



US012383600B2

(12) **United States Patent**
Bishop et al.

(10) **Patent No.:** US 12,383,600 B2
(45) **Date of Patent:** Aug. 12, 2025

(54) **C3B BINDING POLYPEPTIDE**(71) Applicant: **Complement Therapeutics Limited**, London (GB)(72) Inventors: **Paul Bishop**, London (GB); **Simon Clark**, London (GB); **Richard Unwin**, London (GB)(73) Assignee: **Complement Therapeutics Limited**, London (GB)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **18/052,768**(22) Filed: **Nov. 4, 2022**(65) **Prior Publication Data**

US 2023/0310545 A1 Oct. 5, 2023

2008/0221011 A1	9/2008	Gilkeson et al.
2009/0214538 A1	8/2009	Fung et al.
2009/0269338 A1	10/2009	Huang et al.
2009/0324585 A1	12/2009	Robinson et al.
2010/0009393 A1	1/2010	Morgan et al.
2010/0130413 A1	5/2010	Marsh et al.
2010/0240075 A1	9/2010	Granoff et al.
2011/0091450 A1	4/2011	Schwaeble et al.
2011/0190221 A1	8/2011	Francois et al.
2011/0229497 A1	9/2011	Thurman et al.
2012/0107315 A1	5/2012	Behrens et al.
2012/0225056 A1	9/2012	Rother et al.
2012/0258095 A1	10/2012	Demopoulos et al.
2013/0029912 A1	1/2013	Holers et al.
2013/0064820 A1	3/2013	Magro
2013/0078245 A1	3/2013	Holers et al.
2013/0190477 A1	7/2013	Kovacs et al.
2013/0217616 A1	8/2013	Garred et al.
2013/0225795 A1	8/2013	Schmidt et al.
2015/0079084 A1	3/2015	Her et al.
2015/0247143 A1	9/2015	Fitzgerald et al.
2016/0083469 A1	3/2016	Rohrer et al.
2016/0237125 A1	8/2016	Barlow et al.
2016/0311893 A1	10/2016	Patz, Jr. et al.
2016/0326231 A1	11/2016	Hu et al.
2017/0190753 A1	7/2017	Abache
2020/0277360 A1	9/2020	Clark et al.

Related U.S. Application Data

(63) Continuation of application No. 16/962,108, filed as application No. PCT/EP2019/050949 on Jan. 15, 2019, now Pat. No. 11,524,050.

(30) **Foreign Application Priority Data**

Jan. 15, 2018 (GB) 1800620

(51) **Int. Cl.***A61P 27/02* (2006.01)
A61K 38/17 (2006.01)(52) **U.S. Cl.**CPC *A61K 38/1709* (2013.01); *A61P 27/02* (2018.01)(58) **Field of Classification Search**CPC A61K 38/1709; A61K 38/00; A61P 27/02;
C07K 14/47; C07K 14/70596

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,679,546 A	10/1997	Ko et al.
5,981,481 A	11/1999	Fearon et al.
6,458,360 B1	10/2002	Fearon et al.
6,551,595 B1	4/2003	Rosengard et al.
8,664,176 B2	3/2014	Marsh et al.
11,524,050 B2	12/2022	Bishop et al.
2003/0049831 A1	3/2003	Murphy
2003/0086940 A1	5/2003	Costa et al.
2004/0146938 A1	7/2004	Nguyen et al.
2005/0112139 A1	5/2005	Karp
2005/0169921 A1	8/2005	Bell et al.
2005/0222027 A1	10/2005	Chiang et al.
2005/0277158 A1	12/2005	Chen
2006/0140939 A1	6/2006	Fung
2007/0065433 A1	3/2007	Mollnes et al.
2007/0093443 A1	4/2007	Madison et al.
2007/0274989 A1	11/2007	Fung et al.
2008/0075728 A1	3/2008	Newman
2008/0118506 A1	5/2008	An et al.
2008/0193414 A1	8/2008	Proudfoot et al.

FOREIGN PATENT DOCUMENTS

EP	0222611 A2	5/1987
EP	2083841 B1	11/2013
WO	WO-198909220 A1	10/1989
WO	WO-199210205 A1	6/1992
WO	WO-199400571 A1	1/1994

(Continued)

OTHER PUBLICATIONS

U.S. Appl. No. 16/962,108, filed Jul. 14, 2020, Paul Bishop et al.
U.S. Appl. No. 18/048,303, filed Oct. 20, 2022, Paul Bishop et al.
Akkerman et al (Blood, Nov. 12, 2009, vol. 114, No. 20, 4323-4324) (Year: 2009).Clark et al. "Role of Factor H and Related Proteins in Regulating Complement Activation in the Macula and Relevance to Age-Related Macular Degeneration." J Clin Med. 4(1): 18-31 (2015).
Clark et al., Bruch's Membrane Compartmentalizes Complement Regulation in the Eye with Implication for Therapeutic Design in Age-Related Macular Degeneration, Original Research, 2017, vol. 8, pp. 1-10.

Clark et al., Identification of factor H-like protein 1 as the predominant complement regulator in Bruch's membrane: implication for age-related degeneration, J. Immunol., 2014, vol. 193(10), pp. 4962-4970.

(Continued)

Primary Examiner — Sudhakar Katakam

(74) Attorney, Agent, or Firm — Wolf, Greenfield & Sacks, P.C.

(57) **ABSTRACT**

Polypeptides comprising a C3b binding region are disclosed, as well as nucleic acids and vectors encoding such polypeptides, and cells and compositions comprising such polypeptides. Also disclosed are uses and methods using the polypeptides for treating and preventing diseases and conditions.

20 Claims, 17 Drawing Sheets**Specification includes a Sequence Listing.**

(56)

References Cited

FOREIGN PATENT DOCUMENTS

WO	WO-199416719	A1	8/1994	
WO	WO-1995008570	A1	3/1995	
WO	1995/09239	A2	4/1995	
WO	WO-199509239	A3	10/1995	
WO	WO-199944625	A1	9/1999	
WO	WO-200061752	A2	10/2000	
WO	WO-2007149567	A2	12/2007	
WO	2008/048675	A2	4/2008	
WO	WO-2008048675	A3	10/2008	
WO	WO-2008154251	A2	12/2008	
WO	WO-2009056631	A2	5/2009	
WO	WO-2010056399	A1	5/2010	
WO	WO-2011057158	A1	5/2011	
WO	WO-2011163412	A1	12/2011	
WO	WO-2012037370	A1	3/2012	
WO	WO-2012095519	A1	7/2012	
WO	WO-2013082563	A1	6/2013	
WO	2013/151665	A2	10/2013	
WO	WO-2012095519		12/2014	
WO	WO-2015023972	A1	2/2015	
WO	WO-2016022914	A1	2/2016	
WO	WO-2016028150	A1	2/2016	
WO	WO-2017053732	A2 *	3/2017 A61K 47/6911
WO	WO-2017053732	A3	5/2017	
WO	WO-2017194912	A1	11/2017	
WO	WO-2018002131	A1 *	1/2018 A61K 39/3955
WO	WO-2018224663	A1	12/2018	
WO	PCT/EP2019/050949		1/2019	
WO	WO-2019138137	A1	7/2019	

OTHER PUBLICATIONS

- Clark et al., Impaired binding of the age-related macular degeneration-associated complement factor H 402H allotype to Bruch's membrane in human retina, *The Journal of Biological Chemistry*, 2010, vol. 285(39), pp. 30192-30202.
- Fenaille et al., "Site-specific N-glycan characterization of human complement factor H," *Glycobiology* 17(9): 932-944 (2007).
- Foley et al., "Interplay between fibrinolysis and complement: plasmin cleavage of iC3B modulates immune responses," *Journal of Thrombosis and Haemostasis* 13(4): 610-618 (2014).
- Forest et al., Cellular models and therapies for age-related macular degeneration, *Disease Models and Mechanisms*, 2015, vol. 8, pp. 421-427.
- International Search Report and Written Opinion for PCT/EP2018/065199 dated Oct. 19, 2018, 17 pages.

International Search Report and Written Opinion for PCT/EP2019/050949 dated Sep. 10, 2018, 12 pages.

Kirkitadze et al., Independently Melting Modules and Highly Structured Intermodular Junctions within Complement Receptor Type 1, *Biochemistry*, 1999, vol. 38(2), pp. 7019-7031.

Mather et al. "Molecules great and small: the complement system." *Clinical Journal of the American Society of Nephrology* 10(9): 1636-1650 (2015).

Thomas. "Furin at the cutting edge: from protein traffic to embryogenesis and disease." *Nature reviews Molecular cell biology* 3(10): 753-766 (2002).

Tsiftoglou et al. "Human complement factor I does not require cofactors for cleavage of synthetic substrates." *The Journal of Immunology* 173(1): 367-375 (2004).

UK Search Report GB 1800620.5 dated Oct. 17, 2018, 2 pages.

Wang. "Association between complement factor I gene polymorphisms and the risk of age-related macular degeneration: a Meta-analysis of literature." *International Journal of Ophthalmology* 9(2): 298-305 (2016).

Whitmore et al., Complement activation and chroiocapillaris loss in early AMD: Implication for pathophysiology and therapy, 2015, *Prog. Retin Eye Res.*, pp. 1-71.

Wu et al. "Structure of complement fragment C3b-factor H and implications for host protection by complement regulators" *Nature immunology* 10(7): 728-733 (2009).

Zhang et al., Soluble CR1 Therapy Improves Complement Regulation in C3 Glomerulopathy, *J Am Soc Nephrol*, 2013, vol. 24, pp. 1820-1829.

Zipfel et al. "The Role of Complement in AMD." In: Lambris J., Adamis A. (eds) *Inflammation and Retinal Disease: Complement Biology and Pathology*. Advances in Experimental Medicine and Biology, vol. 703. Springer, New York, NY: pp. 9-24 (2010).

Zipfel et al., The Role of Complement in AMD, 2010. *Adv Excp Med Bio.*, vol. 285(39), pp. 30192-30202.

Fernandez-Godino et al., "Changes in extracellular matrix cause RPE cells to make basal deposits and activate the alternative complement pathway," *Hum Mol Genet*. 2018 27(1):147-159.

Rathi et al., "Therapeutic targeting of the complement system in ocular disease," *2023 Drug Discovery Today* 28(11) 1-21.

Hourcade et al., Decay-accelerating factor (DAF), complement receptor 1 (CR1), and factor H dissociate the complement AP C3 convertase (C3bBb) via sites on the type A domain of Bb. *J Biol Chem.* Jan. 11, 2002;277(2):1107-12. doi: 10.1074/jbc.M109322200. Epub Nov. 2, 2001.

Extended European Search Report for Application No. EP 24213690. 1, mailed Mar. 28, 2025.

