)(1),

(2) Date: **May 2, 2025**

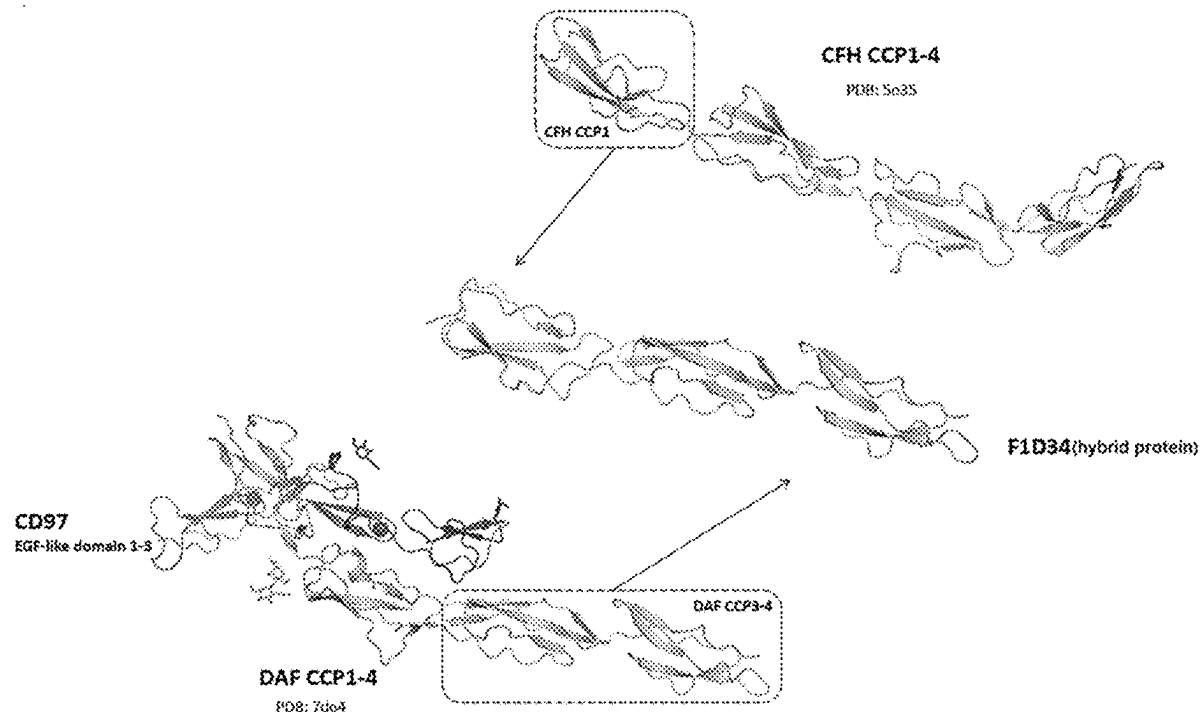
(30) **Foreign Application Priority Data**

Apr. 11, 2022 (CN) 202210374811.2

(57) **ABSTRACT**

Provided are a hybrid protein comprising a CFH functional unit and a DAF functional unit, an encoding nucleic acid thereof, and a method and use thereof for treating related diseases or disorders of a complement system.

Specification includes a Sequence Listing.



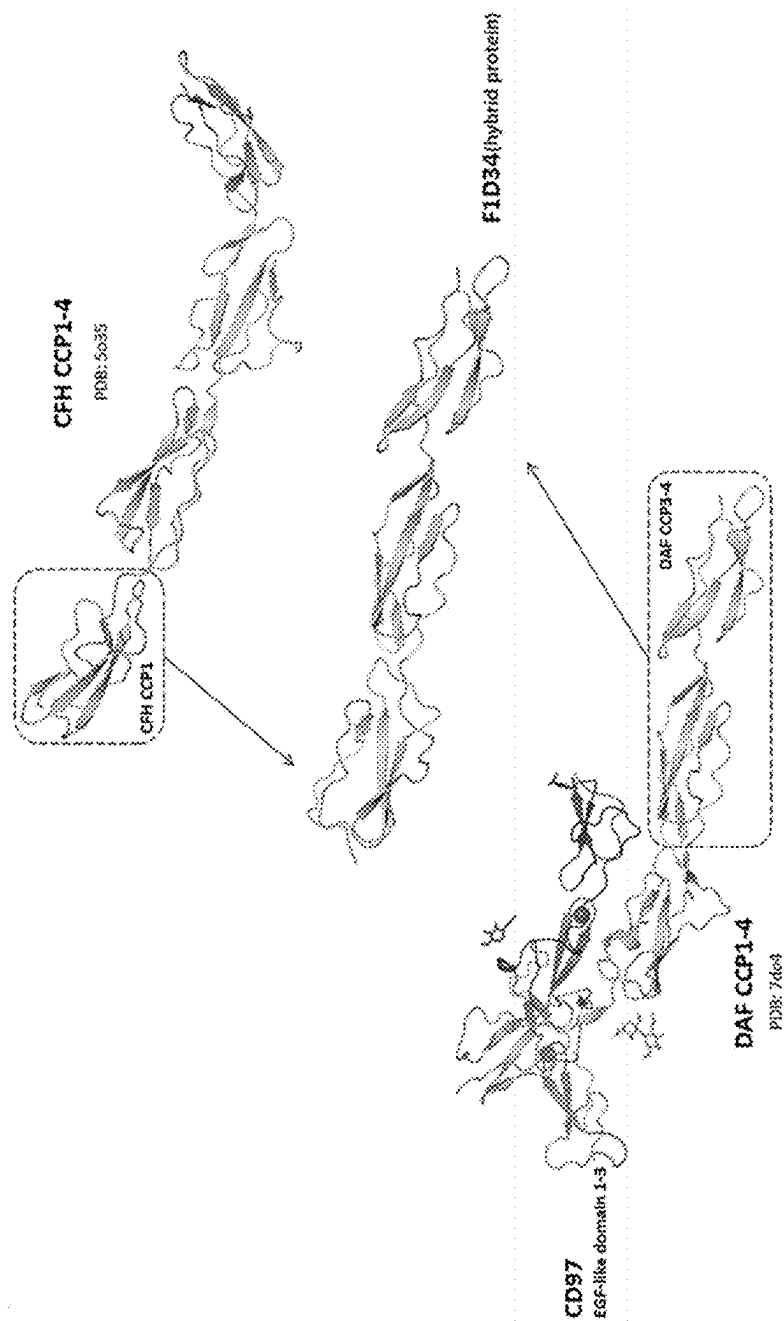


FIG. 1

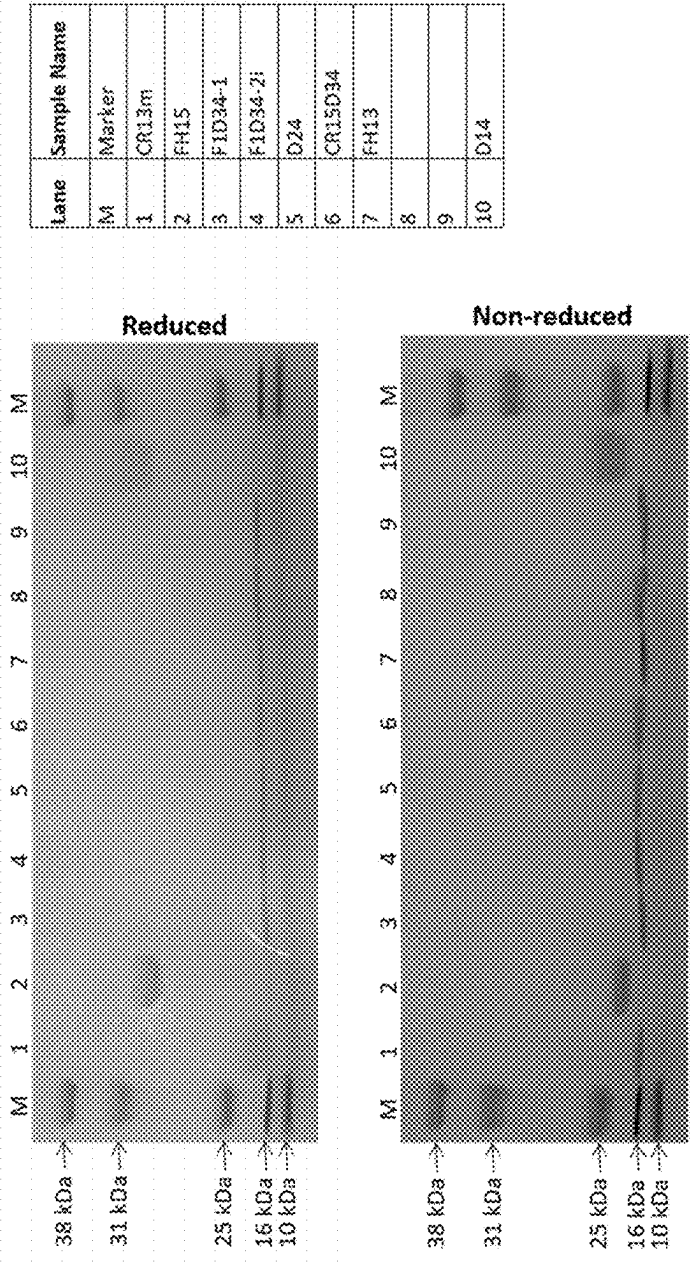


FIG. 2

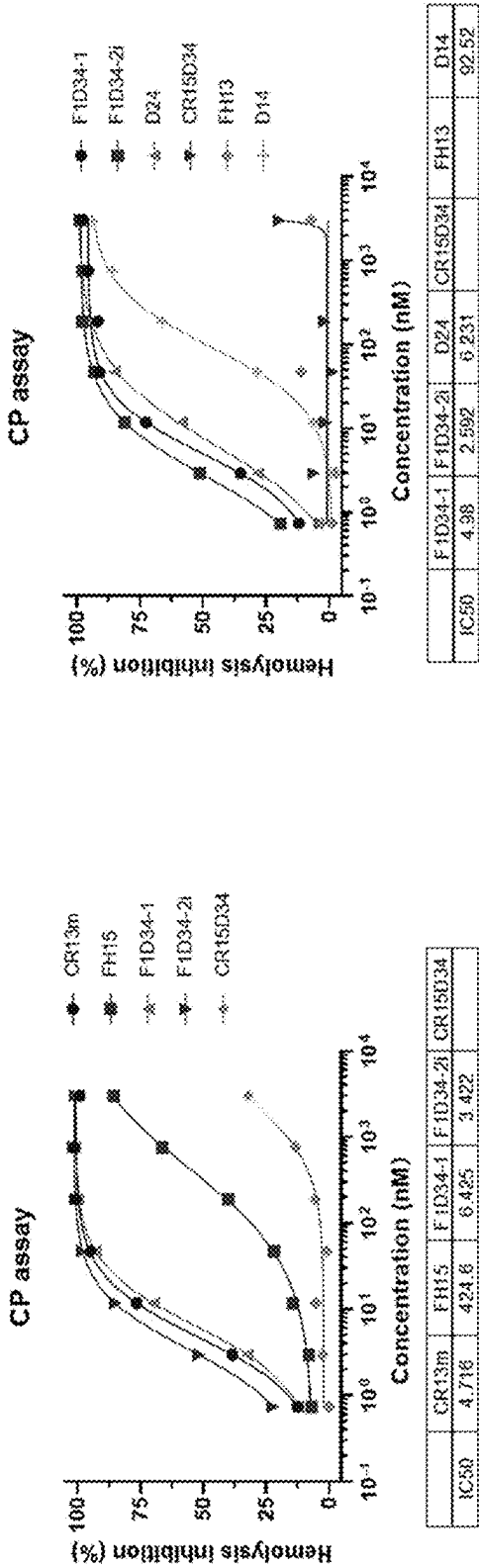


FIG. 3

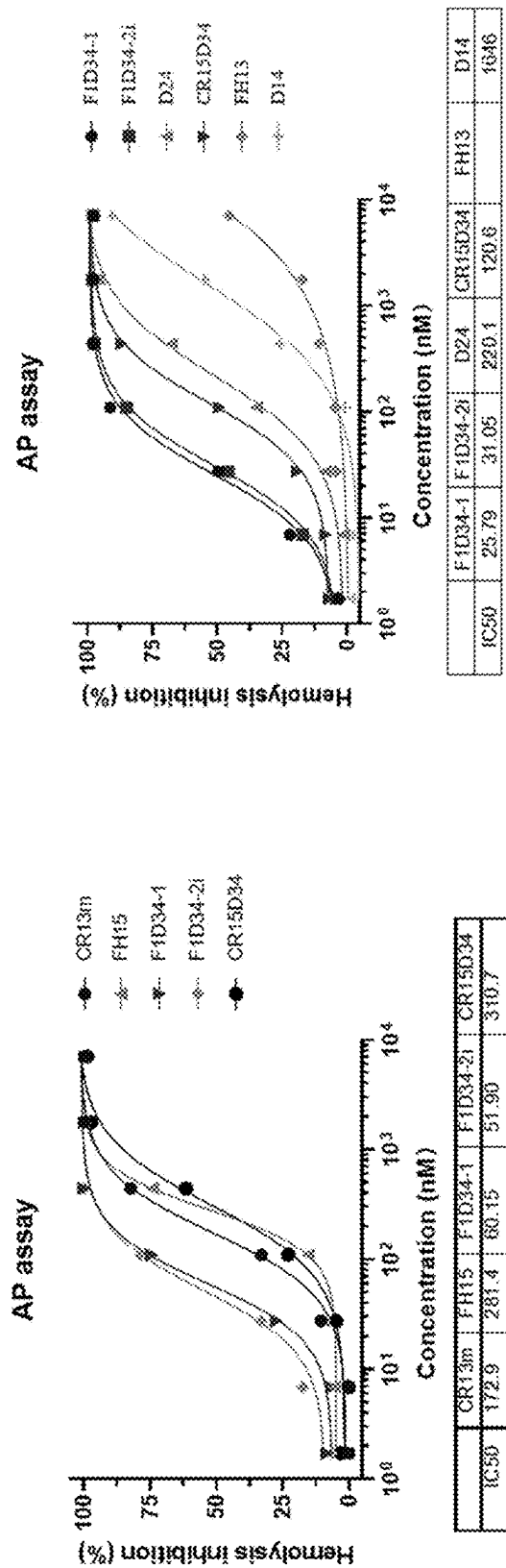


FIG. 4

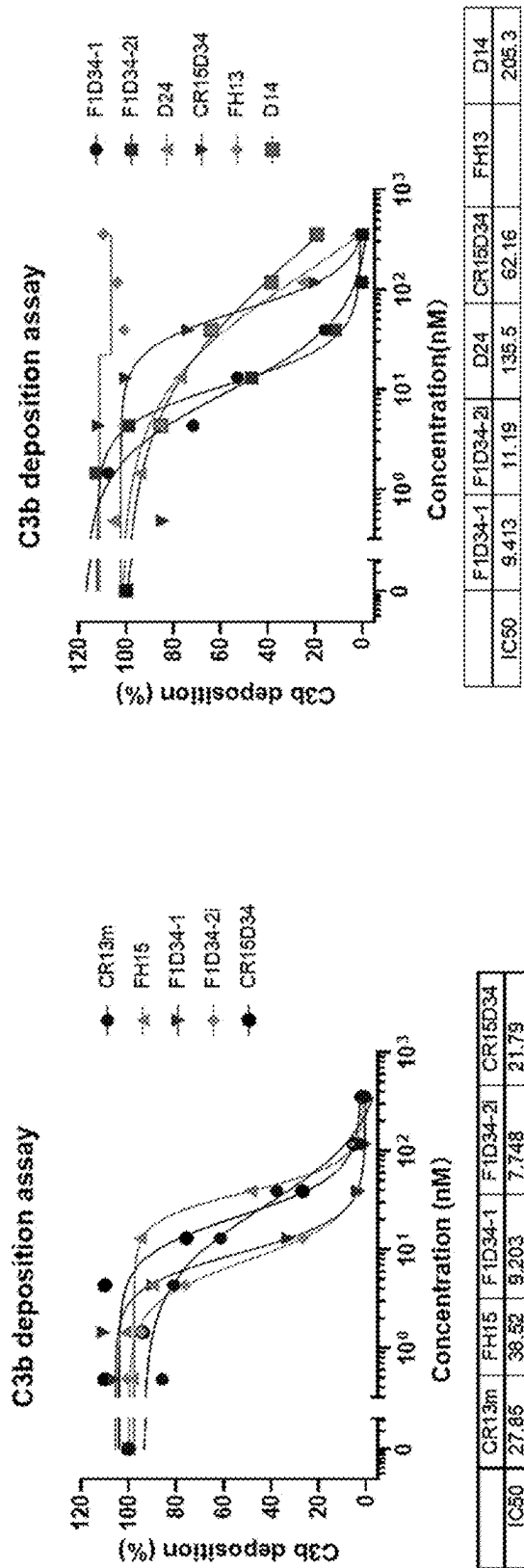


FIG. 5

COMPLEMENT-INHIBITING HYBRID PROTEIN

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a U.S. National Stage of PCT International Application No. PCT/CN2023/086593 filed Apr. 6, 2023, which claims the benefit of Chinese Application No. 202210374811.2 filed Apr. 11, 2022, the content of each of which is incorporated herein by reference in its entirety.

[0002] The present invention relates to a hybrid protein comprising a CFH functional unit and a DAF functional unit, an encoding nucleic acid thereof, and a method and a use thereof for treating a disease or a disorder related to the complement system.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[0003] The contents of the electronic sequence listing “398857.xml” size 48,151 bytes, and Date of Creation Feb. 4, 2025, is herein incorporated by reference in its entirety.

BACKGROUND

[0004] As part of the innate immune system, the complement system comprises more than 30 kinds of soluble protein molecules and more than 30 kinds of molecules such as intrinsic complement components, various regulatory factors, and complement receptors. It is well known that the complement system acts directly on pathogens to promote pathogen-specific adaptive immune responses; and participates in cell differentiation and polarization, tissue regeneration, lipid metabolism, clearance of immune complexes, and apoptosis, etc. As the deficient or inappropriate complement activation regulation will lead to damage to host tissues, the complement system is tightly regulated by a series of proteins, wherein complement activation regulator (RCA) family proteins are primarily responsible for complement regulation; RCA proteins include membrane proteins such as decay-accelerating factor (DAF; CD55), membrane cofactor protein (MCP; CD46) and complement receptor 1 (CR1; CD35), as well as fluid-phase proteins such as Factor H (FH or CFH) and C4b-binding protein (C4BP). The RCA protein structure consists of complement control protein repeat (CCP) modules, of which 2-4 consecutive modules contribute to regulatory functions known as decay-accelerating activity (DAA) and cofactor activity (CFA). RCA proteins act by targeting the C3/C5 convertase, the central enzyme of the complement pathway. RCA proteins bind to these convertases or their non-catalytic subunits to inactivate these enzymes. The DAA was displayed as the irreversible dissociation of the invertase into subunits after binding to the RCA protein, whereas the CFA was displayed as the cleavage and inactivation of the invertase via the recruitment of serine protease Factor I (FI) after the binding RCA protein to the non-catalytic subunits (C3b/C4b), thereby preventing the formation of C3 convertase. Factor H is an opsonin and ligand for complement receptors 2 and 3, acting as a cofactor for Factor I in the catalyzed cleavage of C3b to iC3b, and inhibiting C3b amplification. It accelerates the irreversible dissociation of C3 convertase C3bBb in the alternative pathway, and may also compete with Factor B for C3B binding during the formation of proconvertase. Factor H is

a soluble complement regulator essential for protecting surfaces, including extracellular matrix ECM. FH binds to the peptide hormone adrenomedullin and may prevent its degradation. The FH plays a role in managing cellular senescence, stress, or injury through interactions with C-reactive proteins, pentameric proteins, DNA, histones, annexin II, the malondialdehyde acetaldehyde adduct of proteins, and oxidized lipids. FHLI also has cofactor activity for Factor I and C3bBb decay-accelerating activity (The Complement FactsBook. Edited by: Scott Barnum and Theresa Schein, Copyright® 2018 Elsevier Ltd. All rights reserved. <https://doi.org/10.1016/C2015-0-06595-9>, Chapter 30).