* cited by examiner

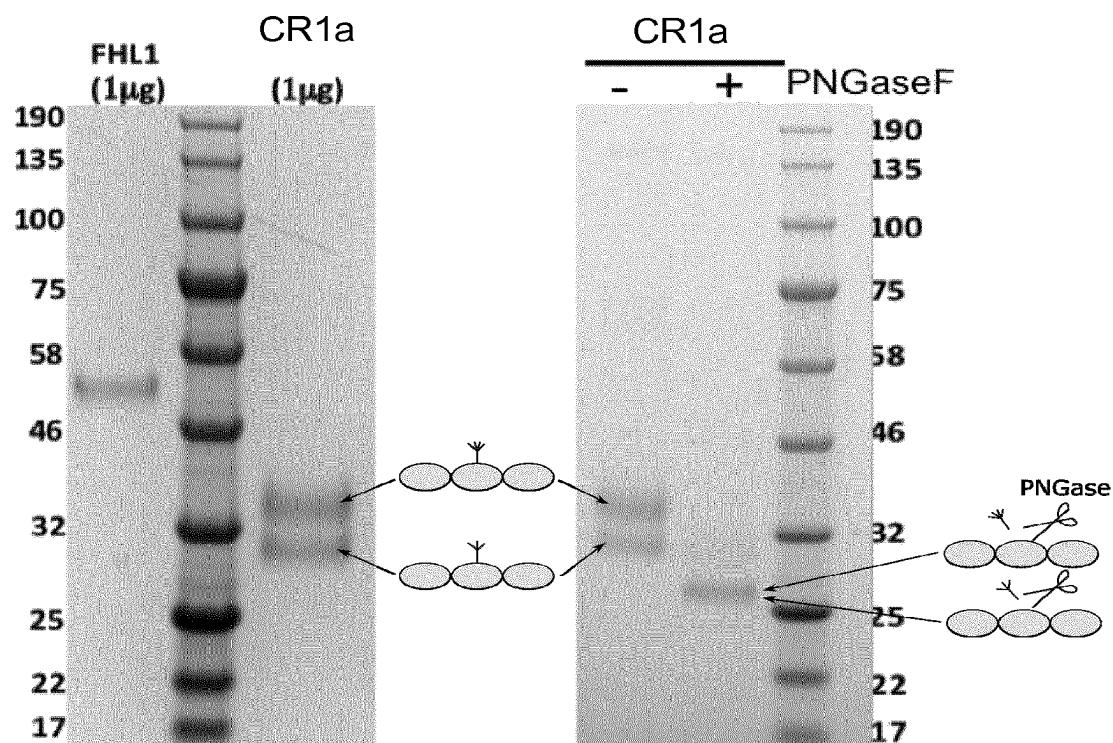


Figure 1A

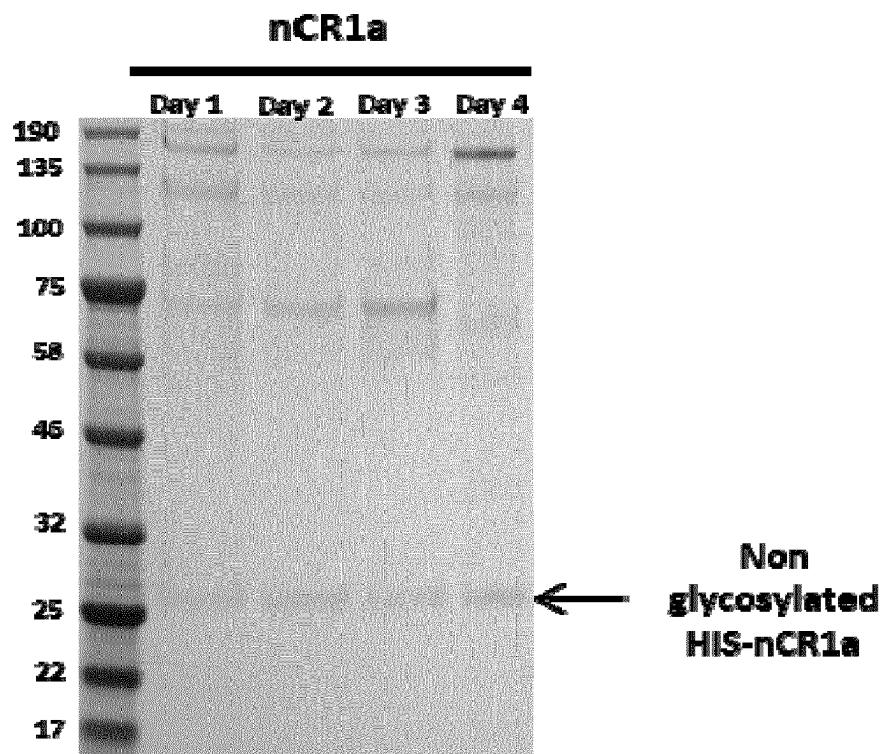
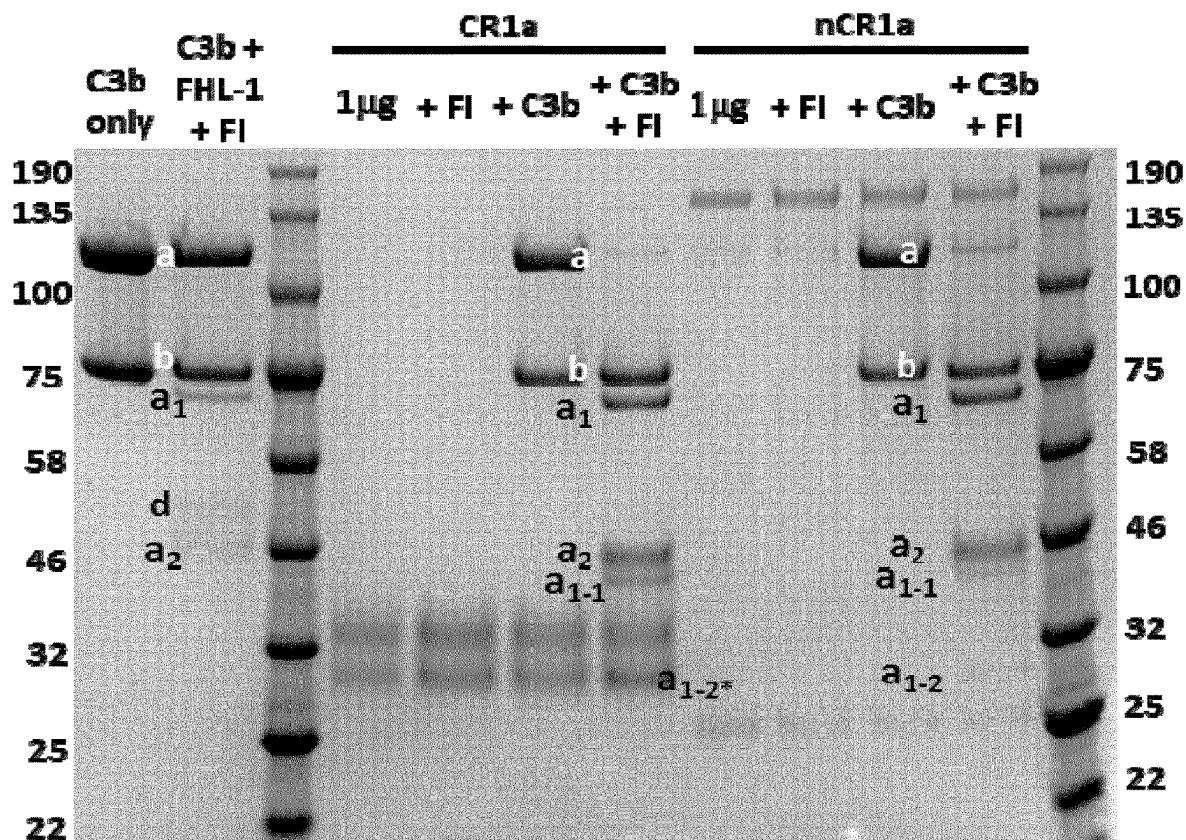


Figure 1B



a - C3b α chain 114 kDa (a = a₁+a₂)

b - C3b β chain 75kDa

a₁ - IC3b 68 kDa (a₁= a₁₋₁+a₁₋₂)

d - FHL-1

a₂ - IC3b 43kDa

a₁₋₁ - C3dg 41kDa

a₁₋₂ - 27kDa (remaining part of 68kDa IC3b fragment)

a_{1-2*} - 27kDa masked by HISCR1a band

Figure 2

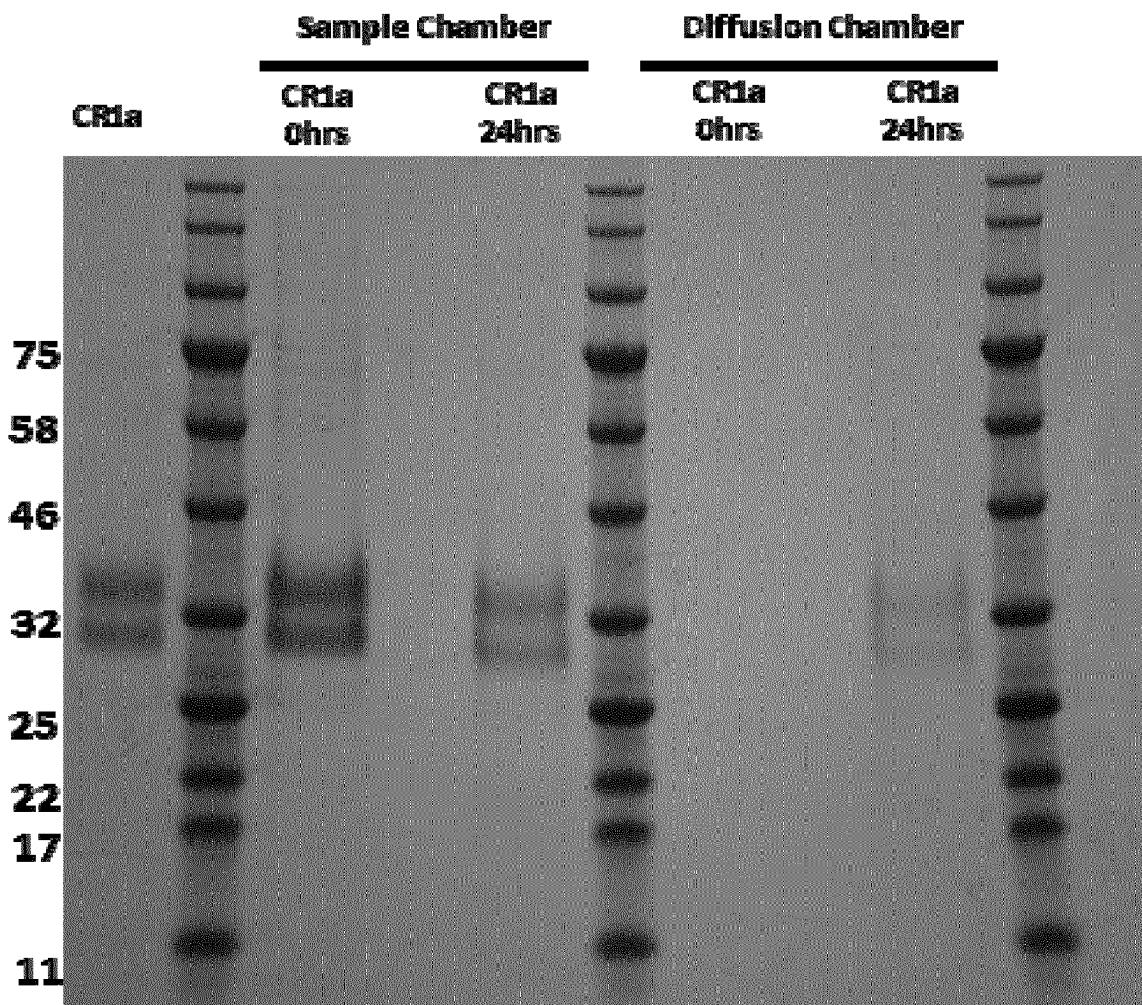


Figure 3A

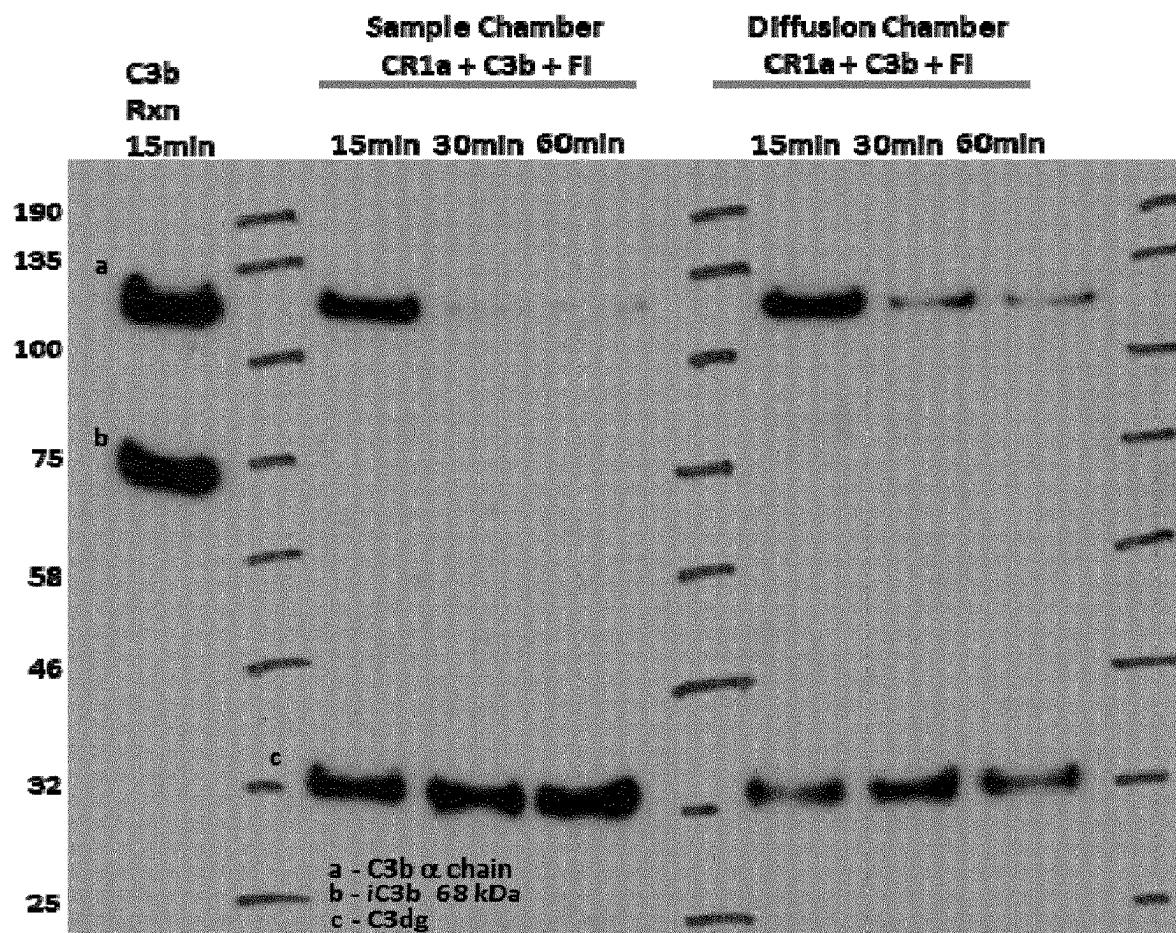


Figure 3B

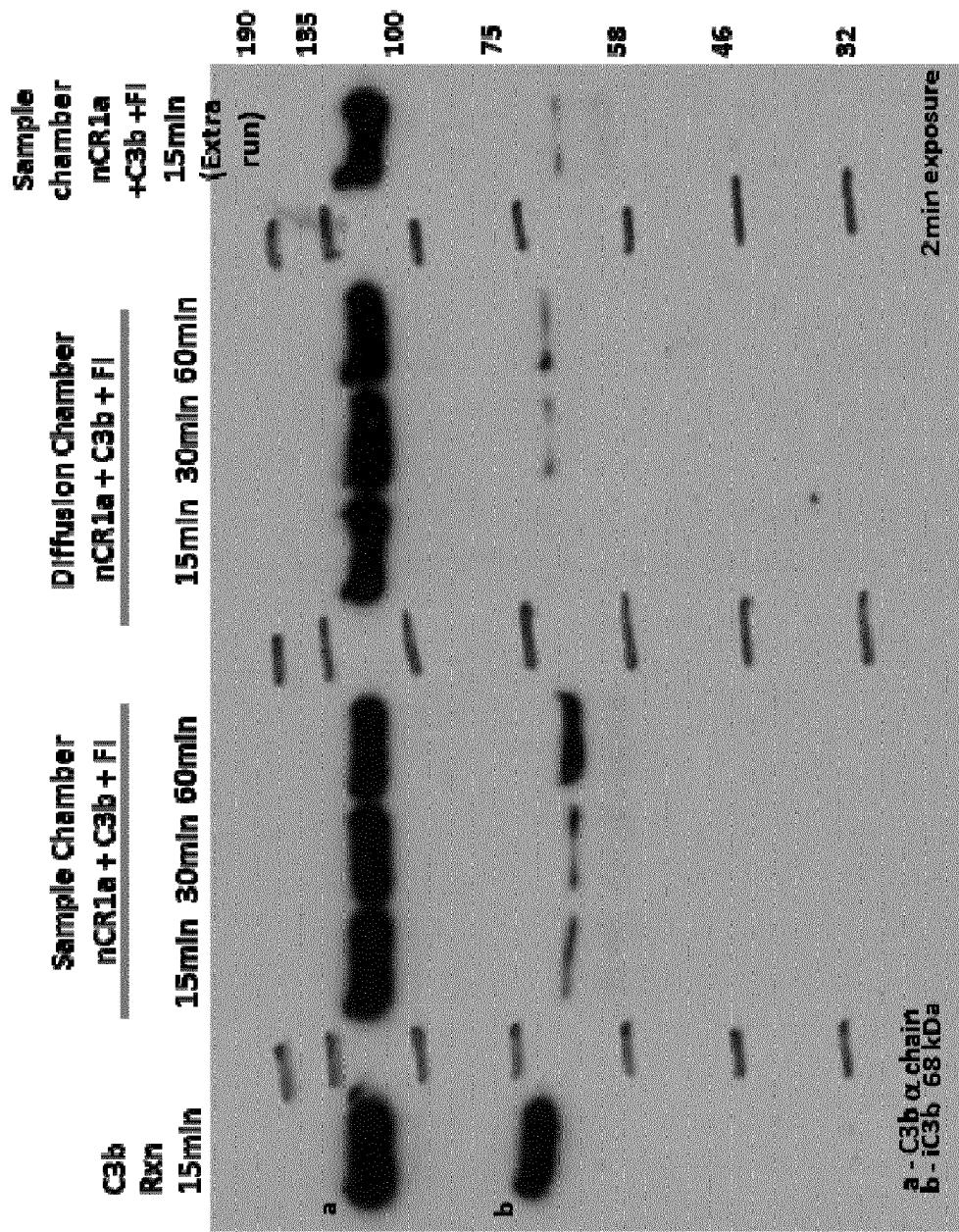


Figure 3C

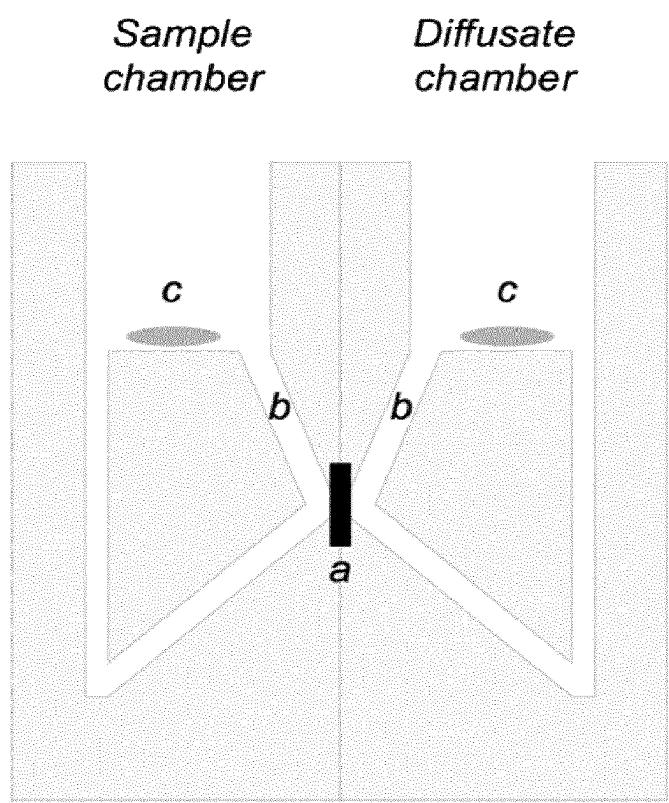


Figure 4

	FH	FHL-1	CR1a
	~155kDa	~50kDa	
KD (M)	5.83E-07	1.17E-06	2.10E-08
KD Error (%)	3.0192	3.1325	0.4186
R ^{max}	0.1792	0.1495	0.1987
R ^{max} Error (%)	1.1926	0.6311	0.0483
Full R ²	0.9770	0.9550	0.8731