[0005] DAF intrinsically protects host cells from autologous complement attack by preventing the formation and accelerating the decay of classical and alternative C3 and C5 convertases, thereby inhibiting the cleavage of C3 and C5. When purified DAF is added to cells, it can be incorporated into the cell membrane, thus displaying functional activity. DAF has also been found to regulate T cell tolerance and thus negatively regulate several animal models of autoimmune diseases (The Complement FactsBook. Edited by: Scott Barnum and Theresa Schein, Copyright 2018 Elsevier Ltd. All rights reserved. <https://doi.org/10.1016/C2015-0-06595-9>, Chapter 25).

[0006] The present invention relates to a hybrid protein with better complement regulation activity, and it is hoped that this hybrid protein can have advantages such as better therapeutic effect.

DESCRIPTION OF THE INVENTION

[0007] The invention relates to a hybrid protein of CFH and DAF, wherein CCP1/SCR1 of CFH is hybridized with CCP of DAF (e.g., CCP3-4/SCR3-4 of DAF) to form the hybrid protein.

[0008] The hybrid proteins have better complement-inhibiting activity, including the activity in inhibiting the Alternative Pathway (AP), the Classical Pathway (CP) and/or C3b deposition.

[0009] In some embodiments, the hybrid protein is effective to block human complement hemolytic activity.

[0010] In some embodiments, the hybrid protein of the present invention is effective in inhibiting hemolysis of the human complement classical pathway and/or hemolysis of the human complement alternative pathway.

[0011] In some embodiments, the hybrid protein of the present invention is effective in inhibiting C3b deposition.

[0012] For example, the hybrid protein of the present invention further comprises less CCP of DAF on the basis of comprising CPP1/SCR1 of CFH, so that the hybrid protein having a smaller molecular weight can be obtained. The hybrid proteins, when in the therapeutic applications, can be administered at the same mass concentration with higher molarity doses, and thus have the potential of better therapeutic efficacy or longer dosing intervals, as compared to other typically used complement regulatory proteins of larger molecular weight (especially compared to hybrid proteins that also comprise other CCPs of CFH), e.g. in ophthalmic, brain medications. In other cases, smaller molecular weights are also more conducive to tissue penetration, such as tumor tissue, blood-brain barrier, and the like.

[0013] In some embodiments, the present invention relates to the following specific items:

[0014] 1. A hybrid protein, which comprises or consists of

[0015] (i) CCP1 of human complement Factor H (CFH),

[0016] (ii) CCP3 and CCP4 of human decay-accelerating factors (DAFs), and optionally a signal peptide and/or a tag.

[0017] 2. The hybrid protein of embodiment 1, wherein the CCP3 and CCP4 of DAF are directly linked together to form CCP3-4.

[0018] 3. The hybrid protein of embodiment 1 or 2, wherein the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-82 of the human CFH protein, optionally, the CCP1 comprises a V62I mutation, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3.

[0019] 4. The hybrid protein of embodiment 1 or 2, wherein the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-84 of the human CFH protein, optionally, the CCP1 comprises a V62I mutation, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3.

[0020] 5. The hybrid protein of any one of embodiments 1-4, wherein the CCP3 of human DAF comprises or consists of the amino acid sequence at positions 161-222 of the human DAF protein, and/or the CCP4 of human DAF comprises or consists of the amino acid sequence at positions 223-285 of the human DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.

[0021] 6. The hybrid protein of any one of embodiments 1-4, wherein the CCP3 of human DAF comprises or consists of the amino acid sequence 163-222 of the human DAF protein, and/or the CCP4 of human DAF comprises or consists of the amino acid sequence 223-285 of the human DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.

[0022] 7. The hybrid protein of embodiment 2, wherein the CCP3-4 of human DAF comprises or consists of the amino acid sequence at positions 161-285 of the human DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.

[0023] 8. The hybrid protein of any one of embodiments 2-4, wherein the CCP3-4 of human DAF comprises or consists of the amino acid sequence at position 163-285 of the human DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.

[0024] 9. The hybrid protein of embodiment 2, wherein

[0025] (1) the CCP1 of human CFH comprises or consists of the amino acid sequences at positions 19-82 of the human CFH protein, and the CCP3 of human DAF comprises or consists of the amino acid sequence 161-222 of the human DAF protein, and the CCP4 of human DAF comprises or consists of the amino acid sequence 223-285 of the human DAF protein;

[0026] (2) the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-84 of the human CFH protein, and the CCP3 of human DAF comprises or consists of the amino acid sequence 163-222 of the human DAF protein, and the CCP4 of human DAF comprises or consists of the amino acid sequence 223-285 of the human DAF protein;

[0027] (3) the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-82 of the human CFH protein, and the CCP3-4 of human DAF comprises or consists of the amino acid sequence at positions 161-285 of the human DAF protein; or

[0028] (4) the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-84 of the human CFH protein, and the CCP3-4 of human DAF comprises or consists of the amino acid sequence at positions 163-285 of the human DAF protein;

[0029] wherein the amino acid positions of the human CFH protein are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3 and the amino acid positions of the human DAF protein are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.

[0030] 10. The hybrid protein of embodiment 9, wherein the CCP1 of human CFH has a V62I mutation.

[0031] 11. The hybrid protein of any one of embodiments 1-10, wherein the human CFH protein is a native human CFH protein or comprises

[0032] (i) the amino acid sequence as set forth in SEQ ID NO: 3 or 5;

[0033] (ii) the amino acid sequence encoded by the nucleotide sequence as set forth in SEQ ID NO: 4 or 6;

[0034] (iii) the amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in (i) or (ii); or

consists of the amino acid sequence as set forth in any one of (i) to (iii).

[0035] 12. The hybrid protein of any one of embodiments 1-11, wherein the human DAF protein is a native human DAF protein or comprises

[0036] (i) the amino acid sequence as set forth in SEQ ID NO: 1;

[0037] (ii) the amino acid sequence encoded by the nucleotide sequence as set forth in SEQ ID NO: 2;

[0038] (iii) the amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in (i) or (ii); or

consists of the amino acid sequence as set forth in any one of (i) to (iii).

[0039] 13. The hybrid protein of any one of embodiments 1-12, wherein

[0040] the CCP1 of human CFH comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 12, 13, 14, or 15, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 12, 13, 14, or 15;

[0041] the CCP3 of human DAF comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 7 or 8, or an amino acid sequence having at

least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 7 or SEQ ID NO: 8;

[0042] the CCP4 of human DAF comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 9, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 9; and/or

[0043] the CCP3-4 of human DAF comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 10 or SEQ ID NO: 11, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 10 or SEQ ID NO: 11.

[0044] 14. The hybrid protein of any one of embodiments 1-13, wherein the tag is, for example, a purification tag such as a hexahistidine tag or a biotin tag, comprising, e.g., the amino acid sequence as set forth in SEQ ID NO: 12.

[0045] 15. The hybrid protein of any one of embodiments 1-14, wherein the signal peptide is secretory, comprising, e.g., the amino acid sequence as set forth in SEQ ID NO: 17.

[0046] 16. The hybrid protein of any one of embodiments 1-15, wherein the hybrid protein comprises or consists of the amino acid sequence of any one of SEQ ID NOs: 18-33, or an amino acid sequence having at least a 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence.

[0047] 17. A nucleic acid molecule encoding the hybrid protein of any one of embodiments 1-16.

[0048] 18. An expression vector comprising the nucleic acid molecule of embodiment 17, preferably the expression vector is pcDNA3.1.

[0049] 19. A host cell comprising the nucleic acid molecule of embodiment 17 or the expression vector of embodiment 18, preferably the host cell is prokaryotic or eukaryotic, such as a 293 cell, e.g., an Expi293 cell.

[0050] 20. A method of preparing the hybrid protein of any one of embodiments 1-16, the method comprising culturing the host cell of embodiment 19 under conditions suitable for expression of the hybrid protein, and optionally further comprising isolating the protein from the host cell or the host cell culture medium, and/or purifying the protein.

[0051] 21. A pharmaceutical composition or medication or formulation comprising the hybrid protein according to any one of embodiments 1-16 or a nucleic acid molecule encoding the same, or the nucleic acid molecule of embodiment 17, and optionally a pharmaceutically acceptable excipient.

[0052] 22. A pharmaceutical combination product comprising

[0053] the hybrid protein according to any one of embodiments 1-16 or a nucleic acid molecule encoding the same, or the nucleic acid molecule of embodiment 17; and the additional therapeutic agents.

[0054] 23. A method of preventing or treating a complement system-associated disease or disorder in a subject, comprising administering to the subject an effective amount of the hybrid protein of any one of embodiments 1-16 or a nucleic acid molecule encoding the same; or the nucleic acid molecule of embodiment 17;

or the pharmaceutical composition or formulation of embodiment 21; or the pharmaceutical combination product of embodiment 22.

[0055] 24. The method of embodiment 23, wherein the disease or disorder is due to an abnormal activation of the complement system or a dysregulation of the complement system.

[0056] 25. The method of embodiment 24, wherein the abnormal activation of the complement system or the dysregulation of the complement system is due to for example, a microbial infection or an increase in auto-immune antibodies, or a decrease, deletion, incapacitation, or interference or blockade of a function of complement regulatory proteins.

[0057] 26. The method of embodiment 23, wherein the complement system-associated disease or disorder is selected from a disease requiring hemolysis inhibition, such as a disease requiring inhibition of hemolysis of the classical pathway of the complement immunity and/or the alternative pathway of the complement immunity; or a disease requiring inhibition of C3b deposition activity, such as dense deposition disease DDD.

BRIEF DESCRIPTION OF DRAWINGS

[0058] FIG. 1 shows the crystal structures of CFH CCP1-5 and DAF CCP1-4 and the structural model of the hybrid protein.

[0059] FIG. 2 shows the SDS-PAGE electrophoresis pattern of each protein to be tested.

[0060] FIG. 3 shows the CP inhibitory activity of each protein to be tested.

[0061] FIG. 4 shows the AP inhibitory activity of each protein to be tested.

[0062] FIG. 5 shows the C3b deposition inhibitory activity of each protein to be tested.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

I. Definition

[0063] It is to be understood that this invention is not limited to the specific methodologies, protocols, and reagents described herein, as these may vary. It should also be understood that the terminology used herein is for describing particular embodiments only, and is not intended to limit the scope of the present invention, which will be limited only by the appended claims. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

Decay-Accelerating Factor (DAF)

[0064] Decay-accelerating Factor (DAF, CD55) is a membrane-associated regulatory protein that protects autologous cells from activation of autologous complement on their surfaces. DAF acts by rapidly dissociating C3 and C5 convertases (central enzymes of the cascade). DAF has the most potent attenuation-accelerating activity of proteins associated with complement regulation and acts on classical pathway enzymes (C4b2a and C4b2a3b) and alternative pathway enzymes (C3bBb and C3BbC3b). However, DAF has no cofactor function.

[0065] Structural analysis of the DAF shows that, starting from its N-terminus, it consists of a unit of 4-60 amino acids in length, followed by a fragment (STP) rich in highly O-glycosylated serine (S) and threonine (T), followed by a post-translationally added glycosylphosphatidylinositol (GPI) anchor.

[0066] In some embodiments, the DAF is a human DAF having the following accession numbers: Genbank accession numbers M31516, M15799, M64653, S72858, or M643567.

[0067] In some embodiments, the DAF is human DAF. In some embodiments, the human DAF is native human DAF. In some embodiments, the amino acid sequence of native human DAF is set forth in SEQ ID NO: 1, with the various modules shown in Table 1 below. Four 60 amino acid long repeating units termed complement control protein repeats (CCP) or Short Consensus Repeats (SCR). CCP1 includes amino acids 35-96; CCP2 includes amino acids 96-160; CCP3 includes amino acids 161-222 and CCP4 includes amino acids 223-285. They provide all the regulatory activities of DAF. The highly O-glycosylated region serves as a cushion that positions the CCP at an appropriate distance suspended above the surface membrane. The GPI anchor allows DAF to move freely on the plasma membrane, making it possible to inactivate invertase complexes anywhere they are assembled. In this context, reference to the amino acid positions of the modules of DAF is made to the amino acid position numbering corresponding to that as set forth in SEQ ID NO: 1.

TABLE 1	
Protein modules of DAF (divided by UniProt way)	
The amino acid position of SEQ ID NO:1	Module
1-34	Lead sequence
35-96	CCP1
96-160	CCP2
161-222	CCP3
223-285	CCP4
287-356	STP

[0068] In some embodiments, the nucleotide sequence of the cDNA encoding DAF is set forth in SEQ. ID NO: 2.

Factor H (CFH or FH)

[0069] “Complement Factor H”, “Factor H”, “FH”, “CFH protein” or “CFH” are used interchangeably and refer to a protein of approximately 150 kDa that is a member of the regulators of the complement activation family and is a complement control protein. CFH is a large soluble glycoprotein that circulates in human plasma and serves to regulate the alternative pathway of the complement system, ensuring that the complement system is directed towards pathogens or other dangerous substances and does not damage host tissues.

[0070] Factor H is predominantly monomeric but can be weakly self-associated (KD=28 μM) and may be oligomeric in the presence of glycosaminoglycans or high concentrations of metal ions. CFH consists of 20 homologous units called complement control protein repeats (CCPs) (SCR or sushi domains), some of which function for cell attachment, while other repeats function to eliminate C3b from the cell surface. Each of the 20 SCRs is about 60 amino acids in

length, arranged head to tail, comprising 4 cysteine residues, forming 2 disulfide bonds for each module. SCR 19 and 20 are involved in C3b binding.

[0071] The splice variant FHL-1 consists of the first 7 CCPs, followed by the C-terminal sequence Ser-Pro-Leu-Thr. Each CCP comprises approximately 60 residues, including four invariant cysteines, forming Cys^I-Cys^{III}, Cys^{II}-Cys^{IV} disulfides. Adjacent modules are connected by sequences of three to eight residues.

[0072] The images from negative staining electron microscopy show that FH molecules adopt a variety of conformations, but are mainly folded in half; analytical ultracentrifugation, small angle X-ray scattering (deposition of the module in PDB, e.g., 3 GAV), and chemical cross-linking also indicated that the module chain itself was bent. The following high resolution structure of several FH fragments was determined, alone or in complex with other molecules (PDB identifier is provided in parenthesis below): CCP 1-2 (2RLP), 2-3 (2RLQ), 1-4 (2WII), 5, 6-7 (e.g., 2W80, 2 YBY), 7 (2JGW and 2JGX), 6-8 (2 UWN and 2V8E), 9 (4K12), 10-11 (4B2R), 11-12 (4B2S), 12-13 (2KMS), 15 (1HFI), 16 (1HCC), 15-16 (1HFI), 18-20 (3SWO), 19-20 (e.g., 2BZM, 2G7I and 4ONT). Each CCP resembles a prolate spheroid having a major axis of about 4 nm and a minor axis of about 2 nm and comprises beta strands in anti-parallel patches that are approximately aligned with the major axis, with the N- and C-termini located at either end of its long axis, facilitating an end-to-end arrangement of tandem CCPs with variable inter-module contact, tilt, and twists.

[0073] In some embodiments, the CFH is a human CFH, e.g., a native human CFH. In some embodiments, the amino acid sequence of CFH has the following accession numbers: HGNC: HGNC: 4883, Ensembl: ENSG00000000971, HPRD: 00601, MIM: 134370 or Vega: OTTHUMG00000035607.

[0074] In some embodiments, the amino acid sequence of CFH is set forth in SEQ ID NO: 3. In some embodiments, the amino acid positions corresponding to each module of the CFH are shown in Table 2 below.

TABLE 2	
Protein modules of CFH (divided by UniProt way)	
The amino acid position of SEQ ID NO: 3	Module
1-18	Lead peptide
19-82	CCP1
83-143	CCP2
144-207	CCP3
208-264	CCP4
265-322	CCP5
324-386	CCP6
387-444	CCP7
446-507	CCP8
515-566	CCP9
567-625	CCP10
628-686	CCP11
689-746	CCP12
751-805	CCP13
809-866	CCP14
868-928	CCP15
929-986	CCP16
987-1045	CCP17
1046-1104	CCP18
1107-1165	CCP19
1170-1230	CCP20

[0075] In some embodiments, the nucleotide sequence of the cDNA encoding CFH is set forth in SEQ. ID NO: 4.

[0076] In some embodiments, the protein sequence encoding the splice variant FHL-1 of CFH is set forth in SEQ ID NO: 5. In some embodiments, the nucleotide sequence of the cDNA encoding the splice variant FHL-1 of CFH is set forth in SEQ ID NO: 6.

[0077] In this context, reference to the amino acid positions of the modules of CFH is made to the amino acid position numbering corresponding to that as set forth in SEQ ID NO: 3.

Other Definitions

[0078] For the purpose of interpreting this specification, the following definitions will be used and, wherever appropriate, terms used in the singular may also include the plural and vice versa. It is to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting.

[0079] The term “about” when used in conjunction with a numerical value is intended to encompass a range of numerical values having a lower limit that is 5% less than the specified numerical value and an upper limit that is 5% greater than the specified numerical value.

[0080] As used herein, the term “and/or” means any one of the options or two or more or all of the options.

[0081] As used herein, the term “comprises” or “includes” means the inclusion of the stated element, integer, or step, but not the exclusion of any other element, integer, or step. When the term “comprises” or “includes” is used herein, it also encompasses a combination of the stated elements, integers, or steps unless otherwise indicated. For example, reference to a protein “comprising” a specific sequence is intended to also encompass the protein consisting of that specific sequence.

[0082] As used herein, the terms “hybrid protein” and “chimeric polypeptide” are used interchangeably and refer to a larger polypeptide formed by the fusion of at least two heterologous polypeptide sequences, optionally through a linker. The hybrid proteins can be produced by recombinant expression.

[0083] “Percent identity (%)” of an amino acid sequence refers to the percentage of amino acid residues in a candidate sequence that are identical to the amino acid residues of the specific amino acid sequence shown in this specification, after aligning the candidate sequence with the specific amino acid sequence shown in this specification and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and without taking into account any conservative substitutions included as part of sequence identity. In some embodiments, the present invention contemplates variants of a protein or polypeptide of the present invention that have a substantial degree of identity, e.g., at least 80%, 85%, 90%, 95%, 97%, 98%, or 99% or more identity, relative to the polypeptide or protein specifically disclosed herein. The variants may comprise conservative changes.