Data Location
167 µg/ml
20180514

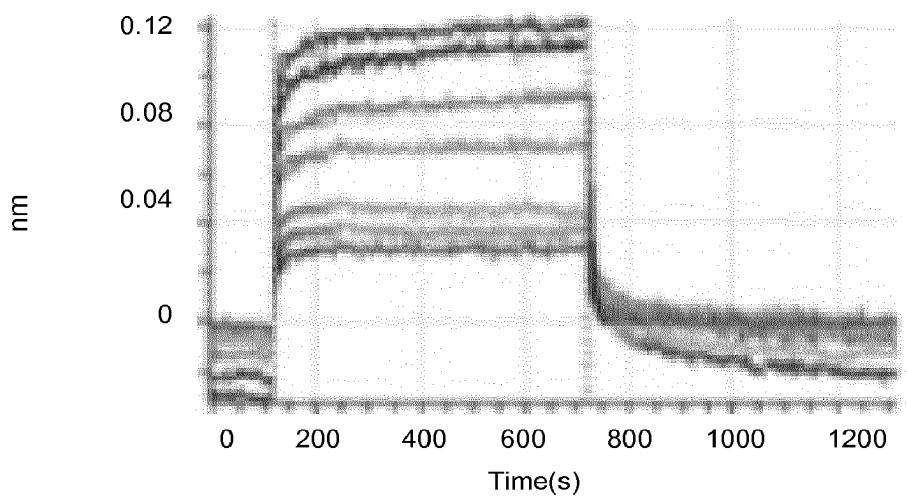
278 µg/ml
20180501

30 µg/ml
20180511

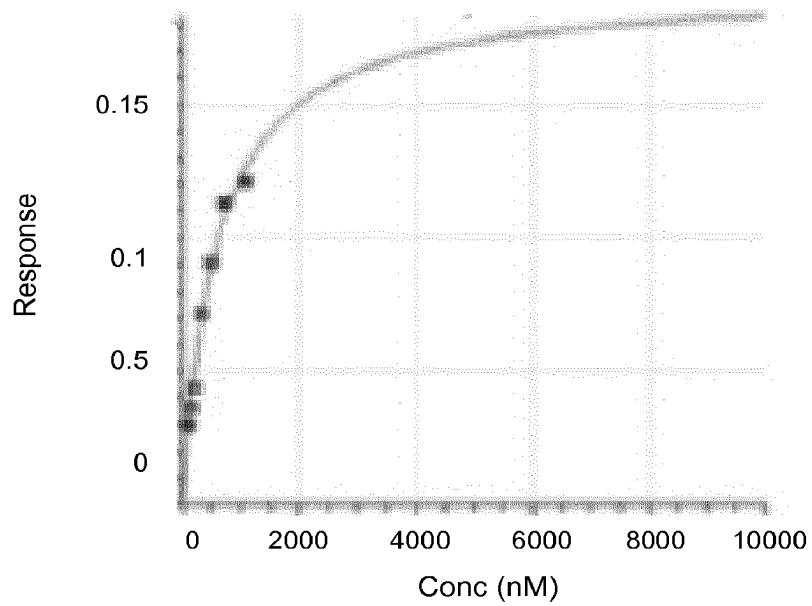
Figure 5

FH

Steady state no

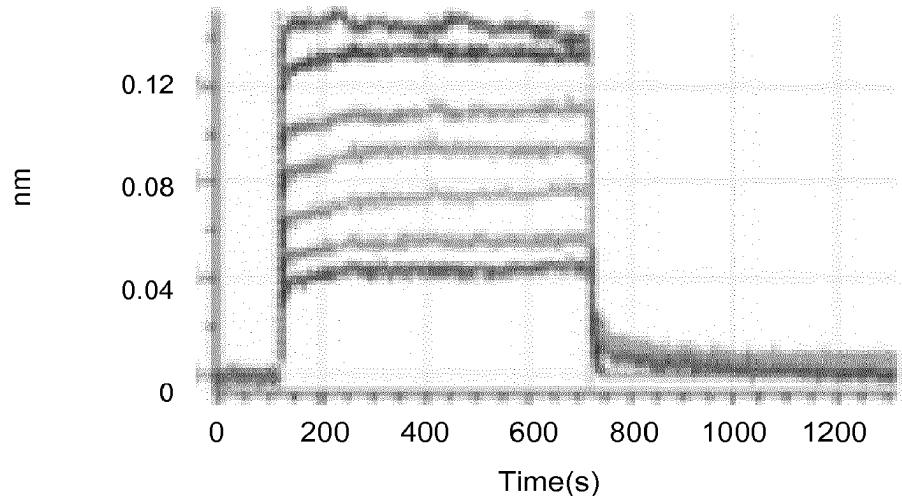


Steady state

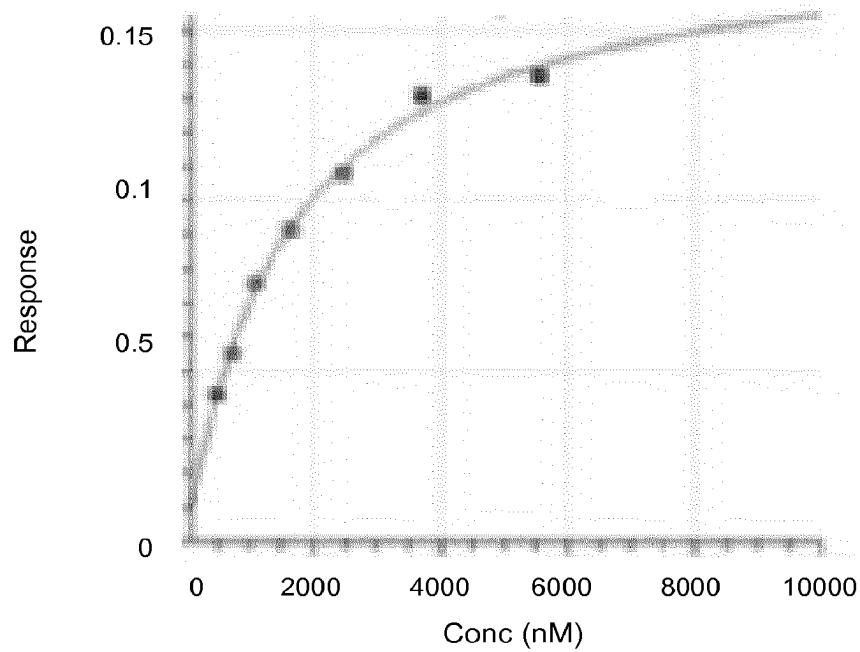
**Figure 5 (cont'd)**

FHL-1

Steady state no

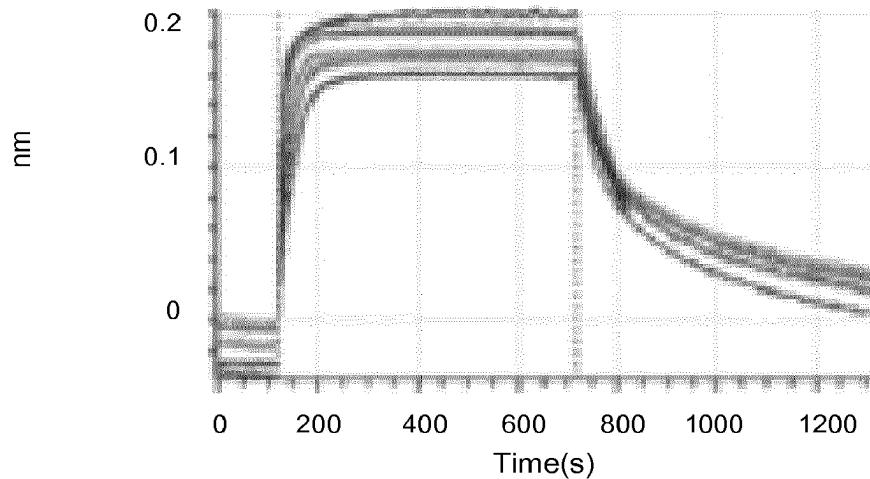


Steady state

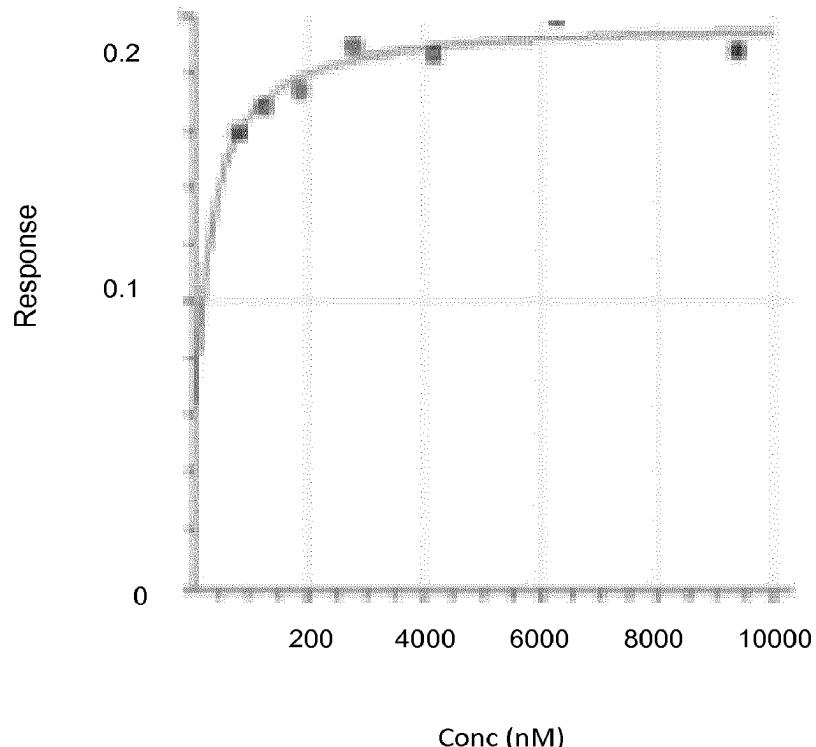
**Figure 5 (cont'd)**

CR1a

Steady state Yes



Steady state

**Figure 5 (cont'd)**

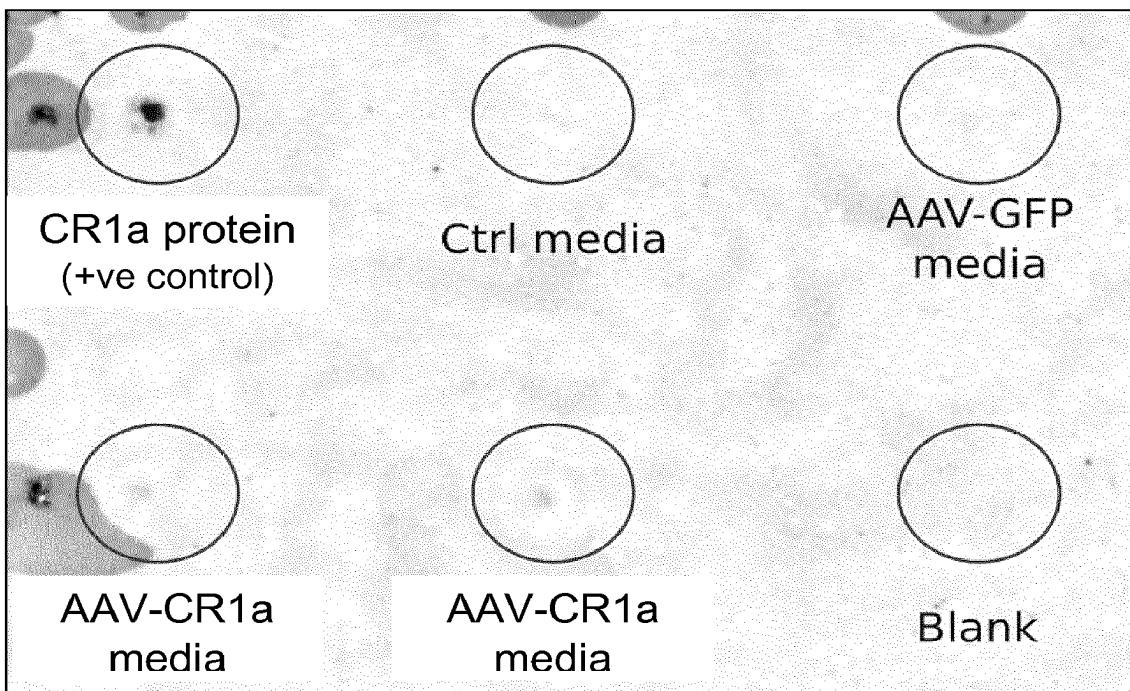
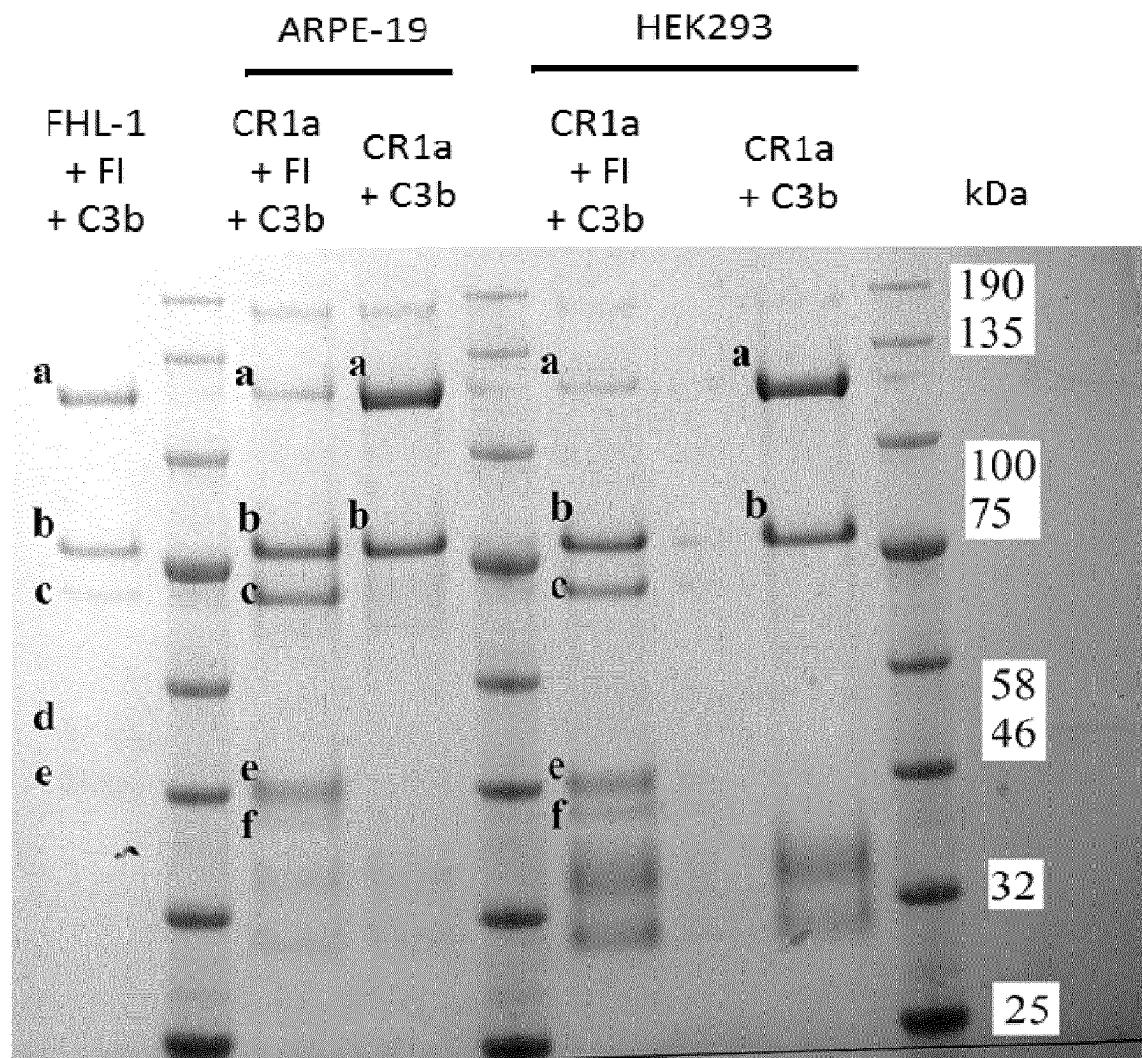


Figure 6



- a** - C3b alpha chain
- b** - C3b beta chain
- c** - iC3b 68 kDa
- d** - FHL-1
- e** - iC3b 43 kDa
- f** - C3dg 41 kDa (part of 68 kDa iC3b)

Figure 7A

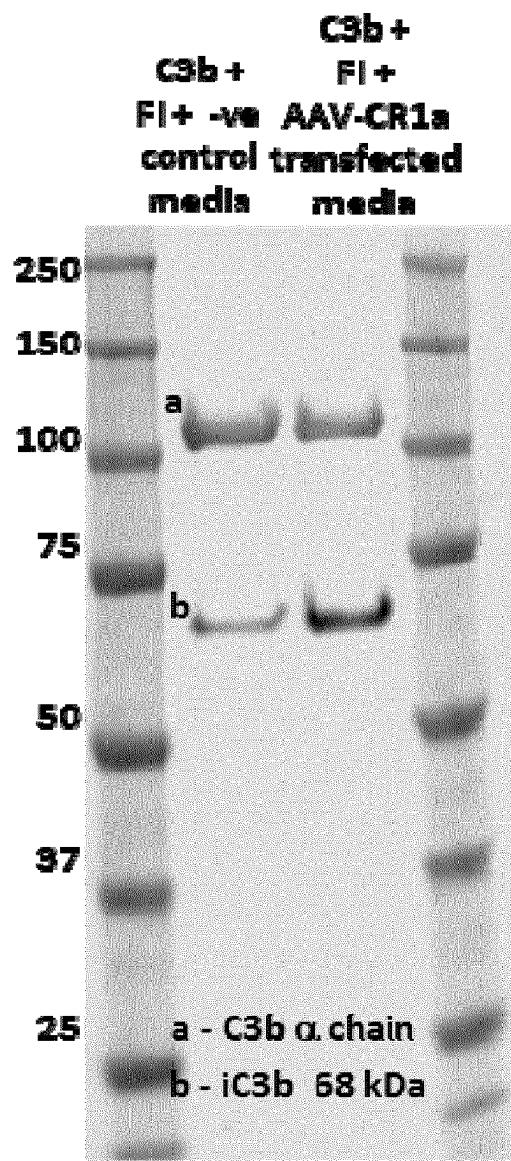


Figure 7B

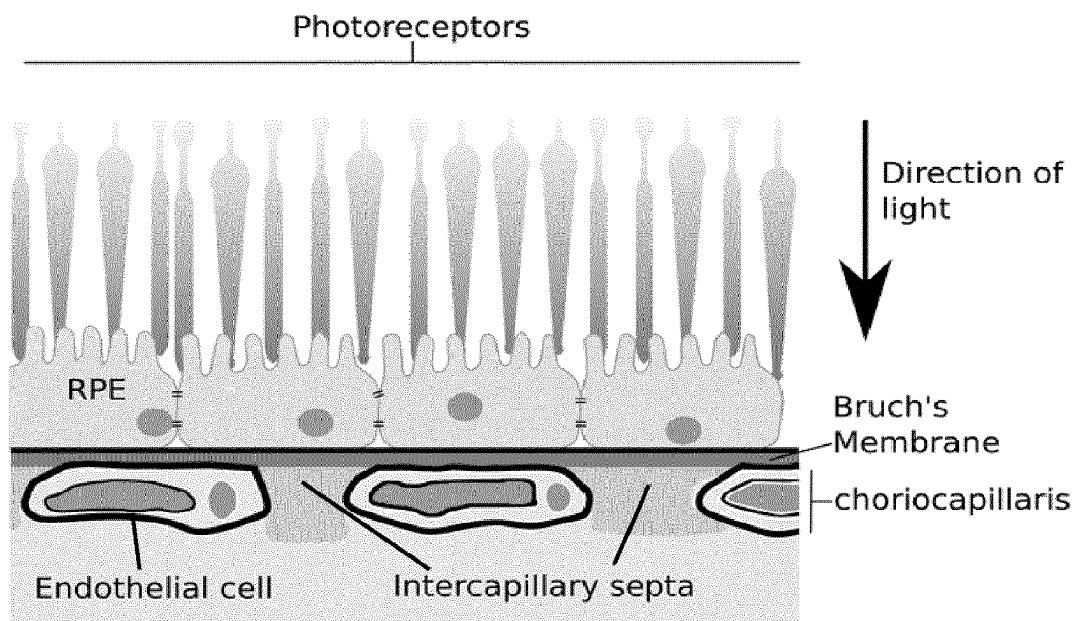


Figure 8

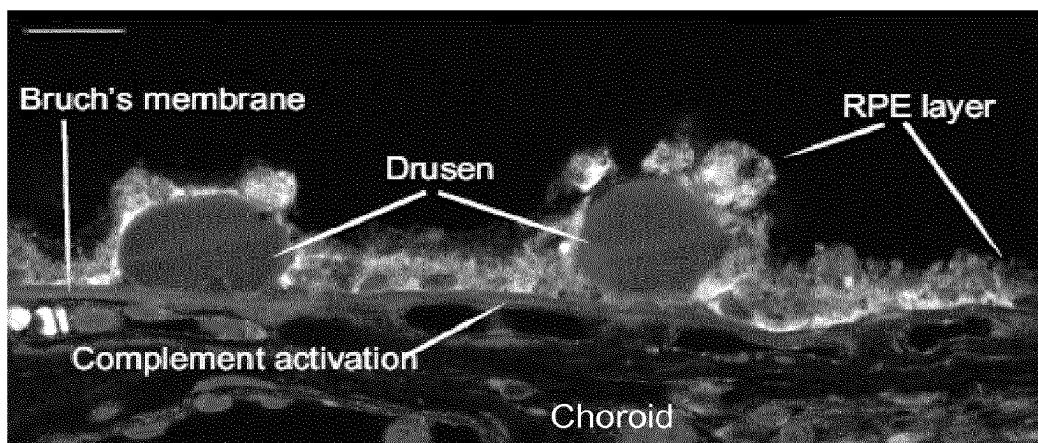


Figure 9

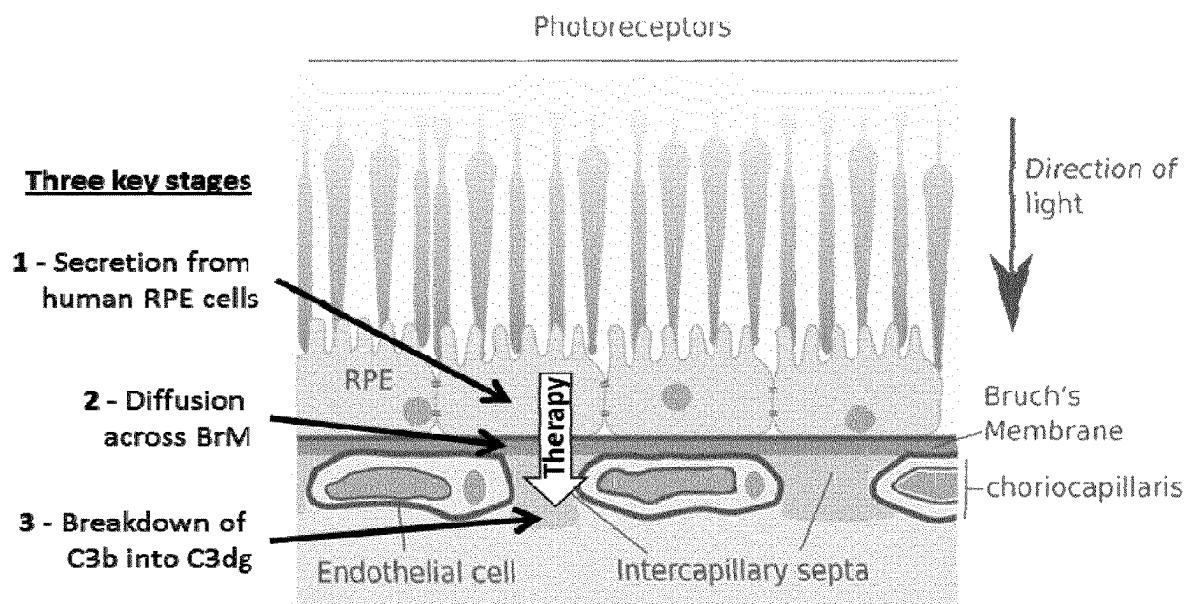


Figure 10

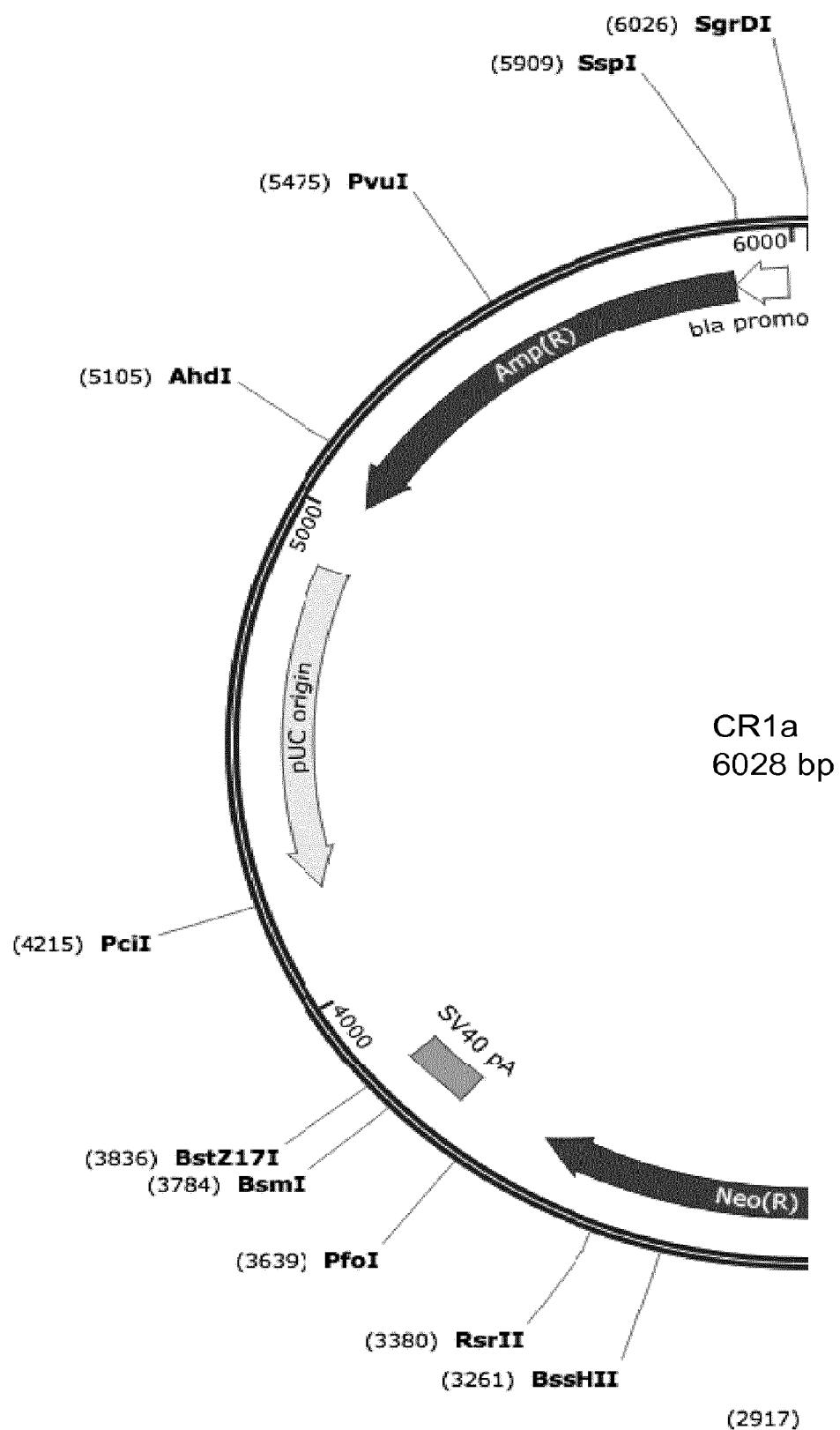


Figure 11

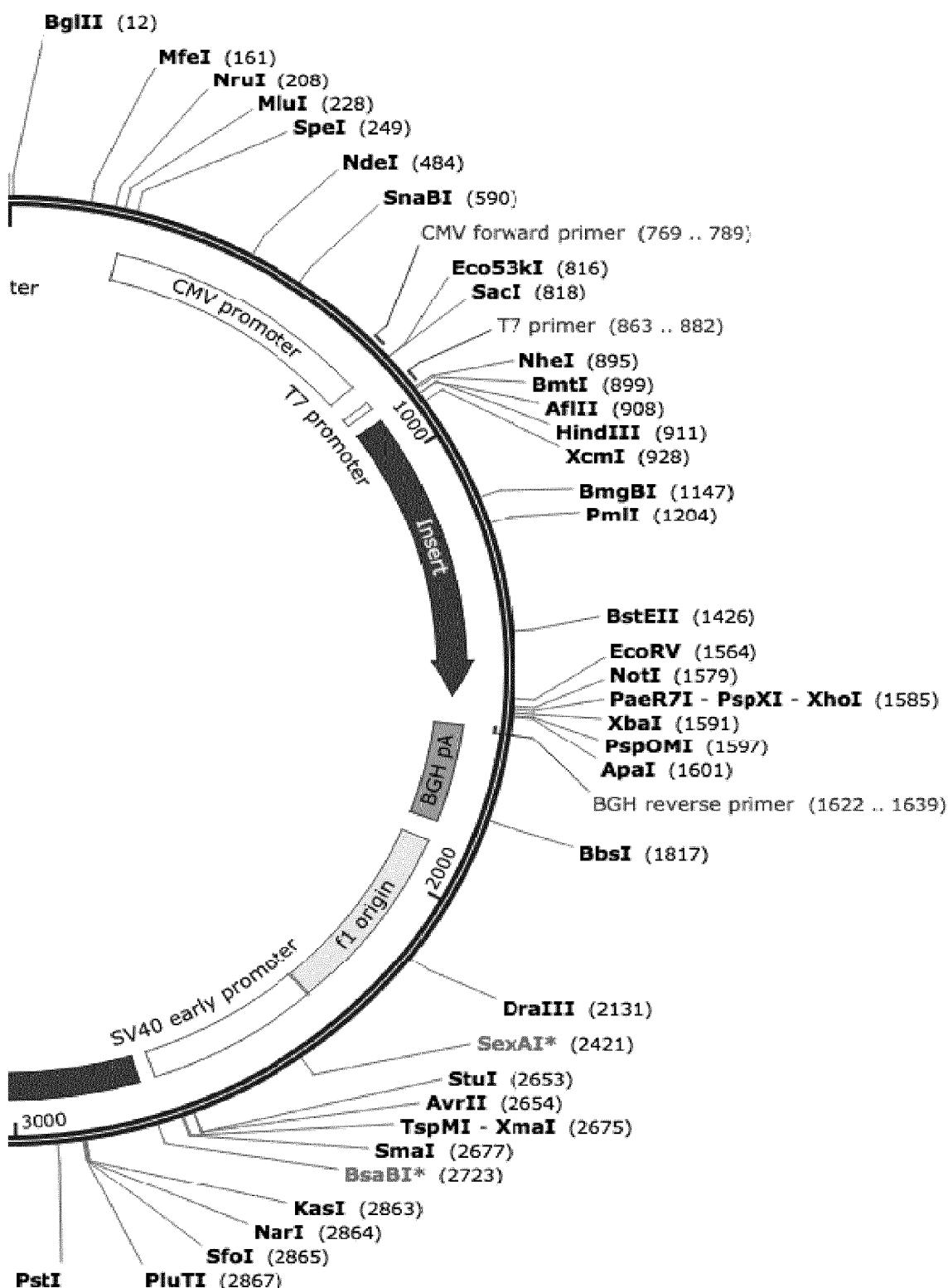


Figure 11 (continued)

1

C3B BINDING POLYPEPTIDE**CROSS-REFERENCE TO RELATED APPLICATIONS**

This application is a continuation of U.S. patent application Ser. No. 16/962,108 filed Jul. 14, 2020, which is a National Phase of International Application No. PCT/EP2019/050949 filed Jan. 15, 2019, which designated the U.S. and that International Application was published under PCT Article 21 (2) in English. This application also includes a claim of priority under 35 U.S.C. § 119 (a) and § 365 (b) to British patent application No. GB 1800620.5 filed Jan. 15, 2018, the entirety of which is hereby incorporated by reference.

REFERENCE TO A SEQUENCE LISTING

This application contains a sequence listing named "SequenceListing" in XML file format, created on Jan. 20, 2021 Nov. 4, 2022, being 73.6 kb in size, which is herein incorporated by reference as though fully disclosed.

FIELD OF THE INVENTION

The present invention relates to the fields of molecular biology, immunology, and medicine. More specifically, the present invention relates to a polypeptide that binds to C3b.

BACKGROUND TO THE INVENTION

Age-related macular degeneration (AMD) is the leading cause of blindness in the developed world: AMD is currently responsible for 8.7% of all global blind registrations and it is estimated that 196 million people will be affected by 2020 (Wong et al. Lancet Glob Heal (2014) 2:e106-16). AMD manifests as the progressive destruction of the macula, the central part of the retina at the back of the eye, leading to loss of central visual acuity. Early stages of the disease see morphological changes in the macula, including first the loss of blood vessels in the choriocapillaris (Whitmore et al., Prog Retin Eye Res (2015) 45:1-29) which are fenestrated blood vessels found in the choroid (a highly vascularized layer that supplies oxygen and nutrition to the outer retina).

AMD is largely a genetic disease. Mutations in genes of the complement system, part of our immune system, are highly associated with increased risk of AMD. Indeed, it has become clear that over-activation of complement is a main driver of disease pathogenesis and many examples of complement over-activation can be seen in the choriocapillaris. The role of complement in AMD is reviewed, for example, by Zipfel et al. Chapter 2, in Lambris and Adamis (eds.), Inflammation and Retinal Disease: Complement Biology and Pathology, Advances in Experimental Medicine and Biology 703, Springer Science+Business Media, LLC (2010), which is hereby incorporated by reference in its entirety. Complement is activated by the deposition onto a surface of protein C3b, a pro-inflammatory breakdown product of immune system protein C3. C3b associates with other proteins to form convertase enzyme complexes for activating and amplifying complement responses, and initiates the amplification loop of the complement cascade, ultimately leading to cell/tissue destruction and a local inflammatory response (all characteristics of AMD).

The choriocapillaris is separated from the metabolically active retinal pigment epithelium (RPE) by Bruch's membrane (BrM), a thin (2-4 µm), acellular, five-layered, extra-

2

cellular matrix. The BrM serves two major functions: the substratum of the RPE and a blood vessel wall. The structure and function of BrM is reviewed e.g. in Curcio and Johnson, Structure, Function and Pathology of Bruch's Membrane, In: Ryan et al. (2013), Retina, Vol. 1, Part 2: Basic Science and Translation to Therapy. 5th ed. London: Elsevier, pp466-481, which is hereby incorporated by reference in its entirety.

C3b activation of complement on acellular structures, such as BrM and the intercapillary septa (extracellular matrix filling the spaces between capillaries in the choriocapillaris), is regulated by proteins 'complement factor H' (FH) and 'complement factor I' (FI). FI prevents complement activation by cleaving C3b to an proteolytically-inactive form, designated iC3b, which is unable to participate in convertase assembly. However, iC3b is an opsonin and therefore a mediator of leucocyte recruitment with a subsequent inflammatory response, whereas the further breakdown products of C3b, iC3dg and C3d, are poor opsonins. In order to cleave C3b, FI requires the presence of a cofactor, examples of which include the blood borne FH protein and the membrane-bound surface co-factor 'complement factor I' (CR1; CD35). CR1 is a membrane receptor expressed on a wide range of cells and is involved in immune complex clearance, phagocytosis, and complement regulation. As well as serving as a co-factor in the FI-mediated cleavage of C3b, CR1 acts as a regulator of complement by accelerating the decay of C3 and C5 convertases. CR1 structure and function is reviewed e.g. in Khera and Das, Mol Immunol (2009) 46(5): 761-772 and Jacquet et al., J Immunol (2013) 190(7): 3721-3731, both of which are hereby incorporated by reference in their entirety.

Hallmark lesions of early AMD, termed drusen, develop within BrM adjacent to the RPE layer (Bird et al, Surv Ophthalmol (1995) 39(5):367-374). Drusen are formed from the accumulation of lipids and cellular debris, and include a swathe of complement activation products (Anderson et al., Prog Retin Eye Res (2009) 29:95-112; Whitcup et al., Int J Inflam (2013) 1-10). The presence of drusen within BrM disrupts the flow of nutrients from the choroid across this extracellular matrix to the RPE cells, which leads to cell dysfunction and eventual death. As the RPE cell monolayer supports the rod and cone cells of the neurosensory retina by providing nutrients and removing waste, their cell death causes dysfunction of photoreceptor cells and subsequent loss of visual acuity.

This represents one of the late stages of AMD, known as 'dry' AMD and also as geographic atrophy, which represents around 90% of AMD cases. In the remaining percentage of cases of late-stage AMD, the presence of drusen promotes choroidal neovascularisation (CNV), where the increased synthesis of vascular endothelial growth factor (VEGF) by RPE cells promotes new blood vessel growth from the choroid/choriocapillaris that breaks through BrM into the retina. These new blood vessels leak and eventually form scar tissue; this is referred to as 'wet' AMD.

'Wet' AMD, while only representing 10% of cases, is the most virulent form of late-stage AMD and has different disease characteristics to 'dry' AMD. There are treatments for wet AMD, where for example the injection of anti-VEGF agents into the vitreous of the eye can slow or reverse the growth of these blood vessels, although it cannot prevent their formation in the first place. Geographic atrophy ("dry" AMD) remains untreatable.

SUMMARY OF THE INVENTION

The present invention provides polypeptides that bind to C3b which are useful for treating or preventing complement-related diseases or conditions.

In one aspect, the present invention provides a polypeptide which is capable of binding C3b, the polypeptide comprising an amino acid sequence having at least 85% identity to SEQ ID NO:4 and wherein the polypeptide has a total length of 450 amino acids or fewer, for use in a method of treating or preventing a complement-related disease or condition.

Also provided is a nucleic acid encoding a polypeptide which is capable of binding C3b, the polypeptide comprising an amino acid sequence having at least 85% identity to SEQ ID NO:4 and wherein the polypeptide has a total length of 450 amino acids or fewer, for use in a method of treating or preventing a complement-related disease or condition.