[0084] For polypeptide sequences, “conservative alterations” include substitutions, deletions, or additions to the polypeptide sequence that do not substantially change the desired functional activity of the polypeptide sequence. For example, conservative substitutions often result in the replacement of an amino acid with a chemically similar amino acid. Conservative substitution tables providing func-

tionally similar amino acids are well known in the art. 8 groups of amino acids conservatively substituted for each other are listed as follows: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M). In some embodiments, the term “conservative sequence alteration” is used to refer to amino acid modifications that do not significantly affect or alter the activity of the parent hybrid protein. For example, conservatively modified variants retain at least 80%, 85%, 90%, 95%, 98%, 99%, or more, e.g., 100-110% or more activity relative to the parent polypeptide or hybrid protein.

[0085] The term “host cell” refers to a cell into which an exogenous polynucleotide has been introduced, including the progeny of such cells. Host cells include “transformants” and “transformed cells,” which include primary transformed cells and progeny derived therefrom. Host cells are any type of cell system that can be used to produce hybrid proteins of the present invention, including eukaryotic cells, such as mammalian cells, insect cells, yeast cells; and prokaryotic cells, such as *E. coli* cells. Host cells include cultured cells, as well as cells within transgenic animals, transgenic plants, or cultured plant tissue or animal tissue.

[0086] The term “expression vector” refers to a vector comprising a recombinant polynucleotide comprising expression control sequences operably linked to the nucleotide sequence to be expressed. The expression vector comprises sufficient cis-acting elements for expression; other elements for expression can be provided by the host cell or in an in vitro expression system. Expression vectors include all those known in the art, including cosmids, plasmids (e.g., naked or included in liposomes), and viruses (e.g., lentiviruses, retroviruses, adenoviruses, viruses, and adeno-associated viruses) into which recombinant polynucleotides are incorporated.

[0087] The terms “individual” or “subject” are used interchangeably and refer to a mammal. Mammals include, but are not limited to, domesticated animals (e.g., cows, sheep, cats, dogs, and horses), primates (e.g., humans and non-human primates such as monkeys), rabbits, and rodents (e.g., mice and rats). In particular, individuals are people.

[0088] The term “treating” includes administering a composition or hybrid polypeptide to prevent or delay the onset of symptoms, complications, or biochemical indication of a disease, alleviate symptoms, or prevent or inhibit further progression of the disease, condition, or disorder. The term “prevention” includes the inhibition of the occurrence or progression of a disease or disorder or symptoms of a particular disease or disorder.

[0089] The term “pharmaceutical excipient” refers to diluents, adjuvants (such as Freund’s adjuvant (complete and incomplete)), excipients, carriers or stabilizers, etc., which are administered with the active substance.

[0090] The term “pharmaceutical composition” refers to a composition that is in a form effective to permit the biological activity of the active ingredients comprised therein and does not comprise additional ingredients that would have unacceptable toxicities to the subject to whom the composition is administered.

[0091] The term “effective amount” refers to an amount or dosage of a hybrid protein of the present invention or a

nucleic acid encoding the same or a composition or combination thereof that, when administered to a patient in single or multiple doses, produces the desired effect in a patient in need of treatment or prevention.

[0092] A “therapeutically effective amount” refers to an amount effective to achieve the desired therapeutic result, at the required doses and for the required period of time. A therapeutically effective amount is also an amount in which any toxic or detrimental effects of the hybrid protein or composition or combination are outweighed by the therapeutically beneficial effects. A “therapeutically effective amount” preferably inhibits a measurable parameter or improves a measurable parameter by at least about 40%, even more preferably at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% or even 100%, relative to an untreated subject.

[0093] A “prophylactically effective amount” refers to an amount effective to achieve the desired prophylactic result, at the required dosage and for the required period of time. Generally, the prophylactically effective amount will be less than the therapeutically effective amount because the prophylactic dose is administered in the subject before or at an earlier stage of the disease.

[0094] These and other aspects and embodiments of the present invention are described in the accompanying drawings (Brief Description of the Drawings immediately follows) and the following detailed description of the present invention and are illustrated in the following examples. Any or all features discussed above and throughout this application may be combined in various embodiments of the present invention. The following examples further illustrate the present invention, however, it is to be understood that the examples are described by way of illustration rather than limitation and that various modifications may be made by those skilled in the art.

II. Hybrid Protein

[0095] The present invention relates to a hybrid protein of CFH and DAF, which comprises at least one functional unit derived from CFH and at least one functional unit derived from DAF.

A. DAF Functional Unit

[0096] In certain embodiments, the hybrid protein of the present invention preferably comprises a functional unit from DAF. Such functional units are capable of dissociating the C3 and C5 convertases and/or accelerating the decay-accelerating activity against C3 convertases in the classical pathway and/or in the alternative pathway. In some embodiments, the DAF functional unit comprises CCPs 3 and 4 of the DAF.

[0097] The amino acid sequence of such CCPs can be identical to the native or naturally occurring amino acid sequence of DAF. Alternatively, the amino acid sequence of such CCPs may be slightly altered, particularly at the amino or carboxy terminus. This alteration occurs when the restriction enzyme site is incorporated into the polynucleotide encoding the CCP. Such alterations also occur when an amino acid is deleted from the N-terminus or C-terminus of the functional unit. For example, in some embodiments, 1-2 amino acids may be deleted at the N-terminus of CCP3.

[0098] Some amino acid substitutions, preferably conservative substitutions, in the sequence may also be introduced

without affecting functional activity. Conservative substitutions may be made, for example, by substituting charged amino acids for one another or substituting hydrophilic amino acids for one another, hydrophobic amino acids for one another, and amino acids of similar mass for one another.

[0099] In one embodiment, the DAF functional unit comprises an amino acid sequence that is at least 95%, 96%, 97%, 98%, or 99% identical to CCP3 and/or 4 of the human DAF protein (or CCP3-4 directly linked together by the C-terminus of CCP3 and the N-terminus of CCP4). In some embodiments, the DAF functional unit comprises CCP3 and/or 4 of the human DAF protein (CCP3-4 directly linked together by the C-terminus of CCP3 and the N-terminus of CCP4). In some embodiments, the DAF functional unit comprises CCP3 and 4 of the human DAF protein. In some embodiments, the DAF function unit is CCP3-4, which is directly linked together by the C-terminus of CCP3 and the N-terminus of CCP4.

[0100] In some embodiments, the human DAF protein is a native human DAF protein. In some embodiments, the native human DAF protein comprises or consists of the amino acid sequence shown as SEQ ID NO: 1. DAF comprises or consists of the amino acid sequence encoded by the DNA sequence as set forth in SEQ ID NO: 2.

[0101] In some embodiments, CCP3 comprises or consists of amino acids 161-222 of DAF. In some embodiments, 1-2 amino acids of the N-terminus of CCP3 may be deleted for attachment of other functional units, e.g., CCP3 comprises or consists of amino acids 163-222 of DAF.

[0102] In some embodiments, CCP3 comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 7 or 8, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 7 or SEQ ID NO: 8.

SEQ ID NO: 7:
KSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGS
SVQWSDPLPECRE

SEQ ID NO: 8:
CPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSV
QWSDPLPECRE

[0103] In some embodiments, CCP4 comprises or consists of amino acids 223-285 of DAF.

[0104] In some embodiments, CCP4 comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 9, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 9.

SEQ ID NO: 9:
IYCPAPPQIDNGIIQGERDHYGRQSVTYACNKGFTMIGEHSIYCTVMN
DEGEWSGPPPECRG

[0105] In some embodiments, CCP3-4 comprises or consists of amino acids 161-285 or amino acids 163-285 of DAF. In some embodiments, the DAF functional unit CCP3-4 comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 10 or SEQ ID NO: 11, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 10 or SEQ ID NO: 11.

SEQ ID NO: 10: Human DAF CCP3-4
KSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGS
SVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFT
MIGEHSIYCTVNNDEGEWSGPPPECRG

SEQ ID NO: 11: Human DAF CCP3-4 (2 amino acids
 are missing from the N-terminus of CCP3)
CPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSV
QWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMI
GEHSIYCTVNNDEGEWSGPPPECRG

[0106] When referring to the amino acid position of DAF, the amino acid sequence position corresponds to the amino acid position of SEQ ID NO: 1.

B. CFH Functional Unit

[0107] In certain embodiments, the hybrid protein of the present invention comprises functional units from CFH. This functional unit is capable of dissociating the alternative pathway C3 convertase C3bBb and/or binding to C3b. In a preferred embodiment, the CFH functional unit comprises CCP1 of CFH.

[0108] The amino acid sequence of the CCP of such CFH may be identical to the native or naturally occurring amino acid sequence of CFH. Alternatively, the amino acid sequence of such CCPs may be slightly altered, particularly at the amino or carboxy terminus. This change occurs when the restriction enzyme site is incorporated into the polynucleotide encoding the CCP. Such changes also occur when amino acids are deleted from or added to the N-terminus or C-terminus of the functional unit.

[0109] For example, in some embodiments, 1-2 amino acids may be deleted at the C-terminus of CCP1.

[0110] In some embodiments, 1-2 amino acids of CCP2 can also be added to the C-terminus of CCP1 to attach functional units of DAF, thus, "CCP1" as described herein encompasses the CCP1 with amino acids (e.g., 1-2 amino acids of CCP2) added at the C-terminus. For example, in the case where 1-2 amino acids are deleted from the N-terminus of the functional unit of DAF, 1-2 amino acids of CCP2 may be added to the C-terminus of CCP1 for linking the functional unit of DAF. In some embodiments, KS are deleted from the CCP3 of DAF, and RP are added to the C-terminus of CCP1 of CFH, both of which are linked to obtain a hybrid protein.

[0111] Some amino acid substitutions, preferably conservative substitutions, in the sequence, may also be introduced without affecting functional activity. Conservative substitutions may be made, for example, by substituting charged amino acids for one another or substituting hydrophilic amino acids for one another, hydrophobic amino acids for one another, and amino acids of similar mass for one another. For example, CCP1 of the CFH of the present invention can comprise a V62I mutation.

[0112] In one embodiment, the CFH functional unit comprises an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to CCP1 of the human CFH protein. In some embodiments, the CFH functional unit comprises CCP1 of the human CFH protein.

[0113] In some embodiments, the human CFH protein is a native human CFH protein. In some embodiments, the native human CFH protein comprises or consists of an amino acid sequence as set forth in SEQ ID NO: 3 or 5. In

some embodiments, the CFH comprises or consists of the amino acid sequence encoded by the DNA sequence as set forth in SEQ ID NO: 4 or 6.

[0114] In some embodiments, CCP1 comprises or consists of amino acids 19-82 of CFH. In some embodiments, CCP1 comprises or consists of amino acids 19 to 84 of CFH.

[0115] In some embodiments, the CFH functional unit comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 12, 13, 14, or 15, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 12, 13, 14, or 15.

SEQ ID NO: 12: Human CFH CCP1 (the C-terminus of CCP1 is added with two N-terminal amino acids RP of CCP2)
EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCR
KGEWVALNPLRKCQKRP

SEQ ID NO: 13: Human CFH CCP1 (CCP1)
EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCR
KGEWVALNPLRKCQK

SEQ ID NO: 14: Human CFH CCP1 (the C-terminus of CCP1 is added with two N-terminal amino acids RP of CCP2 and comprises V62I)
EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCR
KGEWVALNPLRKCQKRP

SEQ ID NO: 15: Human CFH CCP1 (CCP1 comprises V62I)
EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCR
KGEWVALNPLRKCQK

[0116] When referring to the amino acid position of CFH, the amino acid sequence position corresponds to the amino acid position of SEQ ID NO: 3.

C. Other Units

[0117] Optionally, the hybrid protein of the present invention may further comprise a tag, preferably of about 2-10 amino acids, added to the amino or carboxy terminus, e.g. the carboxy terminus, of the hybrid protein. Typically, such additions are made to stabilize the protein or to facilitate secretory expression or purification of the hybrid protein. Such tags are known in the art. Representative examples of such tags include sequences encoding a series of histidine residues (e.g., 2-10 histidines, e.g., 2, 3, 4, 5, 6, or 7 histidines), an epitope tag FLAG, herpes simplex glycoprotein D, beta-galactosidase, a maltose-binding protein, or glutathione S-transferase.

[0118] In some embodiments, the tag is a histidine residue, which may be added at the amino-terminus or the carboxy-terminus, e.g., the carboxy-terminus, of the hybrid protein.

[0119] In some embodiments, the tag is a 6xHis tag of GHHHHHH (SEQ ID NO: 16).

[0120] Optionally, the hybrid protein of the present invention may also include a signal peptide, such as MGWSCILFLVATATGVHS (SEQ ID NO: 17).

[0121] The present invention also encompasses hybrid proteins in which one or more amino acids are altered by post-translational processes or synthetic methods. Examples of such modifications include, but are not limited to, glycosylation, iodination, myristoylation, and PEGylation.

D. Examples of Hybrid Proteins

[0122] In some embodiments, the hybrid protein of the present invention comprises at least one functional unit from CFH and at least one functional unit from DAF.

[0123] In some embodiments, the hybrid protein of the present invention comprises CCP1 of CFH and CCP3 and CCP4 of DAF. In some embodiments, the hybrid protein of the present invention consists of CCP1 of CFH and CCP3 and CCP4 of DAF. In some embodiments, the hybrid protein of the present invention comprises or consists of CCP1 of CFH and CCP3-4 of DAF.

[0124] In some embodiments, 1-2 amino acids (e.g., 1-2 amino acids at the N-terminus of CCP2) are added to the C-terminus of CCP1 of CFH. In some embodiments, 1-2 amino acids are deleted from the N-terminus of CCP3 or CCP3-4 of DAF.

[0125] In some embodiments, the hybrid protein of the present invention comprises CCP1 of CFH and CCP3-4 of DAF, wherein 1-2 N-terminal amino acids of CCP2 are added to the C-terminus of CCP1 and 1-2 amino acids are deleted correspondingly from the N-terminus of CCP3 of DAF. In some specific embodiments, the hybrid protein of the present invention comprises CCP1 of CFH and CCP3-4 of DAF, wherein the N-terminal amino acid RP of CCP2 are added to the C-terminus of CCP1 and the amino acid KS are deleted correspondingly from the N-terminus of CCP3 of DAF.

[0126] In some embodiments, CCP1 of the CFH of the present invention can comprise a V62I mutation.

[0127] In some embodiments, the hybrid protein of the present invention further comprises a signal peptide at the N-terminus, such as the amino acid sequence as set forth in SEQ ID NO: 13, and/or a tag at the C-terminus, such as a histidine tag, such as a 6xHis-tag, such as GHHHHHHH.

[0128] In some embodiments, CCP3 of DAF comprises or consists of amino acids 161-222 of the DAF protein and/or CCP4 comprises or consists of amino acids 223-285 of the DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1. In some embodiments, the CCP3 of DAF is deleted by 2 amino acids from its N-terminus, for example comprising or consisting of amino acids 163-222 of the DAF protein, wherein said amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1. In some embodiments, the CCP3-4 of the DAF comprises or consists of amino acids 161-285 of the DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1. In some embodiments, the CCP3-4 of DAF is deleted at its N-terminus by 2 amino acids, for example comprising or consisting of amino acids 163-285 of the DAF protein, wherein said amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.

[0129] In some embodiments, CCP1 of CFH comprises or consists of amino acids 19-82 of the CFH protein, wherein

the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3. In some embodiments, CCP1 of CFH comprises or consists of amino acids 19-84 of the CFH protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3.

[0130] In some embodiments, CCP3-4 of DAF comprises or consists of amino acids 161-285 of the DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1; and CCP1 of CFH comprises or consists of amino acids 19-82 of the CFH protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3.

[0131] In some embodiments, CCP3-4 of DAF comprises or consists of amino acids 163-285 of the DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1; and CCP1 of CFH comprises or consists of amino acids 19-84 of the CFH protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3.

[0132] In some specific embodiments, the DAF is a human DAF protein, such as a native human DAF protein. In some specific embodiments, the DAF protein comprises or consists of

[0133] (i) the amino acid sequence as set forth in SEQ ID NO: 1;

[0134] (ii) the amino acid sequence encoded by the nucleotide sequence as set forth in SEQ ID NO: 2;

[0135] (iii) the amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in (i) or (ii); or

[0136] consists of the amino acid sequence as set forth in any one of (i) to (iii).

[0137] In some specific embodiments, CFH is a human CFH protein, such as a native human CFH protein. In some specific embodiments, the CFH protein comprises

[0138] (i) the amino acid sequence as set forth in SEQ ID NO: 3 or 5;

[0139] (ii) the amino acid sequence encoded by the nucleotide sequence as set forth in SEQ ID NO: 4 or 6;

[0140] (iii) the amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in (i) or (ii); or

[0141] consists of the amino acid sequence as set forth in any one of (i) to (iii).

[0142] In some embodiments, the hybrid protein of the present invention comprises or consists of the amino acid sequence of any one of SEQ ID NOs: 18-21, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence.

>Human CFH CCP1 or SCR1-Human DAF/CD55 CCP3-4 or SCR3-4 (in bold):
(SEQ ID NO: 18)
EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWALNPLRKQCK
RCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLEGSTSSFCILSGSSVQWSDPLPECREIY
CPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDEGEWSGPPPECRG

-continued

(SEQ ID NO: 19)

EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKQK
KSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLEGSTSSFCLISGSSVQWSDPLPECREIY
CPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG

>Human CFH CCP1 or SCR1-Human DAF/CD55 CCP3-4 or SCR3-4 (in bold),
comprising a V62I mutation:

(SEQ ID NO: 20)

EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKQK
KSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLEGSTSSFCLISGSSVQWSDPLPECREIY
CPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG

(SEQ ID NO: 21)

EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKQK
RPCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIY
CPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG

[0143] In some embodiments, the hybrid protein of the present invention comprises or consists of the amino acid sequence according to any one of SEQ ID NOs: 22-25, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence.

>Human CFH CCP1 or SCR1-Human DAF/CD55 CCP3-4 or SCR3-4 (in bold)-6xHis
(italics):

(SEQ ID NO: 22)

EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKQK

RPCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIY

CPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHH

HHHH

(SEQ ID NO: 23)

EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKQK

KSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIY

CPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHH

HHHH

>Human CFH CCP1 or SCR1-Human DAF/CD55 CCP3-4 or SCR3-4 (in bold)-6xHis
(italics):

comprising a V62I mutation:

(SEQ ID NO: 24)

EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKQK

KSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIY

CPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHH

HHHH

(SEQ ID NO: 25)

EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKQK

RPCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIY

CPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHH

HHHH

[0144] In some embodiments, the hybrid protein of the present invention comprises or consists of the amino acid sequence according to any one of SEQ ID NOs: 26-29, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence.