In another aspect, the present invention provides the use of a polypeptide which is capable of binding C3b, the polypeptide comprising an amino acid sequence having at least 85% identity to SEQ ID NO:4 and wherein the polypeptide has a total length of 450 amino acids or fewer, in the manufacture of a medicament for treating or preventing a complement-related disease or condition.

Also provided is the use of a nucleic acid encoding a polypeptide which is capable of binding C3b, the polypeptide comprising an amino acid sequence having at least 85% identity to SEQ ID NO:4 and wherein the polypeptide has a total length of 450 amino acids or fewer, in the manufacture of a medicament for treating or preventing a complement-related disease or condition.

In another aspect, provided is a method of treating or preventing a complement-related disease or condition, comprising administering to a subject a polypeptide which is capable of binding C3b, the polypeptide comprising an amino acid sequence having at least 85% identity to SEQ ID NO:4 and wherein the polypeptide has a total length of 450 amino acids or fewer.

In another aspect, provided is a method of treating or preventing a complement-related disease or condition in a subject, comprising modifying at least one cell of the subject to express or comprise a polypeptide capable of binding C3b, the polypeptide comprising an amino acid sequence having at least 85% identity to SEQ ID NO:4 and wherein the polypeptide has a total length of 450 amino acids or fewer.

In some embodiments the complement-related disease or condition is an ocular disease or condition.

In some embodiments the treatment or prevention of an ocular disease or condition comprises modifying at least one ocular cell of a subject to express or comprise the polypeptide. In some embodiments the treatment or prevention of an ocular disease or condition comprises modifying at least one ocular cell of a subject to express or comprise a nucleic acid encoding the polypeptide. In some embodiments the treatment or prevention of an ocular disease or condition comprises administering a vector comprising a nucleic acid encoding the polypeptide to at least one ocular cell of a subject. In some embodiments the at least one ocular cell is a retinal pigment epithelial (RPE) cell.

In some embodiments the disease or condition is a disease or condition in which C3b or a C3b-containing complex, an activity/response associated with C3b or a C3b-containing complex, or a product of an activity/response associated with C3b or a C3b-containing complex is pathologically implicated.

In some embodiments the disease or condition is macular degeneration. In some embodiments the disease or condition is selected from one or more of: age-related macular degeneration (AMD), early AMD, intermediate AMD, late AMD, geographic atrophy ('dry' AMD), 'wet' (neovascular) AMD,

choroidal neovascularisation (CNV), glaucoma, autoimmune uveitis, diabetic retinopathy, and early-onset macular degeneration (EOMD).

5 In some embodiments the polypeptide comprises an amino acid sequence having at least 95% identity to SEQ ID NO:4. In some embodiments, X₁ is A or T, X₂ is P or L, and/or X₃ is G or R of SEQ ID NO:4.

In some embodiments the polypeptide has a total length of 50 to 250 amino acids. In some embodiments the polypeptide comprises, or consists of, SEQ ID NO:2 or SEQ ID NO:3.

In some embodiments the polypeptide comprises, or consists of, SEQ ID NO:13.

15 In some embodiments the polypeptide is capable of acting as a co-factor for Complement Factor I. In some embodiments the polypeptide is capable of diffusing across Bruch's membrane (BrM). In some embodiments the polypeptide binds to C3b in the region bound by a co-factor for Complement Factor I. In some embodiments the polypeptide binds 20 to C3b in the region bound by Complement Receptor 1 (CR1).

25 In some embodiments the polypeptide comprises a secretory pathway sequence. In some embodiments the secretory pathway sequence comprises, or consists of, SEQ ID NO:7. In some embodiments the polypeptide comprises, or consists of, SEQ ID NO:47, 49, or 51. In some embodiments the polypeptide comprises a cleavage site for removing the secretory pathway sequence.

30 In one aspect, the present invention provides a polypeptide having at least 80% sequence identity to SEQ ID NO:4, wherein the polypeptide has a length of 700 amino acids or fewer.

In some embodiments, the polypeptide has a length of 50 to 700 amino acids. In some embodiments, the polypeptide has at least 80% sequence identity to SEQ ID NO:4 wherein X₁ is A or T, X₂ is P or L, and/or X₃ is G or R.

35 40 Also provided is a polypeptide which is capable of binding to C3b, the polypeptide comprising an amino acid sequence having at least 85% sequence identity to SEQ ID NO:4 and wherein the polypeptide has a total length of 450 amino acids or fewer.

In some embodiments the polypeptide comprises an amino acid sequence having at least 95% identity to SEQ ID NO:4. In some embodiments X₁ is A or T, X₂ is P or L, and/or X₃ is G or R.

In some embodiments the polypeptide has a total length of 50 to 250 amino acids.

45 50 55 In some embodiments, the polypeptide comprises, or consists of, an amino acid sequence according to SEQ ID NO:2. In some embodiments, the polypeptide comprises, or consists of, an amino acid sequence according to SEQ ID NO:3. In some embodiments, the polypeptide comprises, or consists of, an amino acid sequence according to SEQ ID NO:13.

60 65 70 In some embodiments, the polypeptide is capable of binding to C3b. In some embodiments, the polypeptide binds to C3b in the region bound by a co-factor for Complement Factor I. In some embodiments, the polypeptide binds to C3b in the region bound by Complement Receptor 1 (CR1). In some embodiments, the polypeptide acts as a co-factor for Complement Factor I.

In some embodiments, the polypeptide is capable of diffusing across Bruch's membrane (BrM). In some embodiments, the polypeptide is not glycosylated or is partially glycosylated. In some embodiments, the polypeptide comprises at least one amino acid substitution, e.g. one, two, three or four substitutions, at position 509, 578, 959 and/or

1028 (numbered according to Uniprot: P17927 (SEQ ID NO:1)). In some embodiments, the at least one amino acid substitution is one or more of N509Q, N578Q, N959Q and/or N1028Q (numbered according to Uniprot: P17927 (SEQ ID NO:1)). In some embodiments, the polypeptide comprises, or consists of, SEQ ID NO:5, SEQ ID NO:6, and/or SEQ ID NO:15.

In some embodiments, the polypeptide additionally comprises a secretory pathway sequence. In some embodiments, the secretory pathway sequence comprises, or consists of, SEQ ID NO:7. In some embodiments, the polypeptide additionally comprises a cleavage site for removing the secretory pathway sequence. In some embodiments the polypeptide comprises, or consists of, SEQ ID NO:47, 49 or 51.

In another aspect, the present invention provides a nucleic acid encoding a polypeptide according to the present invention.

In another aspect, the present invention provides a vector comprising a nucleic acid of the present invention.

In another aspect, the present invention provides a cell comprising a polypeptide, nucleic acid, or vector according to the present invention.

In another aspect, the present invention provides a method for producing a polypeptide, comprising introducing into a cell a nucleic acid or a vector according to the present invention, and culturing the cell under conditions suitable for expression of the polypeptide.

In another aspect, the present invention provides a cell, which is obtained or obtainable by the method for producing a polypeptide according to the present invention.

In another aspect, the present invention provides a pharmaceutical composition comprising a polypeptide, nucleic acid, vector or cell according to the present invention. In some embodiments, the pharmaceutical composition comprises a pharmaceutically acceptable carrier, adjuvant, excipient, or diluent.

In another aspect, the present invention provides a polypeptide, nucleic acid, vector, cell or pharmaceutical composition according to the present invention, for use in a method of treating or preventing a disease or condition.

In another aspect, the present invention provides the use of a polypeptide, nucleic acid, vector, cell or pharmaceutical composition according to the present invention, in the manufacture of a medicament for treating or preventing a disease or condition.

In another aspect, the present invention provides a method of treating or preventing a disease or condition, comprising administering to a subject a polypeptide, nucleic acid, vector, cell or pharmaceutical composition according to the present invention.

In another aspect, the present invention provides a method of treating or preventing a disease or condition in a subject, comprising modifying at least one cell of the subject to express or comprise a nucleic acid, vector or polypeptide according to the present invention.

In some embodiments in accordance with various aspects of the present invention, the disease or condition is a disease or condition in which C3b or a C3b-containing complex, an activity/response associated with C3b or a C3b-containing complex, or a product of an activity/response associated with C3b or a C3b-containing complex is pathologically implicated. In some embodiments, the disease or condition is macular degeneration. In some embodiments, the disease or condition is age-related macular degeneration (AMD). In some embodiments, the method for treating or preventing a disease or condition comprises modifying at least one retinal

pigment epithelial (RPE) cell of the subject to express or comprise a nucleic acid, vector, or polypeptide according to the present invention.

In another aspect, the present invention provides a kit of parts comprising a predetermined quantity of a polypeptide, nucleic acid, vector, cell, or pharmaceutical composition according to the present invention.

DESCRIPTION

Complement-based therapies for AMD have thus far concentrated upon injecting complement regulating antibodies into the eye. Such therapies have provided little to no therapeutic benefit as these proteins cannot reach the target area, i.e. the BrM and its underlying vasculature, the choriocapillaris at all, or in effective concentrations.

Complement Factor I (FI)-mediated regulation of complement, i.e. the cleavage of C3b to iC3b (proteolytically-inactive C3b), requires cofactors such as membrane-anchored CR1. However, the present inventors have discovered that it is not necessary to provide full-length membrane-bound CR1, nor even a soluble version of CR1 merely lacking the transmembrane domain, for successful cofactor activity. Instead, the inventors have discovered that short CR1 fragments comprising the CR1 C3b-binding domains are sufficient to enable efficient FI-mediated C3b cleavage.

Thus, the present invention relates to soluble, truncated polypeptides derived from the FI cofactor CR1. The polypeptides comprise domains that are capable of binding to C3b, such that they can act as essential FI cofactors for the regulation of complement activation. A key advantage of the soluble, truncated polypeptides is their ability to pass through BrM, and thus they are able to reach all regions associated with AMD, i.e. the RPE/BrM interface, BrM and the choroid, including the intercapillary septa (the extracellular matrix between the blood vessels of the choriocapillaris). The present invention also provides non-glycosylated polypeptides derived from CR1, which may aid polypeptide passage through BrM. The polypeptides are expressed and secreted easily, enabling in situ expression by cells local to the affected sites and targeting of the polypeptides to areas affected by complement over-activation. In situ expression of the polypeptides may be achieved using gene therapy techniques. In situ expression provides targeted therapy to areas of need without disrupting functioning complement regulation elsewhere in the body.

The present invention enables supplementation of a deteriorating complement regulation system without replacing the endogenous complement regulation currently in place, or interfering in the rest of the complement cascade.

Polypeptides

A polypeptide according to the present invention may comprise, or consist of, one or more C3b binding regions.

A polypeptide according to the present invention has at least 80% sequence identity to SEQ ID NO:4, wherein the polypeptide has a length of 700 amino acids or fewer. In some embodiments, the polypeptide comprises, or consists of, an amino acid sequence having at least 80% sequence identity to SEQ ID NO:4, wherein the polypeptide has a length of 700 amino acids or fewer.

A polypeptide according to the present invention may comprise, or consist of, an amino acid sequence of 700 amino acids or fewer having at least 80% sequence identity to SEQ ID NO:4.

In some embodiments, the polypeptide comprises, or consists of, an amino acid sequence of 650, 600, 550, 500,

450, 400, 350, 300, 250, or 200 amino acids or fewer. In some embodiments, the polypeptide comprises, or consists of, an amino acid sequence having 1 to 200 amino acids, 1 to 250 amino acids, 1 to 300 amino acids, 1 to 350 amino acids, 1 to 400 amino acids, 1 to 450 amino acids, 1 to 500 amino acids, 1 to 550 amino acids, 1 to 600 amino acids, 1 to 650 amino acids, or 1 to 700 amino acids. In some embodiments, the polypeptide has a length of 50 to 700 amino acids. In some embodiments, the polypeptide has a length of 100 to 650 amino acids. In some embodiments, the polypeptide has a length of 100 to 550 amino acids. In some embodiments, the polypeptide has a length of 150 to 450 amino acids. In some embodiments, the polypeptide has a length of 400 to 700 amino acids. In some embodiments, the polypeptide has a length of 700 to 1000 or greater than 1000 amino acids.

"Length" as used herein refers to the total length of the polypeptide; that is, "length" refers to the measurement or extent of the entire polypeptide from end to end, i.e. from the N-terminus to the C-terminus. "Length" as used herein is measured by the number of amino acid residues within the polypeptide.

In some embodiments the polypeptide is a detached/discrete/separate/individual molecule. In some embodiments, the polypeptide is a single contiguous amino acid sequence that is unconnected, i.e. not joined, fused or attached, to another amino acid sequence. In some embodiments the polypeptide is not attached by an amino acid linker or a non-amino acid linker to another polypeptide or amino acid sequence. In some embodiments the polypeptide is not a section, part or region of a longer amino acid sequence, i.e. it is not part of an amino acid sequence that exceeds the maximum, specified, polypeptide length. In some embodiments the polypeptide is not part of, or does not form a section of, a fusion protein. In some embodiments the polypeptide may comprise a sequence provided herein and one or more additional amino acids, as long as the maximum length of the polypeptide is not exceeded. The short length of the polypeptides described herein enables the polypeptides to pass through the BrM and reach sites of complement activation.

In some embodiments the polypeptide has a total length of 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, or 200 amino acids or fewer. In some embodiments the polypeptide has a total length of 450, 440, 430, 420, 410, 400, 390, 380, 370, 360, 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200, 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, or 70 amino acids or fewer. In some embodiments the polypeptide has a total length of 1 to 70 amino acids, 1 to 80 amino acids, 1 to 90 amino acids, 1 to 100 amino acids, 1 to 110 amino acids, 1 to 120 amino acids, 1 to 130 amino acids, 1 to 140 amino acids, 1 to 150 amino acids, 1 to 160 amino acids, 1 to 170 amino acids, 1 to 180 amino acids, 1 to 190 amino acids, 1 to 200 amino acids, 1 to 210 amino acids, 1 to 220 amino acids, 1 to 230 amino acids, 1 to 240 amino acids, 1 to 250 amino acids, 1 to 260 amino acids, 1 to 270 amino acids, 1 to 280 amino acids, 1 to 290 amino acids, 1 to 300 amino acids, 1 to 310 amino acids, 1 to 320 amino acids, 1 to 330 amino acids, 1 to 340 amino acids, 1 to 350 amino acids, 1 to 360 amino acids, 1 to 370 amino acids, 1 to 380 amino acids, 1 to 390 amino acids, 1 to 400 amino acids, 1 to 410 amino acids, 1 to 420 amino acids, 1 to 430 amino acids, 1 to 440 amino acids, or 1 to 450 amino acids. In some embodiments the polypeptide has a total length of 50 to 450 amino acids, 50 to 400 amino acids, 50 to 350 amino acids, 50 to 300 amino acids, 50 to 250 amino acids, 50 to 200 amino acids, 100 to

250 amino acids, 100 to 200 amino acids, 150 to 250 amino acids, or 150 to 200 amino acids. In some embodiments the polypeptide has a total length of one of 61, 72, 194, 212, 231, 388, 406, or 644 amino acids.

- 5 In some embodiments a polypeptide of the present invention has a maximum molecular weight of 80 kDa, whether the polypeptide is covalently/non-covalently bonded to a larger complex, part of a larger complex, or is not part of a larger complex. In some embodiments a polypeptide of the 10 present invention has a molecular weight of 75 kDa or less, 70 kDa or less, 65 kDa or less, 60 kDa or less, 55 kDa or less, 50 kDa or less, 45 kDa or less, 40 kDa or less, 35 kDa or less, 30 kDa or less, 29 kDa or less, 28 kDa or less, 27 kDa or less, 26 kDa or less, 25 kDa or less, 24 kDa or less, 23 kDa or less, 22 kDa or less, 21 kDa or less, 20 kDa or less, 19 kDa or less, 18 kDa or less, 17 kDa or less, 16 kDa or less, 15 kDa or less, 14 kDa or less, 13 kDa or less, 12 kDa or less, 11 kDa or less, or 10 kDa or less. In some embodiments the polypeptide has a maximum molecular weight of 50 kDa, i.e. 50 kDa or less.
- 15 20 In some embodiments the polypeptide has a maximum molecular weight of 26 kDa, i.e. 26 kDa or less. In some embodiments the polypeptide has a maximum molecular weight of 24 kDa, i.e. 24 kDa or less. In some embodiments the polypeptide has a maximum molecular weight of 22 kDa, i.e. 22 kDa or less. In some embodiments the polypeptide has a maximum molecular weight of 20 kDa, i.e. 20 kDa or less.

In some embodiments, a polypeptide of the present invention comprises, or consists of, an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:4. In some embodiments, X₁ is A or T, X₂ is P or L, and/or X₃ is G or R. In some embodiments, X₁ is A, X₂ is P, and/or X₃ is G. In some embodiments, X₁ is A, X₂ is L, and/or X₃ is R. In some embodiments, X₁ is A, X₂ is P, and/or X₃ is R. In some embodiments, X₁ is A, X₂ is L, and/or X₃ is G. In some embodiments, X₁ is T, X₂ is P, and/or X₃ is R. In some embodiments, X₁ is T, X₂ is P, and/or X₃ is G. In some embodiments, X₁ is T, X₂ is L, and/or X₃ is G. In some embodiments, X₁ is T, X₂ is P, and/or X₃ is R.

In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:4, wherein the polypeptide has a length provided herein. For example, in some embodiments the polypeptide comprises, or consists of, an amino acid sequence having 85% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 450, 440, 430, 420, 410, 400, 390, 380, 370, 360, 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200, 190, 180, 170, or 160 amino acids or fewer; in some embodiments the polypeptide comprises, or consists of, an amino acid sequence having 90% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 450, 440, 430, 420, 410, 400, 390, 380, 370, 360, 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200, 190, or 180 amino acids or fewer; in some embodiments the polypeptide comprises, or consists of, an amino acid sequence having 95% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 450, 440, 430, 420, 410, 400, 390, 380, 370, 360, 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200, 190, or 180 amino acids or fewer; in some embodiments the polypeptide comprises, or consists of, an amino acid sequence having 98% sequence identity to SEQ ID

NO:4, wherein the polypeptide has a total length of 450, 440, 430, 420, 410, 400, 390, 380, 370, 360, 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, or 200 amino acids or fewer.

In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 85% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 450 amino acids or fewer. In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 85% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 50 to 450 amino acids. In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 85% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 250 amino acids or fewer. In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 85% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 50 to 250 amino acids. In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 95% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 450 amino acids or fewer. In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 95% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 50 to 450 amino acids. In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 95% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 250 amino acids or fewer. In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 95% sequence identity to SEQ ID NO:4, wherein the polypeptide has a total length of 50 to 250 amino acids.

Human CR1 (UniProt: P17927 (Entry version 181 (25 Oct. 2017), Sequence version 3 (2 Mar. 2010)); SEQ ID NO:1) has a 2,039 amino acid sequence (including an N-terminal, 41 amino acid signal peptide), and comprises 30 complement control protein (CCP) domains (also known as sushi domains or short consensus repeats (SCRs)), with the N-terminal 28 CCPs organised into four long homologous repeat (LHR) domains each comprising 7 CCPs: LHR-A, LHR-B, LHR-C and LHR-D. The C3b binding region of CR1 is found in CCPs 8-10 in LHR-B (UniProt: P17927 positions 491 to 684; SEQ ID NO:2), and CCPs 15-17 in LHR-C (UniProt: P17927 positions 941 to 1134; SEQ ID NO:3). CCPs 8-10 and 15-17 differ in sequence by three amino acid residues, as shown in consensus sequence SEQ ID NO:4.

A polypeptide according to the present invention may comprise, or consist of, an amino acid sequence corresponding to CCPs 8-10 (SEQ ID NO:2) and/or CCPs 15-17 (SEQ ID NO:3). In some embodiments, a polypeptide of the present invention comprises, or consists of, an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:2 and/or SEQ ID NO:3. Such polypeptides may have any length provided herein.

A polypeptide according to the present invention may comprise, or consist of, an amino acid sequence corresponding to CCPs 8-10 and 15-17. The polypeptide may comprise or consist of CCPs 8-10 and 15-17 in their native CR1 sequence (SEQ ID NO:30). The polypeptide may comprise or consist of CCPs 8-10 joined to CCPs 15-17. This may be in a contiguous sequence (SEQ ID NO:13), or achieved by a linker between CCPs 8-10 and 15-17 (e.g. SEQ ID

NO:14). In some embodiments, a polypeptide of the present invention comprises, or consists of, an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:13, SEQ ID NO:14 and/or SEQ ID NO:30. Such polypeptides may have any length provided herein. A polypeptide according to the present invention may comprise, or consist of, an amino acid sequence corresponding to one or more of sequence 'A' (SEQ ID NO:8), sequence 'B' (SEQ ID NO:16) and/or 'Sequence C' (SEQ ID NO:17). In some embodiments, the polypeptide consists of a sequence selected from sequence 'A' (SEQ ID NO:8), sequence 'B' (SEQ ID NO:16) and 'Sequence C' (SEQ ID NO:17). In some embodiments, a polypeptide of the present invention comprises, or consists of, an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to one or more of sequences 'A' (SEQ ID NO:8), 'B' (SEQ ID NO:16) and/or 'C' (SEQ ID NO:17).

In some embodiments, where a polypeptide according to the present invention comprises, or consists of, sequence 'B' (SEQ ID NO:16) or an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to sequence 'B' (SEQ ID NO:16), X₁ is A or T. In some embodiments, X₁ is A. In some embodiments, X₁ is T.

In some embodiments, where a polypeptide according to the present invention comprises, or consists of, sequence 'C' (SEQ ID NO:17) or an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to sequence 'C' (SEQ ID NO:17), X₂ is P or L and/or X₃ is G or R. In some embodiments, X₂ is P and/or X₃ is G. In some embodiments, X₂ is L and/or X₃ is R. In some embodiments, X₂ is G and/or X₃ is P. In some embodiments, X₂ is L and/or X₃ is R.

In some embodiments sequence 'B' corresponds to SEQ ID NO:9 or SEQ ID NO:11. In some embodiments sequence 'C' corresponds to SEQ ID NO:10 or SEQ ID NO:12.

The present invention includes polypeptides comprising sequences 'A', 'B', and/or 'C' as described herein, and combinations thereof, including at least the following combinations (organised from N-terminus to C-terminus):

- A+B
- B+C
- A+C
- C+A
- A+B+C
- B+C+A
- C+A+B
- A+B+C+A
- B+C+A+B
- C+A+B+C
- A+B+C+A+B
- B+C+A+B+C
- A+B+C+A+B+C
- A+B+C+A+B+C+Y (where Y=one or more of A, B and/or C).

In some embodiments, the polypeptide comprises, or consists of, an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%,

11

99%, or 100% sequence identity to any combination of CCP domains or any combination of sequences ‘A’, ‘B’ and/or ‘C’ described herein.

In some embodiments, the combination of CCP domains is a combination found in native CR1. In some embodiments, the combination of CCP domains is not a combination found in native CR1.

In some embodiments, the polypeptide comprises, or consists of, an amino acid sequence having multiple copies of sequence ‘A’, multiple copies of sequence ‘B’, and/or multiple copies of sequence ‘C’. In some embodiments, the polypeptide comprises 1, 2, 3, 4, 5, 6, 7, 8 or more copies of one or more of sequences ‘A’, ‘B’ and/or ‘C’. In some embodiments, the polypeptide comprises 1, 2, 3, 4, 5, 6, 7, 8 or more copies of sequence ‘A’; 1, 2, 3, 4, 5, 6, 7, 8 or more copies of sequence ‘B’; and/or 1, 2, 3, 4, 5, 6, 7, 8 or more copies of sequence ‘C’. In some embodiments, the polypeptide comprises 9, 10, or more copies of one or more of sequences ‘A’, ‘B’ and/or ‘C’.

In some embodiments, a polypeptide according to the present invention lacks substantial sequence identity to one or more of amino acid sequences 1-490, 685-940 and/or 1135-2039 of human CR1 (SEQ ID NO:1). A polypeptide that lacks substantial sequence identity as described herein may have less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, less than 50%, less than 45%, less than 40%, less than 35%, less than 30%, less than 25%, less than 20%, less than 15%, less than 10%, or less than 5% sequence identity to one or more of amino acid sequences 1-490, 685-940 and/or 1135-2039 of human CR1 (SEQ ID NO:1). In some embodiments a polypeptide according to the present invention lacks amino acid sequence having substantial sequence identity to CR1 long homologous repeat (LHR) domains LHR-A and/or LHR-D. In some embodiments a polypeptide according to the present invention lacks amino acid sequence having substantial sequence identity to CR1 CCP domains 1-7, 11-14 and/or 18-30. Amino acid residues of SEQ ID NO:1 are numbered according to Uniprot P17927; Entry version 181 (25 Oct. 2017), Sequence version 3 (2 Mar. 2010).

A polypeptide according to the present invention, and/or described herein, may be isolated and/or substantially purified.

Further Features of the Polypeptide

Polypeptides according to the present invention may comprise modifications and/or additional amino acid sequences. The modifications and/or additional amino acid sequences may be included in the length limitation of a polypeptide provided herein such that the length limitation of that polypeptide is not exceeded.

In some embodiments, an additional amino acid sequence comprises, or consists of, no more than 25, 50, 100, 150, or 200 amino acids, i.e. an additional amino acid sequence comprises, or consists of, 1-25, 1-50, 1-100, 1-150, or 1-200 amino acids. In some embodiments, an additional amino acid sequence comprises more than 200 amino acids. In some embodiments, an additional amino acid sequence comprises no more than 100 amino acids at the C-terminus of a polypeptide according to the present invention, and/or no more than 100 amino acids at the N-terminus of a polypeptide according to the present invention.

In some embodiments, an additional amino acid sequence results in a polypeptide longer than 700 amino acids. In some embodiments, a polypeptide according to the present invention comprises, or consists of, 700 or more amino acids. For example, the polypeptide may comprise, or con-

12

sist of, 700-750, 750-800, 800-850, 850-900, 900-950, 950-100, or more than 1000 amino acids.

In some embodiments, an additional amino acid sequence described herein lacks substantial sequence identity to one or more of amino acid sequences 1-490, 685-940 and/or 1135-2039 of human CR1 (SEQ ID NO:1, numbered according to Uniprot P17927; Entry version 181 (25 Oct. 2017), Sequence version 3 (2 Mar. 2010)). In some embodiments, the additional amino acid sequence lacks substantial sequence identity to CR1 CCP domains 1-7, 11-14 and/or 18-30. In some embodiments, the additional amino acid sequence has less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, less than 50%, less than 45%, less than 40%, less than 35%, less than 30%, less than 25%, less than 20%, less than 15%, less than 10%, or less than 5% sequence identity to one or more of amino acid sequences 1-490, 685-940 and/or 1135-2039 of human CR1 (SEQ ID NO:1). In some embodiments the additional amino acid sequence lacks substantial sequence identity to CR1 long homologous repeat (LHR) domains LHR-A and/or LHR-D.

In some embodiments, a polypeptide may lack amino acid sequence having substantial sequence identity to a region of a co-factor for Complement Factor I (e.g. CR1) other than in the C3b binding region. For example, the polypeptide may lack amino acid sequence having substantial sequence identity to CR1 other than in CR1 CCP domains 8-10 and/or 15-17 (residues 491 to 684 and/or 941 to 1134, respectively, of SEQ ID NO:1). In some embodiments, the polypeptide may lack amino acid sequence having substantial sequence identity to CR1 CCP domains 1-7, 11-14 and/or 18-30. A polypeptide lacking amino acid sequence having substantial sequence identity as described herein may have less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, less than 50%, less than 45%, less than 40%, less than 35%, less than 30%, less than 25%, less than 20%, less than 15%, less than 10%, or less than 5% sequence identity to one or more of amino acid sequences 1-490, 685-940 and/or 1135-2039 of human CR1 (SEQ ID NO:1).

In some embodiments, a polypeptide according to the present invention may comprise a secretory pathway sequence. As used herein, a secretory pathway sequence is an amino acid sequence which directs secretion of polypeptide. The secretory pathway sequence may be cleaved from the mature protein once export of the polypeptide chain across the rough endoplasmic reticulum is initiated. Polypeptides secreted by mammalian cells generally have a signal peptide fused to the N-terminus of the polypeptide, which is cleaved from the translated polypeptide to produce a “mature” form of the polypeptide.