>Signal peptide-human CFH CCP1 or SCR1-human DAF/CD55 CCP3-4 or SCR3-4
(in bold)-6xHis (*italic*):

(SEQ ID NO: 26)

MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSDQTYPEGTQAIYKCRPGYRSLGNVI

MVCRKGEWVALNPLRKCQKRPCNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCL

ISGSSVQWSDPLPECEIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTV

NNDEGEWSGPPPECRGGHHHHHH

(SEQ ID NO: 27)

MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSDQTYPEGTQAIYKCRPGYRSLGNVI

MVCRKGEWVALNPLRKCQKSCNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCL

ISGSSVQWSDPLPECEIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTV

NNDEGEWSGPPPECRGGHHHHHH

>Signal peptide-human CFH CCP1 or SCR1-human DAF/CD55 CCP3-4 or SCR3-4
(in bold)-6xHis (*italic*):
comprising a V62I mutation:

(SEQ ID NO: 28)

MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSDQTYPEGTQAIYKCRPGYRSLGNII

MVCRKGEWVALNPLRKCQKSCNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCL

ISGSSVQWSDPLPECEIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTV

NNDEGEWSGPPPECRGGHHHHHH

(SEQ ID NO: 29)

MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILIGSDQTYPEGTQAIYKCRPGYRSLGNII

MVCRKGEWVALNPLRKCQKRPCNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCL

ISGSSVQWSDPLPECEIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTV

NNDEGEWSGPPPECRGGHHHHHH

[0145] In some embodiments, the hybrid protein of the present invention comprises or consists of the amino acid sequence according to any one of SEQ ID NOs: 30-33, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence.

>Signal peptide-human CFH CCP1 or SCR1-human DAF/CD55 CCP3-4 or SCR3-4
(in bold):

(SEQ ID NO: 30)

MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILIGSDQTYPEGTQAIYKCRPGYRSLGNVI

MVCRKGEWVALNPLRKCQKRPCNPGEIRNGQIDVPGGILFGATISFSCNTGYKLEGSTSSFCL

ISGSSVQWSDPLPECEIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTV

NNDEGEWSGPPPECRG

(SEQ ID NO: 31)

MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSDQTYPEGTQAIYKCRPGYRSLGNVI

MVCRKGEWVALNPLRKCQKSCNPGEIRNGQIDVPGGILFGATISFSCNTGYKLEGSTSSFCL

ISGSSVQWSDPLPECEIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTV

NNDEGEWSGPPPECRG

-continued

>Signal peptide-human CFH CCP1 or SCR1-human DAF/CD55 CCP3-4 or SCR3-4
(in bold):

comprising a V62I mutation:

(SEQ ID NO: 32)

MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNII

MVCRKGEWVALNPLRKCQKSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCL

ISGSSVQWSDPLPECEIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTV

NNDEGEWSGPPPECRG

(SEQ ID NO: 33)

MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNII

MVCRKGEWVALNPLRKCQKRCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCL

ISGSSVQWSDPLPECEIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTV

NNDEGEWSGPPPECRG

III. Polynucleotides, Vectors and Hosts

[0146] The present invention provides a nucleic acid encoding any of the above hybrid proteins of the present invention. Also provided is a vector comprising the nucleic acid. In one embodiment, the vector is an expression vector (e.g., a pcDNA vector, such as pcDNA3.1). Also provided is a host cell comprising the nucleic acid or the vector. In one embodiment, the host cell is eukaryotic. In another embodiment, the host cell is selected from a yeast cell or a mammalian cell (e.g., CHO cells or 293 cells, such as Expi293 cells). In another embodiment, the host cell is prokaryotic.

[0147] In one aspect, the present invention provides a nucleic acid encoding any of the above hybrid proteins. The polypeptide encoded by the nucleic acid is capable of having DAF and/or CFH function when expressed from a suitable expression vector, e.g., dissociating C3 and/or C5 convertase, accelerating decay-accelerating activity against the classical pathway C3 convertase and/or the alternative pathway C3 convertase, dissociating the alternative pathway C3 convertase C3bBb, or binding to C3b, and complement-inhibiting activity, including activity inhibiting the Alternative Pathway (AP), the Classical Pathway (CP), and/or C3b deposition.

[0148] For ease of production and purification, the hybrid protein may be fused at the N-terminus to a secretory signal peptide, and/or a tag peptide, such as a hexahistidine tag or biotin tag, to facilitate the purification.

[0149] As will be appreciated by those skilled in the art, each hybrid protein may be encoded by a variety of nucleic acid sequences due to codon degeneracy.

[0150] In some embodiments, the nucleic acid of the present invention comprises a nucleic acid encoding the amino acid sequence selected from any one of SEQ ID NOs: 18-33, or an amino acid sequence having at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence selected from any one of SEQ ID NOs: 18-33.

[0151] Nucleic acid sequences encoding the molecules of the present invention may be generated using methods well known in the art, for example by de novo solid-phase DNA synthesis, or by PCR amplification.

[0152] In one embodiment, one or more vectors comprising a nucleic acid of the present invention are provided. In one embodiment, the vector is an expression vector, such as

a prokaryotic expression vector or a eukaryotic expression vector. Vectors include, but are not limited to, viruses, plasmids, cosmids, lambda phages, or Yeast Artificial Chromosomes (YACs). In a preferred embodiment, the expression vector is a pcDNA, such as pcDNA3.1.

[0153] In one embodiment, a host cell comprising one or more polynucleotides of the present invention is provided. In some embodiments, host cells comprising the expression vectors of the present invention are provided. As used herein, the term “host cell” refers to any kind of cellular system that can be engineered to produce a protein of the present invention. Host cells suitable for replicating and supporting the expression of the proteins of the present invention are well-known in the art. Such cells can be transfected or transduced with a particular expression vector as desired, and large numbers of cells comprising the vector can be cultivated for seeding large-scale fermenters to obtain sufficient quantities of the protein of the present invention for clinical use.

[0154] Suitable host cells include prokaryotic microorganisms such as *E. coli*, eukaryotic microorganisms such as filamentous fungi or yeast, or various eukaryotic cells such as Chinese hamster ovary Cells (CHO), insect cells, and the like. Mammalian cell lines suitable for suspension culture may be used. Examples of useful mammalian host cell lines include SV40 transformed monkey kidney CV1 line (COS-7); human embryonic kidney lines (HEK 293 or 293F cells or 293 FT cells), baby hamster kidney cells (BHK), monkey kidney cells (CV1), Vero cells (VERO-76), human cervical cancer cells (HELA), canine kidney cells (MDCK), Buffalo rat liver cells (BRL 3A), human lung cells (W138), human liver cells (Hep G2), CHO cells, NSO cells, myeloma cell lines such as YO, NSO, P3X63, and Sp2/0, and the like. Mammalian host cell lines suitable for the production of proteins are known in the art. In a preferred embodiment, the host cell is a CHO or HEK293 cell or an Expi293 cell.

IV. Production and Purification of Molecules of the Invention

[0155] In yet another aspect, the present invention provides a method for producing a hybrid protein of the present invention, the method comprising: culturing a host cell comprising a nucleic acid molecule encoding the protein under conditions suitable for expression of the hybrid pro-

tein, optionally further comprising isolating the protein from the host cell or the host cell culture medium, and/or purifying the protein.

[0156] For recombinant production, a polynucleotide encoding a protein of the present invention may be inserted into a vector for further cloning and/or expression in a host cell. Methods well known to those skilled in the art can be used to construct expression vectors. Once an expression vector comprising a polynucleotide of the present invention has been prepared for expression, the expression vector can be transfected or introduced into a suitable host cell. Various techniques can be used to achieve this purpose, for example, protoplast fusion, calcium phosphate precipitation, electroporation, retroviral transduction, viral transfection, biolistics, liposome-based transfection, or other conventional techniques.

[0157] Proteins prepared as described herein can be purified by known prior art techniques such as nickel column, high-performance liquid chromatography, affinity chromatography, ion exchange chromatography, gel electrophoresis, and the like. The actual conditions used to purify a particular protein also depend on factors such as net charge, hydrophobicity, hydrophilicity, and the like, and will be apparent to those skilled in the art.

[0158] The purity and quantification of the proteins of the present invention can be determined by any of a variety of well-known analytical methods.

V. Assay Method

[0159] The hybrid proteins provided herein can be identified, screened, or characterized for their physical/chemical properties and/or biological activity by a variety of assays known in the art.

[0160] The hemolytic inhibitory effect of the hybrid protein of the present invention can be measured by methods known in the art, such as in vitro assays and/or in vivo animal experiments. For example, a hemolytic assay, such as the method described in Examples 4.1 and 4.2, can be used to test for hemolytic inhibitory effects of the molecule on the classical and/or alternative pathways of complement immunity.

[0161] The inhibitory activity of the hybrid protein of the present invention on the complement C3b deposition can be measured by methods known in the art, such as in vitro assays and/or in vivo animal assays. For example, a hemolysis assay, such as the method described in Example 4.3, can be used to test the inhibitory effect of the molecule on the C3b deposition on the surface of red blood cells.

VI. Pharmaceutical Compositions, Pharmaceutical Combinations, and Kits

[0162] In one aspect, the present invention provides a composition, e.g., a pharmaceutical composition, medicament, or formulation, comprising a hybrid protein of the present invention or a nucleic acid molecule encoding the same.

[0163] In one embodiment, the composition further comprises a pharmaceutical excipient, such as a pharmaceutical carrier, a pharmaceutical excipient, including buffers, as known in the art.

[0164] As used herein, "pharmaceutical carrier" includes any and all solvents, dispersion media, isotonic and absorption delaying agents, and the like that are physiologically

compatible. For the use of pharmaceutical excipients and their use, see also "Handbook of Pharmaceutical Excipients", 8th Edition, R. C. Rowe, P. J. Seskey and S. C. Owen, Pharmaceutical Press, London, Chicago.

[0165] The compositions or medicaments or formulations of the present invention may be in a variety of forms. These forms include, for example, liquid, semi-solid, and solid dosage forms, such as liquid solutions (e.g., injectable solutions), powders or suspensions, liposomal formulations, and suppositories. The preferred form depends on the intended mode of administration and therapeutic use.

[0166] In addition to the hybrid protein of the present invention or the nucleic acid encoding the same, the compositions or medicaments or formulations of the present invention may also comprise additional therapeutic agents required for the particular indication being treated, preferably those that do not adversely affect the activity of each other. Thus, in one embodiment, a composition or formulation or medicament, e.g., a pharmaceutical composition, comprises a combination of one or more molecules of the present invention, and one or more additional therapeutic agents.

[0167] The compositions or medicaments or formulations of the present invention may be in a variety of forms. These forms include, for example, liquid, semi-solid, and solid dosage forms, such as liquid solutions (e.g., injectable solutions), powders or suspensions, liposomal formulations, and suppositories. The compositions or medicaments or formulations of the present invention are suitable for administration by intravenous bolus injection, intravenous infusion, intraperitoneal, intradermal, intramuscular, subcutaneous, and intranasal routes (e.g., by injection or infusion). The preferred form depends on the intended mode of administration and therapeutic use.

[0168] The present invention also provides a pharmaceutical combination or a pharmaceutical combination product comprising a hybrid protein of the present invention or a nucleic acid encoding the same and one or more other therapeutic agents.

[0169] The present invention also provides a kit of parts comprising the pharmaceutical combination, e.g., the kit of parts comprises the following in the same package:

[0170] a first container comprising a pharmaceutical composition comprising a hybrid protein of the present invention or a nucleic acid encoding the same; and

[0171] optionally, a second container comprising a pharmaceutical composition comprising one or more additional therapeutic agents (in some embodiments, the two or more additional therapeutic agents are in the same container, or in separate containers).

VII. Uses and Methods

[0172] In one aspect of the present invention, provided is a method of preventing or treating a complement system-associated disease or disorder in a subject, comprising administering to the subject an effective amount of the hybrid protein or the nucleic acid encoding the same of the present invention or the composition or medicament or formulation comprising the protein.

[0173] In some embodiments, the complement system-associated disease is due to abnormal activation of the complement system or a dysregulation of the complement system. In some embodiments, the abnormal activation of the complement system or dysregulation of the complement

system is due to, for example, microbial infection or an increase in autoimmune antibodies, or to a decrease, deletion, incapacitation, or interference or blockade of a function of complement regulatory proteins, etc.

[0174] In some embodiments, the disease treatment will benefit from inhibiting the activity of the complement system.

[0175] In other aspects, the present invention provides the use of the molecule of the present invention, or a composition comprising the same, in the manufacture or preparation of a medicament for use as described herein, for example, for the prevention or treatment of a complement system-associated disease or disorder as mentioned herein.

[0176] In some embodiments, the complement system-associated disease or disorder can be a disease requiring hemolysis inhibition, e.g., a disease requiring inhibition of hemolysis of the classical pathway of the complement immunity and/or the alternative pathway of the complement immunity.

[0177] In some embodiments, the complement system-associated disease or disorder is a disease in which inhibition of C3b deposition activity is desired, such as Dense Deposit Disease (DDD).

[0178] In some embodiments, the hybrid protein or the nucleic acid encoding the same of the present invention or the composition or medicament or formulation comprising the proteins or nucleic acids delays the onset of a disorder and/or a symptom associated with the disorder.

[0179] In some embodiments, the hybrid protein or the nucleic acid encoding the same of the present invention or the composition or medicament or formulation comprising the proteins or nucleic acids can also be administered in combination with one or more other therapies, e.g., therapeutic modalities and/or other therapeutic agents, for use as

described herein, e.g., for the prevention and/or treatment of the related diseases or disorders as mentioned herein.

[0180] The routes of administration of the hybrid protein or the nucleic acid encoding the same of the present invention or the composition or medicament or formulation comprising the proteins or nucleic acids depend on known methods, e.g., injection or infusion.

[0181] The invention further encompasses the use of the hybrid protein or the nucleic acid encoding the same of the present invention or the composition or formulation comprising the proteins or nucleic acids in the manufacture of a medicament for use as described herein, e.g., for the prevention and/or treatment of a related disease or disorder as mentioned herein.

EXAMPLES

Example 1. Molecular Construction

[0182] In order to construct the protein expression vector of the hybrid molecule involved in the present invention and the complement regulatory molecule of the prior art, the DNA encoding each protein was subcloned into the pcDNA3.1 expression vector (purchased from Biofeng) with the synthesized gene DNA (GENEWIZ Inc., Suzhou/Sangon Biotech (Shanghai) Co., Ltd.) as a template by using a method of conventional molecular cloning technology. The plasmid was verified by sequencing and then was used for transient protein expression. Each encoded protein was added with a signal peptide MGWSCILFLVATATGVHS to the N-terminus and a GHHHHHH 6xHis tag to the C-terminus to facilitate subsequent secretion expression with mammalian cells and nickel column purification.

[0183] The amino acid sequence of each protein is as follows:

>CR13m (human CR1 CCP1-3 or SCR1-3, comprising N29K, S37Y, G79D and D109N mutations, the sequence is derived from patent US9988611_B2):

SEQ ID NO: 34

QCNAPFWLPPFARPTNLITDEFEPPIGTYLKYECPGYYGRPFSSICLKNSVWTGAKDRCCRKSCR

NPPDPVNGMVHVIKDIQFGSQIKYSCTKGYRLIGSSSATCIISGNTVIWDNETPICDRIPCGLP

PTITNGDFISTNRENPHYGSVVTYRCNPGSGGRKFVLFVGEPSIYCTSNDDQVGIWSPAPQCI

GHHHHHH

>D24 (human DAF/CD55 CCP2-4 or SCR2-4, the sequence can take reference from Hui-fen Zhang et al., THE JOURNAL OF BIOLOGICAL CHEMISTRY, Vol. 276, No. 29, Issue of July 20, pp. 27290-27295, 2001):

SEQ ID NO: 35

SECEVPTRLNSASLKQPYITQNYFPVGTVEYECRPGYRREPSLSPKLICLQNLKWSTAVEFCCK

KSCPNGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCISGSSVQWSDPLPECREIY

CPAPPQIDNGIIQGERDHYGRQSVTYACNKGFTMIGEHSIYCTVNNDEGEWSGPPPECRGGHH

HHHH

>F15 (human CFH CCP1-5 or SCR1-5, comprising a V62I mutation, the sequence can take reference from Masha Fridkis-Hareli et al., Blood. 2011 Oct. 27; 118(17): 4705-4713 and Agustin Tortajada et al., Hum Mol Genet. 2009 Sept. 15; 18(18): 3452-3461):

SEQ ID NO: 36

EDCNELPPRRNTEILTGSWSDDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKQCK

RPCGHPGDTFPFGTFTLIGGNVFEYGVKAVYTCNEGYQLLGEINYRECDIDGWINDIPICEVVKC

LPVTAPENGKIVSSAMEPDREYHFGQAVRFVCNSGYKIEGDEEMHCDDGFWSKEKPKCIVEISC

-continued

KSPDVGINSPISQKIIYKENERFQYKCNMGYFYSERGDVCTESGWRPLPSCEEKSCDNPYIPN

GDYSPLRIKHRTGDEITYQCRNGFYPATRGNTAKCTSTGWIPAPRCTLKPGHHHHHH

>F1D34-1 (hybrid molecule 1 of the present invention, human CFH CCP1 or SCR1-human DAF/CD55 CCP3-4 or SCR3-4): SEQ ID NO: 22, the sequence with a signal peptide connected to the N terminus is SEQ ID NO: 26

>F1D34-2i (hybrid molecule 2 of the present invention, human CFH CCP1 or SCR1-human DAF/CD55 CCP3-4 or SCR3-4, comprising a V62I mutation): SEQ ID NO: 24, the amino acid sequence with a signal peptide sequence connected to the N terminus is SEQ ID NO: 28

>CR15D34 (hybrid molecule 3 of the present invention, human CR1 CCP15 or SCR15-human DAF/CD55 CCP3-4 or SCR3-4):

SEQ ID NO: 37

GHCQAPDHFHLFAKLKTQTNASDFPIGTSCLKYECRPEYYGRPFSSITCLDNLVWSSPKDVCKRKSC

PNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCISGSSVQWSDPLPECREIYCPA

PPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDEGEWSGPPPECRGGHHHHH

H

>FH13 (human CFH CCP1-3 or SCR1-3, comprising a V62I mutation, there are two more amino acids at the C-terminus for connecting His purification tag):

SEQ ID NO: 38

EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGWVALNPLRKCKQK

RPGCHPGDTPFGTFTLIGGNVFEYGVKAVYTCNEGYQLLGEINYRECDIDGWINDIPICEVVKC

LPVTAPENGKIVSSAMEPDREYHFGQAVRFVCNSGYKIEGDEEMHCSDDGFWSEKPKCKVEISG

HHHHHH

>D14 (human DAF/CD55 CCP1-4 or SCR1-4, no leader sequence, there are two more amino acids at the C-terminus for connecting His purification tag):

SEQ ID NO: 39

DCGLPPDVPNAQPALEGRISFPEDTVITYKCEESFVKIPGEKDSVICLKGSQWSDIEEFCNRSC

EVPTRLNSASLKQPYITQNYFPVGTVEYECRPGYRREPSLSPKLICLQNLKWSTAVEFCKKKKS

CPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCISGSSVQWSDPLPECREIYCP

APPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDEGEWSGPPPECRGKSGHH

HHHH

Example 2. Protein Expression

[0184] After amplification of the plasmid with strain DH5 alpha (Yeasen Biotechnology), the plasmid was prepared using the NucleoBond Xtra Midi Plus (MACHEREY-NA-GEL, see the product Manual for details). Expi293 cells (purchased from Thermo fisher) were transfected with the prepared plasmids using PEI (Polyscience) for transient expression of proteins (see Jäger, V., et al., BMC Biotechnol 13, 52, 2013. for the transfection method). The harvested supernatant was purified by affinity chromatography in a nickel ion pre-packed column (GE life sciences, HisTrap HP, see product Manual for details), dialyzed against PBS for fluid changes, the fluid change protein samples were sterilized by filtration through 0.2 µm filters, and protein quantification was performed according to the theoretical extinction coefficient using the NanoDrop (Thermofisher) A280 method.

Example 3. Protein Purity Detection (SDS-PAGE)

[0185] 2.5 µg of protein was each taken and added proportionally to the loading buffer (Sangon Biotech) with reducing agent and without reducing agent, respectively.

The samples with loading buffer containing reducing agent were boiled at 95° C. for 5 minutes to fully denature the proteins. For the samples with loading buffer without reducing agent, no heat treatment was performed. Electrophoresis was performed with a 10% Precast-Gel Tris-Glycine precast gel (Sangon Biotech) in a Mini-PROTEAN® Tetra electrophoresis system (BioRad). After completing the electrophoresis, the gel was stained with 0.1% Coomassie Brilliant Blue staining solution and de-stained with an ethanol-glacial acetic acid solution.

[0186] The SDS-PAGE electrophoresis results showed that except for a slight diffusion of the D14 protein band, the other protein bands were relatively clear, indicating that the purity of the proteins under both reducing and non-reducing conditions was good (as shown in FIG. 2). Additionally, due to differences in protein reduction, denaturation, modification status, and electrophoresis conditions, there was a certain discrepancy between the electrophoretic molecular weight and the theoretical molecular weight.

Example 4. Detection of Complement-Inhibiting Activity

4.1 Detection of Inhibitory Activity on Classical Pathway of Complement (CP)

[0187] In this experiment, the degree of inhibition of hemolysin-mediated hemolysis of sheep red blood cells caused by normal human serum complement activity was measured by testing the protein to be tested at different concentrations, and the inhibitory activity of each protein on complement CP was analyzed and compared.

[0188] The experimental process was briefly described as follows:

[0189] (1) Gradient dilution of proteins: the protein was formulated into a 12000 nM solution with GVB++ buffer (Gelatin Flora Buffer comprising Ca^{2+} and Mg^{2+} , Cat. No.: 25-02080, TIANDZ), followed by 4-fold gradient dilution, with a total of 7 gradients;

[0190] (2) Dilution of serum: The normal human serum NHS (Shanghai Schbio Co., Ltd.) was taken out of the -70°C . refrigerator, followed by naturally thawing at 4°C . and then the serum was diluted to 4% with GVB++ buffer;

[0191] (3) Activation of sheep red blood cells: the sheep red blood cells (Nanjing SenBeiJia Biological Technology Co., Ltd.) were washed with GVB++ buffer until the supernatant was transparent, then resuspended with GVB++ buffer, and hemolysin (BM351Y, Beijing Bersee Technology Co., Ltd.) was added to the red blood cell suspension at a ratio of 1:400, followed by incubation at 4°C . for 15 minutes to activate the sheep red blood cells. The activated sheep red blood cells were washed twice with GVB++ and resuspended in GVB++ buffer;

[0192] (4) Incubation: 25 μL each of the diluted protein and serum were mixed in a 96-well plate, then 50 μL of the red blood cell suspension obtained in step (3) was added, resulting in an initial final concentration of the protein of 3000 nM and a final concentration of the serum of 1%. The mixture was blended thoroughly and placed in a 37°C . constant temperature incubator for 1 hour, with a negative control (comprising only serum) and a positive control (comprising only serum and red blood cells);

[0193] (5) Termination: After the incubation, 100 μL of 20 mM EDTA-GVB buffer (GVBE) was added to each reaction well to terminate the reaction;

[0194] (6) Reading value: The 96-well plate was centrifuged at 3000 rpm for 5 minutes. 100 μL of the supernatant was added into a new 96-well flat-bottom plate. The absorbance value at OD405 nm was measured in a multi-function plate reader, and the data were saved;

[0195] (7) Data processing: the obtained OD405 nm reading value was put into the following formula to calculate the red blood cell hemolysis inhibition rate:

$$\text{Hemolysis inhibition rate (\%)} = \frac{(\text{positive control reading value} - \text{experimental group reading value})}{(\text{positive control reading value} - \text{negative control reading value})} * 100\%$$

[0196] Then the IC50 value of the protein was calculated by plotting a four-parameter fit curve, with the final protein

concentration as the abscissa and the hemolysis inhibition rate as the ordinate model. The results were shown in FIG. 3.

[0197] The experimental test results of CP hemolytic inhibitory activity showed that the hybrid proteins of CFH and DAF FID34-1 and FID34-2i both had strong CP inhibitory activity, which was equivalent to that of CR1 activity-enhancing fragment CR13m and DAF fragment D24, significantly better than that of CFH fragment FH15 and DAF fragment D14, and also far superior to that of CR15D34, a hybrid protein of CR1 and DAF. Only weak CP inhibitory activity was detected for CR15D34, while CFH fragment FH13 had little activity (as shown in FIG. 3), indicating that the hybrid protein obtained through domain hybridization had excellent CP inhibitory activity.

4.2 Detection of Inhibitory Activity on Alternative Pathway of Complement (AP)

[0198] In this experiment, the degree of inhibition of rabbit erythrocyte hemolysis caused by normal human serum complement activity was measured by testing the protein to be tested at different concentrations, and the inhibitory activity of each protein on complement AP was analyzed and compared.

[0199] The experimental process was briefly described as follows:

[0200] (1) Gradient dilution of proteins: The protein was formulated into a 28000 nM solution with GVBMG buffer (Gelatin Flora Buffer comprising Mg^{2+} and EGTA, Cat. No.: 25-02090, TIANDZ), followed by 4-fold gradient dilution, with a total of 7 gradients;

[0201] (2) Dilution of serum: The normal human serum NHS (Shanghai Schbio Co., Ltd.) was taken out of the -70°C . refrigerator, followed by naturally thawing at 4°C . and the serum was diluted to 60% with GVBMG;

[0202] (3) Preparation of rabbit red blood cells: rabbit red blood cells (Nanjing SenBeiJia Biological Technology Co., Ltd.) were washed with GVBMG until the supernatant was transparent, and then resuspended with GVBMG;

[0203] (4) Incubation: 25 μL each of the diluted protein and serum was mixed in a 96-well plate, then 50 μL of the red blood cell suspension obtained in step (3) was added, resulting in an initial final concentration of the protein of 7000 nM and a final concentration of the serum of 15%, the mixture was blended thoroughly and placed in a 37°C . constant temperature incubator for 1 hour, with a negative control (comprising only serum) and a positive control (comprising only serum and red blood cells);

[0204] (5) Termination: After the incubation, 100 μL of 10 mM EDTA-GVB buffer (GVBE) was added to each reaction well to terminate the reaction;

[0205] (6) Reading value: The 96-well plate was centrifuged at 3000 rpm for 5 minutes. 100 μL of the supernatant was added into a new 96-well flat-bottom plate. The absorbance value at OD405 nm was measured in a multi-function plate reader, and the data were saved;

[0206] (7) Data processing: The obtained OD405 nm reading value was put into the following formula to calculate the red blood cell hemolysis inhibition rate:

Hemolysis inhibition rate (%)=(positive control reading value-experimental group reading value)/(positive control reading value-negative control reading value)*100%

[0207] Then the IC50 value of the antibody was calculated by plotting a four-parameter fit curve, with the final protein concentration as the abscissa and the hemolysis inhibition rate as the ordinate. The results were shown in FIG. 4.

[0208] The experimental test results of AP hemolytic inhibitory activity showed that the hybrid proteins of CFH and DAF F1D34-1 and F1D34-2i both had strong AP inhibitory activity, which was better than that of CFH fragments FH15 and FH13, and DAF fragments D24 and D14, and was also superior to that of the hybrid protein of CR1 and DAF CR15D34 and the CR1 activity-enhancing fragment CR13m (as shown in FIG. 4), indicating that the hybrid protein obtained through domain hybridization had more excellent AP inhibitory activity.

4.3 Detection of Inhibitory Activity in Complement C3b Deposition

[0209] In this experiment, the inhibition of C3b deposition on the surface of rabbit red blood cells by different complement regulatory proteins after human complement activation was detected, and the inhibitory activity of each protein on C3b deposition was analyzed and compared.

[0210] The experimental process was briefly described as follows:

[0211] (1) Gradient dilution of proteins: The protein was prepared into a 1400 nM solution with GVBMG buffer (Gelatin Flora Buffer comprising Mg²⁺ and EGTA, Cat. No.: 25-02090, TIANDZ), followed by 3-fold gradient dilution, with a total of 7 gradients;

[0212] (2) Dilution of serum: In order to prevent complement activation from causing hemolysis and rupture of rabbit red blood cells, resulting in the inability to accurately detect the deposition of C3b on the cell surface, the C5-depleted normal human serum was used for the experiment. C5-Dpl NHS (Complement Technology) was taken out of the -70° C. refrigerator, followed by naturally thawing at 4° C. and the serum was diluted to 40% with GVBMG;

[0213] (3) Preparation of rabbit red blood cells: 4% rabbit red blood cells (Nanjing SenBeiJia Biological Technology Co., Ltd.) were washed with GVBMG

until the supernatant was transparent, and then resuspended with GVBMG to a cell density of 1×10⁷ cells/mL;

[0214] (4) Incubation: 25 μL each of the diluted protein and serum were mixed in a 96-well plate, then 50 μL of the red blood cell suspension obtained in step (3) was added, resulting in an initial final concentration of the protein of 350 nM and a final concentration of the serum of 10%, the mixture was blended thoroughly and placed in a 37° C. constant-temperature incubator for 1 hour, with a negative control (comprising only complement-inactivated serum and red blood cells) and a positive control (comprising only normal serum and red blood cells);

[0215] (5) Termination: After the incubation, 100 μL of 10 mM EDTA-GVB buffer (GVBE) was added to each reaction well to terminate the reaction;

[0216] (6) Detection: the incubated cells were washed three times with PBS, fluorescein-labeled anti-human C3b/iC3b antibody was added for staining (APC anti-complement C3b/iC3b Antibody, Biolegend), followed by incubation at 4° C. for 30 minutes and washing again 3 times. The APC fluorescence MFI (mean fluorescence intensity) of each sample was measured with a flow cytometer (Cytoflex, Beckman), and the data were saved;

[0217] (7) Data processing: the deposition rate of C3b on the surface of red blood cells was calculated with the following formula:

Deposition rate (%)=experimental group MFI value/positive control MFI value*100%

[0218] Then the IC50 value of the antibody was calculated by plotting a four-parameter fit curve, with the final protein concentration as the abscissa and the deposition rate as the ordinate. The results were shown in FIG. 5.

[0219] The experimental test results of C3b deposition showed that the hybrid proteins of CFH and DAF F1D34-1 and F1D34-2i both had strong C3b deposition inhibitory activity, which was better than that of the CFH fragments FH15 and FH13 and the DAF fragments D24 and D14, and was also better than that of the hybrid protein of CR1 and DAF CR15D34 and the CR1activity-enhancing fragment CR13m (as shown in FIG. 5), indicating that the hybrid protein obtained through domain hybridization had more excellent C3b deposition inhibitory activity.

Sequence listing	
SEQ ID NO: 1	MTVARPSVPAALPLLGLPRLLLLLVLLCLPAVWGDCGLPPDPVNAQPALEGRTSFPEDTVITYKCEESFVKIPGEKDSVICLKGSQWSDIEFCNRSCEVPTRLNSASLKQPYITQNYFPVGTVEYECRPGYRREPSLSKPLTCLQNLKWS TAVEFCKKKSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAP PQIDNGIIQGERDHYGYRQSVTYACNKGFMTIGEHSIYCTVNNDGEWSPPEPCRKSLTSKVPPVTQKPTTVNV PTTEVSPTSQKTTTTPNAQATRSTPVSRTTKHFHETTPNKGSGTTSGTTRLLSGHTCTFLTGLLGLTVMGLL T
SEQ ID NO: 2	ATGACCGTCGCGCGGCGAGCGTGCCCGCGGCGCTGCCCTCCTCGGGGAGCTGCCCGGGCTGCTGCTGCTGGTGC TGTGTGCTGCTGCCGGCCGTGTGGGGTGACTGTGGCCCTCCCCAGATGTACCTAATGCCAGCCAGCTTTGGAAGG CCGTACAAGTTTTCCGAGGATCTGTAATAACGTACAAATGTGAAGAAAGCTTTGTGAAAATTCCTGGCGAGAAG GACTCAGTGATCTGCCTTAAGGGCAGTCAATGGTCAGATATTGAAGAGTTCTGCAATCGTAGCTGCGAGGTGCCAA CAAGGCTAAATTCTGCATCCCTCAACAGCCTTATATCACTCAGAATTATTTCCAGTCGGTACTGTTGTGGAATA TGAGTGCCGTCCAGGTACAGAAGAGAACCCTTCTATCACCAAACTAACTTGCCTTCAGAATTTAAAATGGTCC ACAGCAGTCGAATTTGTAAAAAGAAATCATGCCCTAATCCGGGAGAAATACGAAATGGTCAGATTGATGTACCAG GTGGCATATTATTGGTGCAACCATCTCCTTCTCATGTAACACAGGGTACAAATTATTGGCTCGACTTCTAGTTT TTGTCTTATTTTCAGGCAGCTCTGTCCAGTGGAGTGACCCGTTGCCAGAGTGCAGAGAAATTTATGTCCAGCACCA CCACAAATTGACAATGGAATAATCAAGGGGAACGTGACCATATGGAATATAGACAGCTGTAACGTATGATGATGTA ATAAAGGATTACCATGATTGGAGAGCACTCTATTATTGTACTGTGAATAATGATGAAGGAGAGTGGAGTGGCCC

-continued

Sequence listing

ACCACCTGAATGCAGAGGAAATCTCTAATCTCCAGGTCCACCAACAGTTCAGAAACCTACCACAGTAAATGTT
 CCAACTACAGAAGTCTACCAACTTCTCAGAAAAACCAACAAAAACCAACACCAATGCTCAAGCAACACGGA
 GTACACCTGTTTCCAGGACACCAAGCATTTCATGAAACAACCCCAATAAGGAAGTGGAACCACTTCAGGTAC
 TACCGCTCTTCTATCTGGGCACACGTTTTCACGTTGACAGGTTTGCTTGGGACGCTAGTAACCATGGGCTTGCTG
 ACTTAG

SEQ ID MRLAKI ICLMLWAI CVAEDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNP
 NO: 3 LRKCQKRPCGHPGDTPFGTFTLTGNGVFEYGVKAVYTCNEGYQLLGEINYRECDTDGWTNDIPICEVVKCLPVTAP
 ENKIVSSAMEPDREYHFGQAVRFVCSNGYKIEGDEEMHCSDDGFSKEKPKVEISCKSPDVINGSPISQKIIYK
 ENERFQYKCNMGY EY SERGDAVCTESGWRPLPSCEEKSCDNPYIPNGDYSPLRIKHRTGDEITYQCRNGFYPATRG
 NTAKTCTFGWI PAPRCTLKPCTYDPIKHGGLYHENMRRPYFPVAVGKYYSYCDHEHFETPSGSYWDHIHCTQDGWS
 PAVPCLRKCYFPYLENGYNQNYGRKFVQKSIDVACHPGYALPKAQTTVTCMENGSWSPTRCIRVKTCSSSIDIE
 NGFISESQYTYALKKAKYQCKLGYVTADGETSGSITCGKDGWSAQPTCIKSCDIPVFMNARTKNDFTWFKLNDTL
 DYECHDGYESNTGSTTGSIVCGYNGWSDLPICYERECELPKIDVHLVPDRKKDQYKVGVLKFSCKPGFTIVGPN
 VQCYHFGLSPLDLPICKEQVQSCGPPPELLNGNVKEKTKEEYGHSEVVEYYCNPRFLMKGPNKIQCVDEWTTLPVC
 IVEESTCGDIPELEHGWALSSPPYYGDSVEFNCSSEFTMIGHRSITCIHGVWTQLPQCVAIDKLKKCKSSNLII
 LBEHLKNKEFDHNNINIRYCRGKEGWIHTVCINGRWDPEVNCMAQIQCPPPQIPNSHNMTTTLNYRDEKVS
 VLCQENYLIQEGEEITCKDGRWQSIPLCVEKIPCSQPPQIEHGTINSRSSHQESYAHGKLSYTCGGFRIS EENE
 TTCYMGKWSPPQCEGLPCKSPPEISHGVVAHMSDSYQYGEVITYKCFEGPIDGPAIAKCLGKWSHPPSCKIDT
 CLGTLSPFENAI PMGEKKDVYKAGEQVYTCATYKMDGASNVTCINSRWTGRPTCRDTS CVNPPTVQNAIYVSRQM
 SKYPSGERVRYQCRSPYEMFGDEEVMCLNGNWTEPPQCKDSTGKCGPPPIDNGDITSFPLSVYAPASSVEYQCQN
 LYQLEGNKRITCRNGQWSEPPKCLHPCVISREIMENYNI ALRWTAQKLYSRGTGESVEFVCKRGYRLSSRSHTLRT
 TCWDGKLEYPCTAKR

SEQ ID ATGAGACTTCTAGCAAGATTATTTGCTTATGTTATGGGCTATTTGTGTAGCAGAAGATTGCAATGAACTTCCTC
 NO: 4 CAAGAAGAAATCAGAAATTTCTGACAGGTTCTGGTCTGACCAACATATCCAGAAGCACCAGGCTATCTATAA
 ATGCCGCCCTGGATATAGATCTCTTGAAATGTAATAATGGTATGCAGGAAGGGAGAATGGGTGCTCTTAATCCA
 TTAAGGAATATGCAGAAAGGCCCTGTGGACATCTGGAGATACCTCTTTGGTACTTTTACCCTTACAGGAGGAA
 ATGTGTTTGAATATGGTGTAAGCTGTGTATACATGTAATGAGGGGTATCAATGTCTAGGTAGAGATTAAATACCG
 TGAATGTGACACAGATGGATGGACCAATGATATTCCTATATGTGAAGTTGTGAAGTGTTTACCAGTGACAGCACCA
 GAGAAATGAAAAATTTGTCAGTAGTGCAATGGAACAGATCGGGAATACCATTTTGAGCAAGCAGTACGTTTGTAT
 GTAACTCAGGCTACAGAGATTGAAGAGATGAAGAAATGCATTGTTCAGACGATGGTTTTTGGAGTAAGAGAAAAAC
 AAAGTGTGTGGAATTTTCATGCAAAATCCCCAGATGTTATAAATGGATCTCCTATATCTCAGAAGATTATTTATAAG
 GAGAATGAACGATTTCAATATAAATGTAAACATGGGTTATGAATACAGTGAAAGAGGAGATGCTGTATGCACTGAAT
 CTGGATGGCGTCCGTGCTTCAATGTGAAGAAAAATCATGTGATAATCCTTATATTTCCAAATGGTGACTACTCAC
 TTTAAGGATTAACACAGAACTGGAGATGAATCAGTACCAGTGTAGAAATGGTTTTTATCCTGCAACCCGGGGA
 AATACAGCAAAATGCACAACTACTGGCTGGATACCTGCTCCGAGATGTACCTTGAACCTTTGTGATTATCCAGACA
 TTAACATCAGGTTAGGTTATATCATGAGAATATGCGTAGACCATACTTTCCAGTAGCTGTAGGAAAAATATTACTCCTA
 TTACTGTGATGAACATTTTGAGACTCCGTGAGGAAGTTACTGGGATCACATTTCATGCACACAAGATGGATGGTTCG
 CCAGCAGTACCATGCTCAGAAAAATGTTATTTCTTATTTGGAAAAATGGATATAATCAAAATCATGGAAGAAAGT
 TTGTACAGGTTAAATCTATAGACGTTGCTGCAATCCTGGCTACGCTCTTCCAAAGCGCAGACCCACAGTTACATG
 TATGGAGAAATGGCTGGTCTCTACTCCAGATGCATCCGTGTCAAAACATGTTCCAAATCAAGTATAGATATTGAG
 AATGGGTTTATTTCTGAATCTCAGTATACATATGCTTAAAGAAAAAGCGAAATATCAATGCAAACTAGGATATG
 TAACAGATCAGGTTAGGTAACATCAGGATCAATTACATGTGGGAAGATGGATGGTCAGCTCAACCCACGTGCATTAA
 ATCTTGTGATATCCAGTATTTATGAATGCCAGAATAAAAAATGACTTCACATGGTTTAAAGCTGAATGACACATTG
 GACTATGAATGCCATGATGGTTATGAAAGCAATCTGGAAGCACCAGTGGTTCCATAGTGTGTGGTTACAATGGTT
 GGTCGTGATTTACCCATATGTTATGAAAGAGAAATGCGAACTTCTTAAATAGATGTACATTAGTTCTGTATCGCAA
 GAAAGACCAAGTATAAAGTTGGAGAGGTGTTGAAATCTCCTGCAAAACAGGATTTACAATAGTTGGACCTTAATCC
 GTTCAGTGCTACCACTTTGGATTGTCTCCTGACCTCCCAATATGTAAAGAGCAAGTACAATCATGTGGTCCACCTC
 CTGAACCTCTCAATGGGAATGTTAAGGAAAAAACGAAAGAAATATGGACACAGTGAAGTGGTGGAAATATTATTG
 CAATCCTAGATTTCTAATGAAGGGACCTAATAAAATTCATGTGTTGATGGAGAGTGGACAACTTTACCAGTGTGT
 ATTTGTGGAGGAGAGTACCTGTGGAGATATACCTGAACCTGAACATGGCTGGGCCAGCTTTCTTCCCTCTCTTATT
 ACTATGGAGATTCAGTGGAAATCAATTGCTCAGAATCATTTACAATGATTGGACACAGATCAATTACGTGATTCA
 TGGAGTATGGACCAACTTCCCAGTGTGTGGCAATAGATAAACTTAAGAAGTGAATCATCAAAATTAATTATA
 CTTGAGGAACATTTAAAAAACAGGAAGGAATTCGATCATAATTCTAACATAAGGTACAGATGTAGAGGAAAAAGAG
 GATGGATACACACAGCTTCGCATAAATGGAAGATGGGATCCAGAAGTGAACCTGCTCAATGGCACAAATACAATTATG
 CCCACCTCCACCTCAGATTCCCAATCTCACAAATATGACAACCACACTGAATTATCGGGATGGAGAAAAAGTATCT
 GTTCTTTGCCAAGAAAAATTATCTAATTCAGGAAGGAGAAATAATACATGCAAGATGGAAGATGGCAGTCAATAC
 CACTCTGTGTTGAAAAAATTCATGTTTCAACACACCTCAGATAGAACACGGAACCATTAATTCATCCAGGTCTTC
 ACAAGAAAGTTATGCACATGGGACTAAATGAGTTATACCTGTGAGGGTGGTTTCAGGATATCTGAAGAAAAATGAA
 ACAACATGCTACATGGGAAAAATGGAGTTCTCCACCTCAGTGTGAAGGCCCTTCTTGTAAATCTCCACCTGAGATTT
 CTTAGTGTGTTGAGCTCACATGTCAGAGATTATCAGTATGGAGAAGAAGTTACGTACAAATGTTTGAAGGTTT
 TGAATGATGGGCTGCAATTTGCAAAATGCTTAGGAAAAATGGTCTCACCTCATGATGCAAAAAACAGAT
 TGTCTCAGTTTACCTAGCTTTGAAAAATGCCATACCATGGGAGAGAGAAGGATGTGTATAAGGCGGGTGAGCAAG
 TGACTTACACTTGTGCAACATATTACAAAAATGGATGGAGCCAGTAAATGTAACATGCATTAAATAGCAGATGGACAGG
 AAGGCCAACATGCAGAGACACCTCCTGTGTGAATCCGCCACAGTACAAAACTTATATAGTGTGAGACAGATG
 AGTAAATATCCATCTGGTGAAGAGTACGTTATCAATGTAGGAGCCCTTATGAAATGTTTGGGGATGAAGAAGTGA
 TGTGTTTAAATGGAACCTGGACGGAACCACTCAATGCAAGATTTCTACAGGAAAAATGGGGCCCCCTCCACCTAT
 TGACAAATGGGACATTACTTCTTCCCGTTGTCAGTATATGCTCCAGTTTATCAGTTGAGTACCAATGCCAGAAC
 TTGTATCAACTTGAGGGTAACAAGCGAATAACATGTAGAAATGGCAATGGTCAGAACCAACCAAAATGCTTACATC
 CGTGTGTAATATCCGAGAAATATGGAATAATATAACATAGCATTAAAGGTGACGCAACAGAAAGCTTTATTC
 GAGAACAGGTGAATCAGTTGAATTTGTGTGTAACCGGGATATCGTCTTTCATCAGCTTCTCACATTGCGAACA
 ACATGTTGGGATGGGAACTGGAGTATCCAACCTGTGCAAAAAGATAG

-continued

Sequence listing

SEQ ID NO: 5	MRLAKI ICLMLWAI CVAEDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNP LRKCQKRPCGHPGDTFPGTFTLTGGNVFEYGVKAVYTCNEGYQLLGEINYRECDTDGWTNDIPICEVVKCLPVTAP ENGKIVSSAMEPDREYHPGQAVRFVCSNGYKIEGDEEMHCSDDGFSKEKPKCVEISCKSPDVINGSPISQKIIYK ENERFQYKCNMGYEYSERGDAVCTESGWRPLPSCEEKSCDNPYIPNGDYSPLRIKHRTGDEITYQCRNGFYPATRG NTAKCTSTGWI PAPRCTLKPCDYPDIKHGGLYHENMRRPYFPVAVGKYYSYCDHEHFETPSGSYWDHIHCTQDGWS PAVPCLRKCYFPYLENGYNQNYGRKFVQGSIDVACHPGYALPKAQTTVTTCMENGSWSPTRCIRVSFTL
SEQ ID NO: 6	ATGAGACTTCTAGCAAAGATTATTTGCCTTATGTTATGGGCTATTTGTGTAGCAGAAGATTGCAATGAACCTCCTC CAAGAAGAAATACAGAAATTTGACAGGTTCTGGTCTGACCAAACATATCCAGAAGGCACCCAGGCTATCTATAA ATGCCGCCCTGGATATAGATCTCTTGGAATGTAATAATGGTATGCAGGAAGGGAGAATGGGTGCTCTTAATCCA TTAAGGAAATGTCAGAAAAGGCCCTGTGGACATCCTGGAGATACTCCTTTTGGTACTTTTACCCTTACAGGAGGAA ATGTGTTGAATATGGTGAAAAGCTGTGTATACATGTAATGAGGGGTATCAATTGCTAGGTGAGATTAATTACCG TGAATGTGACACAGATGGATGGACCAATGATATTCCTATATGTGAAGTGTGAAGTGTGTACCAGTGACAGCACCA GAGAATGGAAAAAATTGTCTAGTAGTCAATGGAAACAGATCGGGAATACCATTTTGGACAAGCAGTACGGTTGTAT GTAACCTCAGGCTACAAGATTGAAGGAGATGAAGAAATGCATTGTTTCAGACGATGGTTTTTGGAGTAAGAGAAAACC AAAGTGTGTGGAAATTTTCATGCAAATCCCCAGATGTTATAAATGGATCTCCTATATCTCAGAAGATTATTTATAAG GAGAATGAACGATTTCAATATAAATGTAACATGGGTTATGAATACAGTGAAAGAGGAGATGCTGTATGCACTGAAT CTGGATGGCGTCCGTTGCCTTCATGTGAAGAAAAATCATGTGATAATCCTTATATTTCCAAATGGTGACTACTCACC TTTAAGGATTAACACAGAACTGGAGATGAAATCACGTACCAGTGTAGAAATGGTTTTTATCCTGCAACCCGGGGA AATACAGCAAAATGCACAACTACTGGCTGGATACCTGCTCCGAGATGTACCTTGAAACCTTGTGATTATCCAGACA TTAAACATGGAGGTCTATATCATGAGAATATGCGTAGACCATACTTTCCAGTAGCTGTAGGAAAAATATTACTCCTA TTACTGTGTGAAACATTTTGAAGACTCCGTGAGAAAGTTACTGGGATCACATTATTGACACACAAGATGGATGGTTCG CCAGCAGTACCATGCCTCAGAAAAATGTTATTTTCTTATTTGGAAAAATGGATATAATCAAATCATGGAAGAAAGT TTGTACAGGGTAAATCTATAGACGTTGCCTGCCATCCTGGCTACGCTCTTCCAAAAGCGCAGACCACAGTTACATG TATGGAGAATGGCTGGTCTCTACTCCAGATGCATCCGTGTCAGCTTTACCCCTCTGA
SEQ ID NO: 7	KSCPNGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSVQWSDPLPECRE
SEQ ID NO: 8	CPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSVQWSDPLPECRE
SEQ ID NO: 9	IYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 10	KSCPNGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSVQWSDPLPECREIYCPAPPQIDNGII QGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 11	CPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQ GERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 12	EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKQKRP
SEQ ID NO: 13	EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKQK
SEQ ID NO: 14	EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKQKRP
SEQ ID NO: 15	EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKQK
SEQ ID NO: 16	GHHHHHH
SEQ ID NO: 17	MGWSCIIILFVATATGVHS
SEQ ID NO: 18	EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKQKRPCNPGEIRNG QIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQS VTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 19	EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKQKSKSCPNGEIRNG QIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQS VTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 20	EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKQKSKSCPNGEIRNG QIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQS VTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 21	EDCNELPPRRNTEILTGSWSDQTYPEGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKQKRPCNPGEIRNG QIDVPGGILFGATISFSCNTGYKLFGSTSSFLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQS VTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG

-continued

Sequence listing	
SEQ ID NO: 22	EDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKCQKRPCPNPGEIRNG QIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQS VTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 23	EDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNVIMVCRKGEWVALNPLRKCQKSCPNPGEIRNG QIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQS VTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 24	EDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKCQKSCPNPGEIRNG QIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQS VTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 25	EDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKCQKRPCPNPGEIRNG QIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQS VTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 26	MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNVIMVCRKGEWVALN PLRKCQKRPCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPP QIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 27	MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNVIMVCRKGEWVALN PLRKCQKSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPP QIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 28	MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNIIMVCRKGEWVALN PLRKCQKSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPP QIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 29	MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNIIMVCRKGEWVALN PLRKCQKRPCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPP QIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 30	MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNVIMVCRKGEWVALN PLRKCQKRPCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPP QIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 31	MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNVIMVCRKGEWVALN PLRKCQKSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPP QIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 32	MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNIIMVCRKGEWVALN PLRKCQKSCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPP QIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 33	MGWSCIIILFLVATATGVHSEDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNIIMVCRKGEWVALN PLRKCQKRPCPNPGEIRNGQIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPP QIDNGIIQGERDHYGYRQSVTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRG
SEQ ID NO: 34	QCNAPEWLPFARPTNLTDEFEPPIGTLYKYECPGYGRPFSIIICLKNVVTGAKDRCKRCKSRNPPDPVNGMVHV IKDIQFGSQIKYSC TKGYRLIGSSSATCIIISGNTVIWDNETPICDRIPCGLPPTITNGDFISTNRENFHYGSVVTY RCNPGSGGRKFVLFVGEPSIYCTSNDDQVGIWSGPAPQCIGHHHHHH
SEQ ID NO: 35	SCEVPTRLNSASLQPYITQNYFPVGTVEYECRPGYRREPSLSPKLTCLQNLKWS TAVEFCCKKSCPNPGEIRNG QIDVPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQS VTYACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 36	EDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKCQKRPCGHPGDTFPG TFTLTGGNVFEYGVKAVYTCNEGYQLLGEINYRECDTDGWTNDIPICEVVKCLPVTAPENGKIVSSAMEPDREYHF GQAVRFVCNSGYKIEGDEEMHCSDDGFWSEKPKCIVEISCKSPDVINGSPISQKIIYKENERFYKCNMGYEYSER GDAVCTESGWRPLPSCEEKSCDNPIYPNGDYSPLRIKHRTGDEITYQCRNGFYPATRGNTAKCTSTGWIIPAPRCTL KPGHHHHH
SEQ ID NO: 37	GHCQAPDHF LFAKLKTQTNASDFPIGTS LKYECPREYYGRPFSITCLDNLVWSSPKDVCKRKSCPNPGEIRNGQID VPGGILFGATISFSCNTGYKLFGSTSSFCLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQSVTY ACNKGFTMIGEHSIYCTVNNDGEWSGPPPECRGGHHHHH
SEQ ID NO: 38	EDCNELPPRRNTEILTGSWSDQTYPEPGTQAIYKCRPGYRSLGNIIMVCRKGEWVALNPLRKCQKRPCGHPGDTFPG TFTLTGGNVFEYGVKAVYTCNEGYQLLGEINYRECDTDGWTNDIPICEVVKCLPVTAPENGKIVSSAMEPDREYHF GQAVRFVCNSGYKIEGDEEMHCSDDGFWSEKPKCIVEISGHHHHH

-continued

Sequence listing

SEQ ID NO: 39 DCGLPPDVPAQPALEGRSTSPFEDTVITYKCEESFVKIPGEKDSVCLKGSQWSDIEEFNCNRSCEVPTRLNSASLK
 QPYITQNYFPVGTVVEYECRPGYRREPSLSPKLTCLQNLKWTAVEFCKKSCPNPGEIRNGQIDVPGGILFGATI
 SFSCNTGKYLFGSTSSFLISGSSVQWSDPLPECREIYCPAPPQIDNGIIQGERDHYGYRQSVTYACNKGFTMIGE
 HSIYCTVNNDEGEWSGPPPECRGKSGHHHHHH

SEQUENCE LISTING

Sequence total quantity: 39

SEQ ID NO: 1 moltype = AA length = 381

FEATURE Location/Qualifiers

source 1..381

mol_type = protein

organism = synthetic construct

SEQUENCE: 1

MTVARPSVPA ALPLLGLLPR LLLLVLLCLP AVWGDCGLPP DVPNAQPALE GRTSFPEDTV 60
 IYKCEESFV KIPGEKDSV CLKGSQWSDI EEFNCNRSCEV PTRLNSASLK QPYITQNYFP 120
 VGTVVEYECR PGYRREPSLS PKLTCLQNLK WTAVEFCKK KSCPNPGEIR NGQIDVPGGI 180
 LFGATISFSC NTGKYLFGST SSFLISGSS VQWSDPLPEC REIYCPAPPQ IDNGIIQGER 240
 DHYGYRQSVT YACNKGFTMI GEHSIYCTVN NDEGEWSGPP PECRGKSLTS KVPPTVQKPT 300
 TVNVPTTEVS PTSQKTTTKT TTPNAQATRS TPVSRTTKHF HETTPNKGSG TTSGTTRLLS 360
 GHTCFTLTGL LGTLVTMGLL T 381

SEQ ID NO: 2 moltype = DNA length = 1146

FEATURE Location/Qualifiers

source 1..1146

mol_type = other DNA

organism = synthetic construct

SEQUENCE: 2

atgaccgtcg cgcgccgag cgtgcccgcg gcgctgcccc tctcgggga gctgcccgcg 60
 ctgctgctgc tgggtgctgt gtgcctgccc gccgtgtggg gtgactgtgg ccttccccca 120
 gatgtacctg atgcccagcc agctttggaa ggccgtacaa gttttcccga ggatactgta 180
 ataacgtaca aatgtgaaga aagctttgtg aaaattcctg gcgagaagga ctccagtgtc 240
 tgccctaaag gcagtcgaat gtcagatatt gaagagttct gcaatcgtag ctgcgagggtg 300
 ccaacaaggc taaattctgc atccctcaaa cagccttata tcaactcaga ttattttcca 360
 gtccgtactg ttgtggaata tgagtcccg ctaggttaca gaagagaacc ttctctatca 420
 ccaaaactaa cttgccttca gaatttaaaa tgggtccacag cagtogaatt ttgtaaaaag 480
 aaatcatgcc ctaatccggg agaatacaga aatggtcaga ttgatgtacc aggtggcata 540
 ttatttggtg caaccatctc cttctcatgt aacacaggg acaaatatt tggtcgact 600
 tctagttttt gtcttatttc aggcagctct gtccagtggg gtgacccgtt gccagagtgc 660
 agagaaattt attgtccagc accaccacaa attgacaatg gaataattca aggggaacct 720
 gaccattatg gatatagaca gctgttaacg tatgcatgta ataaaggatt caccatgatt 780
 ggagagcact ctattttatg tactgtgaat aatgatgaag gagagtggag tggccacca 840
 cctgaatgca gaggaataat tctaacttcc aaggtccac caacagttca gaaacctacc 900
 acagtaaatg ttccaactac agaagtctca caaacttctc agaaaaccac cacaaaaacc 960
 accacaccaa atgctcaagc aacacggagt acacctgttt ccaggacaac caagcatatt 1020
 catgaacaa ccccaataa aggaagtggg accacttcag gtactaccgc tcttctatct 1080
 gggcacacgt gtttcacgtt gacaggtttg cttgggacgc tagtaaccat gggcttgctg 1140
 acttag 1146

SEQ ID NO: 3 moltype = AA length = 1231

FEATURE Location/Qualifiers

source 1..1231

mol_type = protein

organism = synthetic construct

SEQUENCE: 3

MRLLAKIICL MLWAICVAED CNELPPRRNT EILTGSWSDQ TYPEGTQAIY KCRPGYRSLG 60
 NVIMVCRKGE WVALNPLRKC QKRPCGHPGD TPFGTFTLTG GNVFEYGVKA VYTCNEGYQL 120
 LGIENYRECD TDGWTNDIPI CEVVKCLPVT APENKIVSS AMEPDREYHF GQAVRFVCNS 180
 GYKIEGDEEM HCSDDGFWSK EKPKCVEISC KSPDVINGSP ISQKIIYKEN ERFQYKCNMG 240
 YEYSEGRDAV CTESGWRPLP SCSEKSCDNP YIPNGDYSPL RIKHRTGDEI TYQCRNGFYF 300
 ATRGNTAKCT STGWIPAPRC TLKPCDYPDI KHGGLYHENM RRPYFPVAVG KYYSYVCDEH 360
 FETPSGSYWD HIHCTQDWS PAVPCLRKCY FPYLENGYNQ NYGRKFVQKG SIDVACHPGY 420
 ALPKAQTTVT CMENGWSPTP RCIRVKTCCK SSIDIENGFI SESQYTYALK EKAKYQCKLG 480
 YVTADGETSG SITCGKDGWS AQPTCIKSCD IPVFMNARTK NDFTWFKLND TLDYECHEG 540
 ESNTGSTTGS IVCYNGWSD LPICYERECE LPKIDVHLVP DRKKDQYKVG EVLKFSCPKG 600
 FTIVGPNVQ CYHFGLSFDL PICKEQVQSC GPPPELLNGN VKEKTKEEY HSEVVEYYCN 660
 PRFLMKGNPK IQCVDGEWTT LPVCIVEEST CGDIPELEHG WAQLSSPPY YGDSVEFNCS 720
 ESFTMIGHRS VTCIHGVWTQ LPQCVAIDKL KKCKSSNLII LEEHLKKNKE FDHNSNIRYR 780
 CRKKEGWHT VCINGRWDP NCNSMAQIQ LCPPPPQIPNS HNMTTTLNRY DGEKVSVLQ 840
 ENYLIQEGEE ITCKDGRWQS IPLCCEKIPC SQPPQIEHGT INSSRSSQES YAHGTKLSYT 900

-continued

CEGGFRISEE	NETTCYMGKW	SSPPQCEGLP	CKSPPEISHG	VVAHMSDSYQ	YGEVITYKCF	960
EGFGIDGPAI	AKCLGEKWSH	PPSCIKTDCL	SLPSFENAIP	MGEKKDVYKA	GEQVTTYTCAT	1020
YYKMDGASNV	TCINSRWTR	PTCRDTSCVN	PPTVQNAIYV	SRQMSKYPSG	ERVRYQCRSP	1080
YEMFGDEEVM	CLNGNWTEPP	QCKDSTGKCG	PPPIDNGDI	TSFPLSVYAP	ASSVEYQCQN	1140
LYQLEGNKRI	TCRNQWSEF	PKCLHPCVIS	REIMENYNIA	LRWTAKQKLY	SRTGESVEFV	1200
CKRGYRLSSR	SHTLRTTCWD	GKLEYPTCAK	R			1231

SEQ ID NO: 4 moltype = DNA length = 3696
 FEATURE Location/Qualifiers
 source 1..3696
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 4

atgagacttc	tagcaaaagat	tatttgccct	atgttatggg	ctattttgtgt	agcagaagat	60
tgcaatgaac	tccctccaag	aagaaataca	gaaattctga	caggttcctg	gtctgaccaa	120
acatatccag	aaggcaccga	ggctatctat	aaatgccgcc	ctggatatag	atctcttgga	180
aatgtaataa	tggtagtcag	gaaggagaa	tgggttgctc	ttaatccatt	aaggaaatgt	240
cagaaaaggc	cctgtggaca	tccctggagat	actccttttg	gtacttttac	ccttacagga	300
ggaaatgtgt	tgaatatgg	tgtaaaagct	gtgtatacat	gtaatgaggg	gtatcaattg	360
ctaggtgaga	ttaattaccg	tgaatgtgac	acagatggat	ggaccaatga	tattcctata	420
tgtgaagtgt	tgaagtgttt	accagtgaca	gcaccagaga	atggaaaaat	tgtcagtagt	480
gcaatggaac	cagatcggga	ataccatttt	ggacaagcag	tacggtttgt	atgtaactca	540
ggctacaaga	tgaaggaga	tgaagaaatg	cattgttcag	acgatggttt	ttggagttaa	600
gagaaaccaa	agtgtgtgga	aatttcatgc	aaatccccag	atgttataaa	tggatctcct	660
atatctcaga	agattattta	taaggagaat	gaacgatttc	aataataatg	taacatgggt	720
tatgaataca	gtgaaagagg	agatgctgta	tgcactgaat	ctggatggcg	tccgttgcc	780
tcatgtgaag	aaaaatcatg	tgataatcct	tatttccaa	atggtgacta	ctcaccttta	840
aggattaaac	acagaactgg	agatgaaatc	acgtaccagt	gtagaaatgg	ttttatcct	900
gcaacccggg	gaaatacagc	aaaatgcaca	agtaactggc	ggatacctgc	tccgagatgt	960
accttgaaac	cttgtgatta	tccagacatt	aaacatggag	gtctatatca	tgagaatag	1020
cgtagaccat	actttccagt	agctgtagga	aaatattact	cctattactg	tgatgaacat	1080
tttgagactc	cgtcaggaag	ttactgggat	cacattcatt	gcacacaaga	tggatggctg	1140
ccagcagtac	catgctcag	aaaatgttat	ttccttatt	tggaaaaatg	atataatcaa	1200
aatcatggaa	gaaagtgtgt	acagggtaaa	tctatagacg	ttgcttgcca	tcttggttac	1260
gctcttccaa	aagcgcagac	cacagttaca	tgtatggaga	atggctggct	tccactccc	1320
agatgcattc	gtgtcaaaac	atgttccaaa	tcaagtatag	atattgagaa	tgggtttatt	1380
tctgaatctc	agtatacaca	tgcttaaaaa	gaaaaagcga	aatatcaatg	caaaactagga	1440
tatgtaacag	cagatgggtga	aacatcagga	tcaattacat	gtgggaaaga	tggatgggtca	1500
gctcaaccca	cgtgcattaa	atcttgtgat	atcccagat	ttatgaatgc	cagaactaaa	1560
aatgacttca	catggtttga	gctgaatgac	acattggact	atgaatgcca	tgatggttat	1620
gaaagcaata	ctggaaagc	cactgggttc	atagtgtgtg	gttacaatgg	ttggctctgat	1680
ttaccatcat	gttatgaaag	agaatgcgaa	cttcctaaaa	tagatgtaca	cttagttcct	1740
gatcgcaaga	aagcagcaga	taaaagtggg	gaggtgttga	aattctcctg	caaaccagga	1800
tttacaatag	ttggacctaa	ttccgttcag	tgctaccact	ttggattgtc	tccctgacct	1860
ccaatatgta	aagagcaagt	acaatcatgt	ggctccacct	ctgaactcct	caatgggaat	1920
gttaaggaaa	aaacgaaaga	agaatatgga	cacagtgaa	tggtggaata	ttattgcaat	1980
cctagatttc	taattgaagg	acctaataaa	attcaatgtg	ttgatggaga	gtggacaact	2040
taccagtggt	gtattgtgga	ggagagtacc	tgtggagata	tacctgaact	tgaacatggc	2100
tgggcccagc	tttcttcccc	tctctattac	tatggagatt	cagtgggaat	caattgctca	2160
gaatcattta	caatgattgg	acacagatca	attacgtgta	ttcatggagt	atggacccaa	2220
cttccccagt	gtgtggcaat	agataaactt	aagaagtgca	aatcatcaaa	tttaattata	2280
cttgaggaa	atttaaaaaa	caagaaggaa	ttcgatcata	attctaact	aaggtagaca	2340
tgtagaggaa	aagaaggatg	gatacacaca	gtctgcataa	atgggaagatg	ggatccagaa	2400
gtgaactgct	caatggcaca	aatacaatta	tgcccacctc	cacctcagat	tcccaattct	2460
cacaatatga	caaccacact	gaattatcgg	gatggagaaa	aagtatctgt	tctttgcaa	2520
gaaaattatc	taattcagga	aggagaagaa	attacatgca	aagatggaag	atggcagtc	2580
ataccactct	gtgttgaaaa	aattccatgt	tcacaaccac	ctcagataga	acacgggaac	2640
attaattcat	ccaggctctc	acaagaaagt	tatgcacatg	ggactaaatt	gagttatact	2700
tgtgagggtg	gtttcaggat	atctgaagaa	aatgaaacaa	catgctacat	gggaaaaatg	2760
agttctccac	ctcagtggtga	aggccttcct	tgtaaatctc	cacctgagat	ttctcatggt	2820
gttgtagctc	acatgtcaga	cagttatcag	tatggagaag	aagttacgta	caaatgtttt	2880
gaagggtttg	gaatttgatg	gctgtcaatt	gcataaatgct	taggagaaaa	atgggtctcac	2940
cctccatcat	gcataaaaac	agattgtctc	agtttaccta	gctttgaaaa	tgccataccc	3000
atgggagaga	agaaggatgt	gtataaggcg	ggtagagca	tgacttacac	ttgtgcaaca	3060
tattacaaaa	tggatggagc	cagtaatgta	acatgcatta	atagcagatg	gacaggaagg	3120
ccaacatgca	gagacacctc	ctgtgtgaat	ccgcccacag	tacaaaatgc	ttatatagtg	3180
togagacaga	tgagtaataa	tccatctggt	gagagagtac	gttatcaatg	taggagccct	3240
tatgaaatgt	ttggggatga	agaagtgtgt	tgtttaaatg	gaaactggac	ggaaccacct	3300
caatgcaaa	attctacagg	aaaatgtggg	ccccctccac	ctattgacaa	tggggacatt	3360
acttcattcc	cgtgtgcagt	atatgtctca	gcttcacag	ttgagtacca	atgccagaac	3420
ttgtatcaac	ttgagggtta	caagcgaata	acatgtagaa	atggacaatg	gtcagaacca	3480
ccaaaatgct	tacatccgtg	tgtaatatcc	cgagaaatta	tggaaaatta	taacatagca	3540
ttaagggtga	cagccaaaca	gaagctttat	tgcagaaacag	gtgaatcagt	tgaatttgtg	3600
tgtaaacggg	gatatcgtct	ttcatcacgt	tctcacacat	tgcgaacaac	atgttgggat	3660
gggaaactgg	agtatccaac	ttgtgcaaaa	agatag			3696

SEQ ID NO: 5	moltype = AA length = 449		
FEATURE	Location/Qualifiers		
source	1..449		
	mol_type = protein		
	organism = synthetic construct		
SEQUENCE: 5			
MRLLAKIICL MLWAICVAED	CNELPPRRNT	EILTGSWSDQ TYPEGTQAIY KCRPGYRSLG 60	
NVIMVCRKGE WVALNPLRKC	QKRPCGHGPD	TPFGTFPLTG GNVFPEYGKVA VYTCNEGYYQL 120	
LGIMINYRECD TDGWNLDIPI	CEVVKCLPVT	APENGKIYSS AMEPDREYHF GQAVRFVCNS 180	
GKIEGDEEM HCSDDGFWSK	EKPKCVEISC	KSPDVINGSP ISQKIIYKEN ERFQYKCNMG 240	
YEYSERGDV CTESGWRPLP	SCEEKSCDNP	YIPNGDYSPL RIKHRTGDEI TYQCRNGFYP 300	
ATRGNTAKCT STGWIPAPRC	TLKPCDYPDI	KHGLLYHENM RRPYFPVAVG KYYSYYCDEH 360	
PETPSGSYWD HIHCTQDGWS	PAVPCLRKCY	FPYLENGYNQ NYGRKFVQGK SIDVACHPGY 420	
ALPKAQTTVT CMENGWSPTP	RCIRVSFTL		449
SEQ ID NO: 6	moltype = DNA length = 1350		
FEATURE	Location/Qualifiers		
source	1..1350		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 6			
atgagacttc tagcaaaagt	tatttgcctt	atgttatggg ctattttgtgt agcagaagaat 60	
tgcaatgaac ttcctccaa	agaagaaataca	gaaattctga caggttctgt gtcctgaccaa 120	
acatatccag aaggcaccca	ggcatcttat	aaatgcgcgc ctggatatag atctcttgga 180	
aatgtaataa tggtagcgac	gaaggggagaa	tgggttgctc ttaattccatt aaggaaatgt 240	
cgaaaaaggc cctgtggaca	tctctggagat	actcttttgg gtacttttac ccttacagga 300	
ggaataatgtgt tgaatatgg	tgtaaaagct	gtctctacat gtaatgaggg gtaacaaatg 360	
ctaggttgaga ttaattaccg	tgaatgtgac	acagatggat ggaccaatga tattcctata 420	
tgtgaagttg tgaagttgtt	accagtgaca	gcaccagaga atggaaaaat tgctcagtagt 480	
cgcaatggaac cagatcggga	ataccatttt	ggacaagcac tgcggtttgt atgtaactca 540	
ggctacaaga ttgaaggaga	tgaagaaatg	cattgttctag acgatgtgtt ttggagtaaa 600	
gagaaaccaa agtgtgtgga	aatttctatg	aaatccccag atgttataaa tggatctcct 660	
atatactcaga agattattta	tgaaggagaat	gaacgatctt aatataaatg taacatgggt 720	
tatgaataca gtgaaagagg	agatgctgta	tgcactgaat ctggatggcg tccgttgctc 780	
tcattgtgaag aaaaaatcat	tgataatcct	tatatctcaa atgggtgacta ctcacacctta 840	
agatattaaac acagaactgg	agatgaaatc	acgtaccagt gtagaataatg tttttatcct 900	
gcaacccggg gaaatacacg	aaaatgcaca	agtactggct ggatacctgc tccgagatgt 960	
accttgaaac cttgtgatta	tccagacatt	aaacatggag gtctatataca tgagaatatg 1020	
cgtagaccat actttccagt	agctgttagga	aaatattact cctattactg tgatgaacat 1080	
tttgagactc cgtcaggag	ttactgggat	cacattcatt gcacacaaga tggatggctg 1140	
ccagcagtag catgcctcat	aaaatgttat	tttccctatt tggaaaaatg atataactca 1200	
aatcatgggaa gaaagtttgt	acagggtata	tctatagacg ttgcctgcca tctcggctac 1260	
gctcttccaa aagcgcagac	cacagttaca	tgtatggaga atggctggct tctactctcc 1320	
agatgcatcc gtgtcagctt	taccctctga		1350
SEQ ID NO: 7	moltype = AA length = 62		
FEATURE	Location/Qualifiers		
source	1..62		
	mol_type = protein		
	organism = synthetic construct		
SEQUENCE: 7			
KSCPNGEIR NGQIDVPGGI	LFGATISFSC	NTGYKLFGST SSFCLISGSS VQWSDPLPEC 60	
RE			62
SEQ ID NO: 8	moltype = AA length = 60		
FEATURE	Location/Qualifiers		
source	1..60		
	mol_type = protein		
	organism = synthetic construct		
SEQUENCE: 8			
CPNPGEIRNG QIDVPGGILF	GATISFS CNT	GKYLFGSTSS FCLISGSSVQ WSDPLPECRE 60	
SEQ ID NO: 9	moltype = AA length = 63		
FEATURE	Location/Qualifiers		
source	1..63		
	mol_type = protein		
	organism = synthetic construct		
SEQUENCE: 9			
IYCPAPPQID NGIIQGERDH	YGYRQSVTYA	CNKGFTMIGE HSIYCTVNND EGEWSGPPPE 60	
CRG			63
SEQ ID NO: 10	moltype = AA length = 125		
FEATURE	Location/Qualifiers		
source	1..125		
	mol_type = protein		
	organism = synthetic construct		

-continued

SEQUENCE: 10
KSCPNPGEIR NGQIDVPGGI LFGATISFSC NTGYKLFGST SSFCLISGSS VQWSDPLPEC 60
REIYCPAPPQ IDNGIIQGER DHYGRQSVT YACNKGFTMI GEHSIYCTVN NDEGEWSGPP 120
PECRG 125

SEQ ID NO: 11 moltype = AA length = 123
FEATURE Location/Qualifiers
source 1..123
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 11
CPNPGEIRNG QIDVPGGILF GATISFSCNT GYKLFGSTSS FCLISGSSVQ WSDPLPECRE 60
IYCPAPPQID NGIIQGERDH YGRQSVTYA CNKGFTMIGE HSIYCTVNND EGEWSGPPPE 120
CRG 123

SEQ ID NO: 12 moltype = AA length = 66
FEATURE Location/Qualifiers
source 1..66
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 12
EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNVIMVCRK GEWVALNPLR 60
KCQKRP 66

SEQ ID NO: 13 moltype = AA length = 64
FEATURE Location/Qualifiers
source 1..64
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 13
EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNVIMVCRK GEWVALNPLR 60
KCQK 64

SEQ ID NO: 14 moltype = AA length = 66
FEATURE Location/Qualifiers
source 1..66
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 14
EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNIIMVCRK GEWVALNPLR 60
KCQKRP 66

SEQ ID NO: 15 moltype = AA length = 64
FEATURE Location/Qualifiers
source 1..64
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 15
EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNIIMVCRK GEWVALNPLR 60
KCQK 64

SEQ ID NO: 16 moltype = AA length = 7
FEATURE Location/Qualifiers
source 1..7
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 16
GHHHHHH 7

SEQ ID NO: 17 moltype = AA length = 19
FEATURE Location/Qualifiers
source 1..19
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 17
MGWSCIIILFL VATATGVHS 19

SEQ ID NO: 18 moltype = AA length = 189
FEATURE Location/Qualifiers
source 1..189
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 18
EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNVIMVCRK GEWVALNPLR 60
KCQKRPCPNP GEIRNGQIDV PGGILFGATI SFSCNTGYKL FGSTSSFCLI SGSSVQWSDP 120
LPECREIYCP APPQIDNGII QGERDHYGYR QSVTYACNKG FTMIGEHSIY CTVNNDGEW 180

-continued

SGPPPECRG 189

SEQ ID NO: 19 moltype = AA length = 189
 FEATURE Location/Qualifiers
 source 1..189
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 19
 EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNVIMVCRK GEWVALNPLR 60
 KCQKKSCLNP GEIRNGQIDV PGGILFGATI SFSCNTGYKL FGSTSSFCL I SGSSVQWSDP 120
 LPECREIYCP APPQIDNGII QGERDHYGYR QSVTYACNKG FTMIGEHSIY CTVNNDEGEW 180
 SGPPPECRG 189

SEQ ID NO: 20 moltype = AA length = 189
 FEATURE Location/Qualifiers
 source 1..189
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 20
 EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNIIMVCRK GEWVALNPLR 60
 KCQKKSCLNP GEIRNGQIDV PGGILFGATI SFSCNTGYKL FGSTSSFCL I SGSSVQWSDP 120
 LPECREIYCP APPQIDNGII QGERDHYGYR QSVTYACNKG FTMIGEHSIY CTVNNDEGEW 180
 SGPPPECRG 189

SEQ ID NO: 21 moltype = AA length = 189
 FEATURE Location/Qualifiers
 source 1..189
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 21
 EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNIIMVCRK GEWVALNPLR 60
 KCQKRPCNP GEIRNGQIDV PGGILFGATI SFSCNTGYKL FGSTSSFCL I SGSSVQWSDP 120
 LPECREIYCP APPQIDNGII QGERDHYGYR QSVTYACNKG FTMIGEHSIY CTVNNDEGEW 180
 SGPPPECRG 189

SEQ ID NO: 22 moltype = AA length = 196
 FEATURE Location/Qualifiers
 source 1..196
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 22
 EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNVIMVCRK GEWVALNPLR 60
 KCQKRPCNP GEIRNGQIDV PGGILFGATI SFSCNTGYKL FGSTSSFCL I SGSSVQWSDP 120
 LPECREIYCP APPQIDNGII QGERDHYGYR QSVTYACNKG FTMIGEHSIY CTVNNDEGEW 180
 SGPPPECRG HHHHHH 196

SEQ ID NO: 23 moltype = AA length = 196
 FEATURE Location/Qualifiers
 source 1..196
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 23
 EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNVIMVCRK GEWVALNPLR 60
 KCQKKSCLNP GEIRNGQIDV PGGILFGATI SFSCNTGYKL FGSTSSFCL I SGSSVQWSDP 120
 LPECREIYCP APPQIDNGII QGERDHYGYR QSVTYACNKG FTMIGEHSIY CTVNNDEGEW 180
 SGPPPECRG HHHHHH 196

SEQ ID NO: 24 moltype = AA length = 196
 FEATURE Location/Qualifiers
 source 1..196
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 24
 EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNIIMVCRK GEWVALNPLR 60
 KCQKKSCLNP GEIRNGQIDV PGGILFGATI SFSCNTGYKL FGSTSSFCL I SGSSVQWSDP 120
 LPECREIYCP APPQIDNGII QGERDHYGYR QSVTYACNKG FTMIGEHSIY CTVNNDEGEW 180
 SGPPPECRG HHHHHH 196

SEQ ID NO: 25 moltype = AA length = 196
 FEATURE Location/Qualifiers
 source 1..196
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 25
 EDCNELPPRR NTEILTGSWS DQTYPEGTQA IYKCRPGYRS LGNIIMVCRK GEWVALNPLR 60
 KCQKRPCNP GEIRNGQIDV PGGILFGATI SFSCNTGYKL FGSTSSFCL I SGSSVQWSDP 120

-continued

LPECREIYCP	APPQIDNGII	QGERDHYGYR	QSVTYACNKG	FTMIGEHSIY	CTVNNDEGEW	180
SGPPECRGG	HHHHH					196

SEQ ID NO: 26	moltype = AA	length = 215
FEATURE	Location/Qualifiers	
source	1..215	
	mol_type = protein	
	organism = synthetic construct	

SEQUENCE: 26						
MGWSCIIILFL	VATATGVHSE	DCNELPPRRN	TEILTGSWSD	QTYPEGTQAI	YKCRPGYRSL	60
GNVIMVCRKG	EWVALNPLRK	CQKRPCPNPG	EIRNGQIDVP	GGILFGATIS	FSCNTGYKLF	120
GSTSSFCLIS	GSSVQWSDPL	PECREIYCPA	PPQIDNGIIQ	GERDHYGYRQ	SVTYACNKGK	180
TMIGEHSIYC	TVNNDEGEWS	GPPPECRGGH	HHHHH			215

SEQ ID NO: 27	moltype = AA	length = 215
FEATURE	Location/Qualifiers	
source	1..215	
	mol_type = protein	
	organism = synthetic construct	

SEQUENCE: 27						
MGWSCIIILFL	VATATGVHSE	DCNELPPRRN	TEILTGSWSD	QTYPEGTQAI	YKCRPGYRSL	60
GNVIMVCRKG	EWVALNPLRK	CQKRPCPNPG	EIRNGQIDVP	GGILFGATIS	FSCNTGYKLF	120
GSTSSFCLIS	GSSVQWSDPL	PECREIYCPA	PPQIDNGIIQ	GERDHYGYRQ	SVTYACNKGK	180
TMIGEHSIYC	TVNNDEGEWS	GPPPECRGGH	HHHHH			215

SEQ ID NO: 28	moltype = AA	length = 215
FEATURE	Location/Qualifiers	
source	1..215	
	mol_type = protein	
	organism = synthetic construct	

SEQUENCE: 28						
MGWSCIIILFL	VATATGVHSE	DCNELPPRRN	TEILTGSWSD	QTYPEGTQAI	YKCRPGYRSL	60
GNVIMVCRKG	EWVALNPLRK	CQKRPCPNPG	EIRNGQIDVP	GGILFGATIS	FSCNTGYKLF	120
GSTSSFCLIS	GSSVQWSDPL	PECREIYCPA	PPQIDNGIIQ	GERDHYGYRQ	SVTYACNKGK	180
TMIGEHSIYC	TVNNDEGEWS	GPPPECRGGH	HHHHH			215

SEQ ID NO: 29	moltype = AA	length = 215
FEATURE	Location/Qualifiers	
source	1..215	
	mol_type = protein	
	organism = synthetic construct	

SEQUENCE: 29						
MGWSCIIILFL	VATATGVHSE	DCNELPPRRN	TEILTGSWSD	QTYPEGTQAI	YKCRPGYRSL	60
GNVIMVCRKG	EWVALNPLRK	CQKRPCPNPG	EIRNGQIDVP	GGILFGATIS	FSCNTGYKLF	120
GSTSSFCLIS	GSSVQWSDPL	PECREIYCPA	PPQIDNGIIQ	GERDHYGYRQ	SVTYACNKGK	180
TMIGEHSIYC	TVNNDEGEWS	GPPPECRGGH	HHHHH			215

SEQ ID NO: 30	moltype = AA	length = 208
FEATURE	Location/Qualifiers	
source	1..208	
	mol_type = protein	
	organism = synthetic construct	

SEQUENCE: 30						
MGWSCIIILFL	VATATGVHSE	DCNELPPRRN	TEILTGSWSD	QTYPEGTQAI	YKCRPGYRSL	60
GNVIMVCRKG	EWVALNPLRK	CQKRPCPNPG	EIRNGQIDVP	GGILFGATIS	FSCNTGYKLF	120
GSTSSFCLIS	GSSVQWSDPL	PECREIYCPA	PPQIDNGIIQ	GERDHYGYRQ	SVTYACNKGK	180
TMIGEHSIYC	TVNNDEGEWS	GPPPECRG				208

SEQ ID NO: 31	moltype = AA	length = 208
FEATURE	Location/Qualifiers	
source	1..208	
	mol_type = protein	
	organism = synthetic construct	

SEQUENCE: 31						
MGWSCIIILFL	VATATGVHSE	DCNELPPRRN	TEILTGSWSD	QTYPEGTQAI	YKCRPGYRSL	60
GNVIMVCRKG	EWVALNPLRK	CQKRPCPNPG	EIRNGQIDVP	GGILFGATIS	FSCNTGYKLF	120
GSTSSFCLIS	GSSVQWSDPL	PECREIYCPA	PPQIDNGIIQ	GERDHYGYRQ	SVTYACNKGK	180
TMIGEHSIYC	TVNNDEGEWS	GPPPECRG				208

SEQ ID NO: 32	moltype = AA	length = 208
FEATURE	Location/Qualifiers	
source	1..208	
	mol_type = protein	
	organism = synthetic construct	

SEQUENCE: 32						
MGWSCIIILFL	VATATGVHSE	DCNELPPRRN	TEILTGSWSD	QTYPEGTQAI	YKCRPGYRSL	60

-continued

GNIIMVCRKG	EWVALNPLRK	CQKKSCPNPG	EIRNGQIDVP	GGILFGATIS	FSCNTGYKLF	120
GSTSSFCLIS	GSSVQWSDPL	PECREIYCPA	PPQIDNGIIQ	GERDHYGYRQ	SVTYACNKGK	180
TMIGEHSIYC	TVNNDEGEWS	GPPPECRG				208

SEQ ID NO: 33	moltype = AA	length = 208
FEATURE	Location/Qualifiers	
source	1..208	
	mol_type =	protein
	organism =	synthetic construct

SEQUENCE: 33						
MGWSCIIIFL	VATATGVHSE	DCNELPPRRN	TEILTGWSWD	QTYPEGTQAI	YKCRPGYRSL	60
GNIIMVCRKG	EWVALNPLRK	CQKRPCPNPG	EIRNGQIDVP	GGILFGATIS	FSCNTGYKLF	120
GSTSSFCLIS	GSSVQWSDPL	PECREIYCPA	PPQIDNGIIQ	GERDHYGYRQ	SVTYACNKGK	180
TMIGEHSIYC	TVNNDEGEWS	GPPPECRG				208

SEQ ID NO: 34	moltype = AA	length = 199
FEATURE	Location/Qualifiers	
source	1..199	
	mol_type =	protein
	organism =	synthetic construct

SEQUENCE: 34						
QCNAPFWLFF	ARPTNLTFDE	EFPIGTLYKY	ECRPGYYGRP	FSIICLKNSV	WTGAKDCRRR	60
KSCRNPDPV	NGMVHVIKDI	QFGSQIKYSC	TKGYRLIGSS	SATCIISGNT	VIWDNETPIC	120
DRIPCGLPPT	ITNGDFISTN	RENFHYGSVV	TYRCNPGSGG	RKVPFELVGE	SIYCTSNDDQ	180
VGIWSGPAPQ	CIGHHHHHH					199

SEQ ID NO: 35	moltype = AA	length = 196
FEATURE	Location/Qualifiers	
source	1..196	
	mol_type =	protein
	organism =	synthetic construct

SEQUENCE: 35						
SCEVPTRLNS	ASLKQPYITQ	NYFPVGTVVE	YECRPGYRRE	PSLSPKLTCL	QNLKWSTAVE	60
FCKKKSCLNP	GEIRNGQIDV	PGGILFGATI	SFSCNTGYKL	FGSTSSFCL	SGSSVQWSDP	120
LPECREIYCP	APPQIDNGII	QGERDHYGYR	QSVTYACNKG	FTMIGEHSIY	CTVNNDEGEW	180
SGPPPECRGG	HHHHHHH					196

SEQ ID NO: 36	moltype = AA	length = 313
FEATURE	Location/Qualifiers	
source	1..313	
	mol_type =	protein
	organism =	synthetic construct

SEQUENCE: 36						
EDCNELPPRR	NTEILTGWS	DQTYPEGTQA	IYKCRPGYRS	LGNIIMVCRK	GEWVALNPLR	60
KCQKRPCGHP	GDTPFGFTL	TGGNVFEYGV	KAVYTCNEGY	QLLGEINYRE	CDTDGWTNDI	120
PICEVVKCLP	VTAPENGKIV	SSAMEPDREY	HFGQAVRFVC	NSGYKIEGDE	EMHCSDDGFW	180
SKEKPKCVEI	SCKSPDVI	SPISQKIIYK	ENERFQYKCN	MGYEYSEGRD	AVCTESGWRP	240
LPSCEEKSCD	NPYIPNGDYS	PLRIKHRTGD	EITYQCRNGF	YPATRGNTAK	CTSTGWIPAP	300
RCTLKPGHHH	HHH					313

SEQ ID NO: 37	moltype = AA	length = 193
FEATURE	Location/Qualifiers	
source	1..193	
	mol_type =	protein
	organism =	synthetic construct

SEQUENCE: 37						
GHCQAPDHL	FAKLKTQNA	SDFPIGTSK	YECRPEYYGR	PFSITCLDNL	VWSSPKDVCK	60
RKSCPNPGEI	RNGQIDVPG	ILFGATISFS	CNTGYKLEGS	TSSFCLISGS	SVQWSDPLPE	120
CREIYCPAPP	QIDNGIIQGE	RDHYGYRQSV	TYACNKGFTM	IGEHSIYCTV	NNDEGEWSGP	180
PPECRGHHH	HHH					193

SEQ ID NO: 38	moltype = AA	length = 198
FEATURE	Location/Qualifiers	
source	1..198	
	mol_type =	protein
	organism =	synthetic construct

SEQUENCE: 38						
EDCNELPPRR	NTEILTGWS	DQTYPEGTQA	IYKCRPGYRS	LGNIIMVCRK	GEWVALNPLR	60
KCQKRPCGHP	GDTPFGFTL	TGGNVFEYGV	KAVYTCNEGY	QLLGEINYRE	CDTDGWTNDI	120
PICEVVKCLP	VTAPENGKIV	SSAMEPDREY	HFGQAVRFVC	NSGYKIEGDE	EMHCSDDGFW	180
SKEKPKCVEI	SGHHHHHH					198

SEQ ID NO: 39	moltype = AA	length = 260
FEATURE	Location/Qualifiers	
source	1..260	
	mol_type =	protein

-continued

```

organism = synthetic construct
SEQUENCE: 39
DCGLPPDVVN AQPALGRTS FPEDTVITYK CEESFVKIPG EKDSVICLKG SQWSDIEEFC 60
NRSCVPTLRL NSASLKQPYI TQNYFPVGTV VEYECRPGYR REPSLSPKLT CLQNLKWSTA 120
VEFCKKKSCP NPGEIRNGQI DVPGGILFGA TISFSCNTGY KLFGSTSSFC LISGSSVQWS 180
DPLPECREIY CPAPPQIDNG IIQGERDHYG YRQSVTYACN KGFTMIGHS IYCTVNNDEG 240
EWSGPPPECR GKSQHSHHHH 260

```

1. A hybrid protein, which comprises or consists of
 - (i) CCP1 of human complement Factor H (CFH),
 - (ii) CCP3 and CCP4 of human decay-accelerating factors (DAFs), and
 optionally a signal peptide and/or a tag.
2. The hybrid protein of claim 1, wherein the CCP3 and CCP4 of DAF are directly linked together to form CCP3-4.
3. The hybrid protein of claim 1 or 2, wherein the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-82 of the human CFH protein, optionally, the CCP1 comprises a V62I mutation, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3.
4. The hybrid protein of claim 1 or 2, wherein the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-84 of the human CFH protein, optionally, the CCP1 comprises a V62I mutation, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3.
5. The hybrid protein of any one of claims 1-4, wherein the CCP3 of human DAF comprises or consists of the amino acid sequence at positions 161-222 of the human DAF protein, and/or the CCP4 of human DAF comprises or consists of the amino acid sequence at positions 223-285 of the human DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.
6. The hybrid protein of any one of claims 1-4, wherein the CCP3 of human DAF comprises or consists of the amino acid sequence 163-222 of the human DAF protein, and/or the CCP4 of human DAF comprises or consists of the amino acid sequence 223-285 of the human DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.
7. The hybrid protein of claim 2, wherein the CCP3-4 of human DAF comprises or consists of the amino acid sequence at positions 161-285 of the human DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.
8. The hybrid protein of any one of claims 2-4, wherein the CCP3-4 of human DAF comprises or consists of the amino acid sequence at position 163-285 of the human DAF protein, wherein the amino acid positions are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.
9. The hybrid protein of claim 2, wherein
 - (1) the CCP1 of human CFH comprises or consists of the amino acid sequences at positions 19-82 of the human CFH protein, and the CCP3 of human DAF comprises or consists of the amino acid sequence 161-222 of the human DAF protein, and the CCP4 of human DAF comprises or consists of the amino acid sequence 223-285 of the human DAF protein;
 - (2) the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-84 of the human CFH protein, and the CCP3 of human DAF comprises or consists of the amino acid sequence 163-222 of the human DAF protein, and the CCP4 of human DAF comprises or consists of the amino acid sequence 223-285 of the human DAF protein;
 - (3) the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-82 of the human CFH protein, and the CCP3-4 of human DAF comprises or consists of the amino acid sequence at positions 161-285 of the human DAF protein; or
 - (4) the CCP1 of human CFH comprises or consists of the amino acid sequence at positions 19-84 of the human CFH protein, and the CCP3-4 of human DAF comprises or consists of the amino acid sequence at positions 163-285 of the human DAF protein;
 wherein the amino acid positions of the human CFH protein are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 3 and the amino acid positions of the human DAF protein are numbered corresponding to the amino acid positions as set forth in SEQ ID NO: 1.
10. The hybrid protein of claim 9, wherein the CCP1 of human CFH has a V62I mutation.
11. The hybrid protein of any one of claims 1-10, wherein the human CFH protein is a native human CFH protein or comprises
 - (i) the amino acid sequence as set forth in SEQ ID NO: 3 or 5;
 - (ii) the amino acid sequence encoded by the nucleotide sequence as set forth in SEQ ID NO: 4 or 6;
 - (iii) the amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in (i) or (ii); or
 consists of the amino acid sequence as set forth in any one of (i) to (iii).
12. The hybrid protein of any one of claims 1-11, wherein the human DAF protein is a native human DAF protein or comprises
 - (i) the amino acid sequence as set forth in SEQ ID NO: 1;
 - (ii) the amino acid sequence encoded by the nucleotide sequence as set forth in SEQ ID NO: 2;
 - (iii) the amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in (i) or (ii); or
 - (iv) consists of the amino acid sequence as set forth in any one of (i) to (iii).
13. The hybrid protein of any one of claims 1-12, wherein the CCP1 of human CFH comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 12, 13, 14, or 15, or an amino acid sequence having at least

- 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 12, 13, 14, or 15;
- the CCP3 of human DAF comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 7 or 8, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 7 or SEQ ID NO: 8;
- the CCP4 of human DAF comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 9, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 9; and/or
- the CCP3-4 of human DAF comprises or consists of the amino acid sequence as set forth in SEQ ID NO: 10 or SEQ ID NO: 11, or an amino acid sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence as set forth in SEQ ID NO: 10 or SEQ ID NO: 11.
- 14.** The hybrid protein of any one of claims **1-13**, wherein the tag is, for example, a purification tag such as a hexahistidine tag or a biotin tag, comprising, e.g., the amino acid sequence as set forth in SEQ ID NO: 12.
- 15.** The hybrid protein of any one of claims **1-14**, wherein the signal peptide is secretory, comprising, e.g., the amino acid sequence as set forth in SEQ ID NO: 17.
- 16.** The hybrid protein of any one of claims **1-15**, wherein the hybrid protein comprises or consists of an amino acid sequence of any one of SEQ ID NOs: 18-33, or an amino acid sequence having at least a 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence.
- 17.** A nucleic acid molecule encoding the hybrid protein of any one of claims **1-16**.
- 18.** An expression vector comprising the nucleic acid molecule of claim **17**, preferably the expression vector is pcDNA3.1.
- 19.** A host cell comprising the nucleic acid molecule of claim **17** or the expression vector of claim **18**, preferably the host cell is prokaryotic or eukaryotic, such as a 293 cell, e.g., an Expi293 cell.
- 20.** A method of preparing the hybrid protein of any one of claims **1-16**, the method comprising culturing the host cell

of claim **19** under conditions suitable for expression of the hybrid protein, and optionally further comprising isolating the protein from the host cell or the host cell culture medium, and/or purifying the protein.

21. A pharmaceutical composition or medicament or formulation comprising the hybrid protein according to any one of claims **1-16** or a nucleic acid molecule encoding the same, or the nucleic acid molecule of claim **17**, and optionally a pharmaceutically acceptable excipient.

22. A pharmaceutical combination product comprising the hybrid protein according to any one of claims **1-16** or a nucleic acid molecule encoding the same, or the nucleic acid molecule of claim **17**; and the additional therapeutic agents.

23. A method of preventing or treating a complement system-associated disease or disorder in a subject, comprising administering to the subject an effective amount of the hybrid protein of any one of claims **1-16** or a nucleic acid molecule encoding the same; or the nucleic acid molecule of claim **17**; or the pharmaceutical composition or formulation of claim **21**; or the pharmaceutical combination product of claim **22**.

24. The method of claim **23**, wherein the disease or disorder is due to an abnormal activation of the complement system or a dysregulation of the complement system.

25. The method of claim **24**, wherein the abnormal activation of the complement system or the dysregulation of the complement system is due to, for example, a microbial infection or an increase in autoimmune antibodies, or a decrease, deletion, incapacitation, or interference or blockade of a function of a complement regulatory protein.

26. The method of claim claim **23**, wherein the complement system-associated disease or disorder is selected from a disease requiring hemolysis inhibition, such as a disease requiring inhibition of hemolysis of the classical pathway of the complement immunity and/or the alternative pathway of the complement immunity; or a disease requiring inhibition of C3b deposition activity, such as dense deposition disease DDD.

* * * * *