In some embodiments, the secretory pathway sequence may comprise or consist of a leader sequence (also known as a signal peptide or signal sequence). Leader sequences normally consist of a sequence of 5-30 hydrophobic amino acids, which form a single alpha helix. Secreted proteins and proteins expressed at the cell surface often comprise leader sequences. The leader sequence may be present in the newly-translated polypeptide (e.g. prior to processing to remove the leader sequence). Leader sequences are known for many proteins, and are recorded in databases such as GenBank, UniProt, Swiss-Prot, TrEMBL, Protein Information Resource, Protein Data Bank, Ensembl, and InterPro, and/or can be identified/predicted e.g. using amino acid sequence analysis tools such as SignalP (Petersen et al., 2011 Nature Methods 8: 785-786) or Signal-BLAST (Frank and Sippl, 2008 Bioinformatics 24: 2172-2176).

13

In some embodiments, the secretory pathway sequence is derived from Complement Factor H (FH). In some embodiments, the secretory pathway sequence comprises or consists of SEQ ID NO:7. In some embodiments, the secretory pathway sequence of the polypeptide of the present invention comprises, or consists of, an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to the amino acid sequence of SEQ ID NO:7.

In some embodiments a polypeptide according to the present invention comprises, or consists of, an amino acid sequence corresponding to SEQ ID NO:47, 49, and/or 51. In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:47, 49 and/or 51. Such polypeptides may have any length provided herein.

In some embodiments, a polypeptide according to the present invention may additionally comprise a cleavage site for removing the secretory pathway sequence from the polypeptide. In some embodiments, the cleavage site for removing the secretory pathway sequence from the polypeptide is a cleavage site for an endoprotease. In some embodiments, the cleavage site is for an endoprotease expressed by the cell in which the polypeptide is expressed. In some embodiments, the cleavage site is a signal peptidase cleavage site. In some embodiments, the cleavage site is a protease cleavage site, e.g. a cleavage site for an endoprotease expressed by cells expressing the polypeptide. In some embodiments, the cleavage site is a cleavage site for an endoprotease expressed by RPE cells.

A polypeptide according to the present invention may comprise one or more linker sequences between amino acid sequences. A linker sequence may be provided between any two or more of sequences ‘A’, ‘B’ and/or ‘C’. In some embodiments, a polypeptide according to the present invention comprises, or consists of, an amino acid sequence A+B+C-[LINKER]-A+B+C. In some embodiments, a polypeptide comprises, or consists of, SEQ ID NO:14.

Linker sequences are known to the skilled person, and are described, for example in Chen et al., *Adv Drug Deliv Rev* (2013) 65(10): 1357-1369, which is hereby incorporated by reference in its entirety. In some embodiments, a linker sequence may be a flexible linker sequence. Flexible linker sequences allow for relative movement of the amino acid sequences which are linked by the linker sequence. Flexible linkers are known to the skilled person, and several are identified in Chen et al., *Adv Drug Deliv Rev* (2013) 65(10): 1357-1369. Flexible linker sequences often comprise high proportions of glycine and/or serine residues.

In some embodiments, the linker sequence comprises at least one glycine residue and/or at least one serine residue. In some embodiments the linker sequence consists of glycine and serine residues. In some embodiments, the linker sequence has a length of 1-2, 1-3, 1-4, 1-5, 1-10, 1-15, 1-20, 1-25, 1-30 or 1-35 amino acids.

In some embodiments, a polypeptide according to the present invention comprises a non-amino acid linker. In some embodiments, a polypeptide according to the present invention may comprise two or more polypeptides linked by conjugation, e.g. by nucleophilic substitutions (e.g., reactions of amines and alcohols with acyl halides, active esters), electrophilic substitutions (e.g., enamine reactions) and additions to carbon-carbon and carbon-heteroatom multiple bonds (e.g., Michael reaction, Diels-Alder addition). These

14

and other useful reactions are discussed in, for example, March, *Advanced Organic Chemistry*, 3rd Ed., John Wiley & Sons, New York, 1985; Hermanson, *Bioconjugate Techniques*, Academic Press, San Diego, 1996; and Feeney et al., *Modification of Proteins; Advances in Chemistry Series*, Vol. 198, American Chemical Society, Washington, D.C., 1982.

In some embodiments, a polypeptide according to the present invention comprises a cleavable linker.

It may be desirable for a polypeptide according to the present invention to lack certain properties of CR1. For example, it may be desirable for the polypeptide to lack regions that would otherwise inhibit diffusion through Bruch’s membrane (BrM) or that would interfere with the action of native co-factor family proteins.

A polypeptide according to the present invention lacks the CR1 transmembrane domain (SEQ ID NO:32). A polypeptide according to the present invention may lack the CR1 cytoplasmic tail (SEQ ID NO:33). In a preferred embodiment, a polypeptide according to the present invention is soluble.

A polypeptide according to the present invention may lack regions which could otherwise be exploited by pathogenic bacteria to subvert the host immune system. Bacteria have developed molecules on their surface that can bind and recruit soluble complement factor H from the blood. This enables the bacteria to effectively coat themselves in a complement regulator and evade a host immune response. Polypeptides according to the present invention may lack bacterial binding sites such that they cannot be used by invading pathogens to evade an immune response.

In some embodiments a polypeptide according to the present invention comprises one or more sites for glycosylation. In some embodiments a polypeptide according to the present invention is glycosylated.

In some embodiments, a polypeptide according to the present invention is not glycosylated. In some embodiments, a polypeptide according to the present invention lacks one or more sites for glycosylation. In some embodiments, the polypeptide of the present invention lacks one or more sites for N-linked glycosylation. In some embodiments, a polypeptide according to the present invention lacks N-linked glycans. In some embodiments, a polypeptide according to the present invention is expressed and/or secreted by cells that are unable to glycosylate or fully glycosylate polypeptides. For example, cells may lack functional glycosyl transferase enzymes. In some embodiments, the polypeptide is aglycosyl (i.e. is not glycosylated). In some embodiments, the polypeptide has been deglycosylated, e.g. by treatment with a glycosidase (e.g. Peptide N-Glycosidase). Deglycosylation is preferably non-denaturing. In some embodiments a polypeptide according to the present invention is partially glycosylated, non-glycosylated or de-glycosylated.

In some embodiments, a polypeptide according to the present invention lacks sequence conforming to the consensus sequence of SEQ ID NO:27. In some embodiments, the polypeptide according to the present invention comprises one or more sequences conforming to the consensus sequence of SEQ ID NO:27 that have been mutated to remove sites for N-glycosylation. In some embodiments, the Asn (N) residue in one or more consensus sequences according to SEQ ID NO:27 is substituted with another amino acid residue, e.g. a residue selected from: Ala (A), Cys (C), Asp (D), Glu (E), Phe (F), Gly (G), His (H), Ile (I), Lys (K), Leu (L), Met (M), Pro (P), Gln (Q), Arg (R), Ser (S), Thr (T), Val (V), Trp (W) or Tyr (Y). In some embodiments, the Asn (N) residue in one or more consensus sequences according

to SEQ ID NO:27 is substituted with a Gln (Q) residue. In some embodiments, residue X₂ of SEQ ID NO:27 is, or is mutated to be, an amino acid that is not Ser (S) or Thr (T).

In some embodiments, a polypeptide comprising, or consisting of, an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:2 or SEQ ID NO:4 comprises one or more amino acid substitutions at position 509 and/or position 578 (numbered according to Uniprot: P17927). In some embodiments, the one or more amino acid substitutions are selected from N509Q and/or N578Q. In some embodiments, a polypeptide comprising, or consisting of, an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:3 comprises one or more amino acid substitutions at position 959 and/or position 1028 (numbered according to Uniprot: P17927). In some embodiments, the one or more amino acid substitutions are selected from N959Q and/or N1028Q. In some embodiments, a polypeptide comprising, or consisting of, an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:13 comprises one or more amino acid substitutions at position 509, 578, 959 and/or position 1028 (numbered according to Uniprot: P17927). In some embodiments, the one or more amino acid substitutions are selected from N509Q, N578Q, N959Q and/or N1028Q. Such polypeptides may have any length provided herein.

In some embodiments, the polypeptide comprises, or consists, of an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:15, and/or SEQ ID NO:31. Such polypeptides may have any length provided herein.

In some embodiments a polypeptide according to the present invention comprises a secretory pathway sequence and one or more sequences conforming to the consensus sequence of SEQ ID NO:27 that have been mutated to remove sites for N-glycosylation. In some embodiments a polypeptide according to the present invention comprises, or consists of, an amino acid sequence corresponding to SEQ ID NO:48, 50, 52, 53, and/or 54. In some embodiments the polypeptide comprises, or consists, of an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:48, 50, 52, 53, and/or 54. Such polypeptides may have any length provided herein.

In some embodiments, a polypeptide according to the present invention may comprise amino acid sequence(s) to facilitate expression, folding, trafficking, processing, purification or detection of the polypeptide. For example, the polypeptide may comprise a sequence encoding a protein tag, e.g. a His, (e.g. 6×His), FLAG, Myc, GST, MBP, HA, E, or Biotin tag, optionally: at the N- or C-terminus of the polypeptide; in a linker; or at the N- or C-terminus of a linker. In some embodiments the polypeptide comprises a detectable moiety, e.g. a fluorescent, luminescent, immuno-detectable, radio, chemical, nucleic acid or enzymatic label. In some embodiments, the detectable moiety facilitates detection of the polypeptide in a sample obtained from a subject, e.g. following administration to the subject of the polypeptide, nucleic acid, vector, cell or pharmaceutical

composition according to the present invention. The sample may be any biological sample obtained from a subject. In some embodiments the sample is a liquid biopsy, such as ocular fluid (tear fluid, aqueous humour, or vitreous), blood, plasma, etc. In some embodiments the sample is a cytological sample or a tissue sample such as a surgical sample, e.g. of ocular cells/tissue.

In some embodiments, the polypeptide according to the present invention may be detected and/or distinguished from endogenous CR1 by Western blotting, mass spectrometry and/or enzyme digestion, e.g. by a specific peptidase. In some embodiments, the polypeptide may comprise a point mutation to generate peptides by enzyme digestion that are distinct from post-digestion peptides from endogenous CR1.

In some embodiments, a polypeptide according to the present invention may additionally comprise a cleavage site for removing a protein tag. For example, it may be desired to remove a tag used for purification of the polypeptide following purification. In some embodiments the cleavage site may e.g. be a Tobacco Etch Virus (TEV) protease cleavage site, for example as shown in SEQ ID NO:34.

In some embodiments a polypeptide according to the present invention comprises, or consists of, an amino acid sequence corresponding to SEQ ID NO:40, 42, 44, and/or 46. In some embodiments the polypeptide comprises, or consists of, an amino acid sequence having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:40, 42, 44, and/or 46. Such polypeptides may have any length provided herein.

As used herein, a “polypeptide” includes molecules comprising more than one polypeptide chain, which may be associated (e.g. covalently or non-covalently) into a complex. That is, a “polypeptide” within the meaning of the present invention encompasses molecules comprising one or more polypeptide chains. The polypeptide of the invention may in various different embodiments and at different stages of expression/production in vitro or in vivo comprise e.g. a signal peptide, protein tag, cleavage sites for removal thereof, etc. The polypeptide of the invention may comprise any CR1 CCP sequence described herein, or any combination of CR1 CCP domains 8 (SEQ ID NO:8), 9 (SEQ ID NO:9), 10 (SEQ ID NO:10), 15 (SEQ ID NO:8), 16 (SEQ ID NO:11), and/or 17 (SEQ ID NO:12), or any combination of sequences ‘A’, ‘B’ and/or ‘C’ described herein, optionally in combination with one or more of any of the further features of the polypeptide of the invention described herein (e.g. signal peptide, linker, detection sequence, lack of glycosylation site, substituted amino acid residue, protein tag, cleavage site for removing a protein tag, secretory pathway sequence, cleavage site for removing a secretory pathway sequence).

Sequence Identity

As used herein, an amino acid sequence which corresponds to a reference amino acid sequence may comprise at least 60%, e.g. one of at least 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to the reference sequence.

Pairwise and multiple sequence alignment for the purposes of determining percent identity between two or more amino acid or nucleic acid sequences can be achieved in various ways known to a person of skill in the art, for instance, using publicly available computer software such as ClustalOmega (Söding, J. 2005, Bioinformatics 21, 951-960), T-coffee (Notredame et al. 2000, J. Mol. Biol. (2000) 302, 205-217), Kalign (Lassmann and Sonnhammer 2005,

US 12,383,600 B2

17

BMC Bioinformatics, 6(298)) and MAFFT (Katoh and Standley 2013, Molecular Biology and Evolution, 30(4) 772-780 software. When using such software, the default

18

parameters, e.g. for gap penalty and extension penalty, are preferably used.
Sequences

Sequences		
SEQ ID NO:	Description	Sequence
1	Human Complement Receptor 1 (UniProt: P17927; Entry version 181 (25 Oct. 2017), Sequence version 3 (02 Mar. 2010)); residues 1-2039 Including signal sequence [CCP domains '8-10' and 15-17 indicated by single underline, individual CCP domains 8, 9, 10, 15, 16, 17 indicated by double underline; amino acid differences between CCPs 8-10 and 15-17 indicated with wavy underline]	MGASSPRSPRPEVGPPAPGLPFCGGSLLAVVLLALPVAWGQCNAPEW LPFARPTNLTEDEFPIGTYLNIECRPGYSGRPFSIICLKNSVWTGAK DRCRCKSCRNPPDPVNGMVH1KGIQFGSQIKYSCTKGRLIGSSSAT CIIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRRENFHGSVVT YRCNPNGSGGRKVFELVGEPSIYCTSNDQVGIVWSPAPQCIIIPNKCTP PNVENGILVSDNRSLSLNEVVEFRCQPGFMKGPRRVKCQALNKWEP ELPSCSRVCQPPPDLVHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAA SMRCPQGDWSPAAPTCEVKSCDDFMGOLLNGRVLFPVNLOLGAKVDF VCDEGFLQKGSACVLAGMESLNNSVPVCEQIFCPSPPVIPNGRH TGKPLEVFFGKTVNTCDPHDRGTSFDLIGESTIRTSDPQNGVW SSPAPRCGILGHQAPDHFLFAKLKTQTNASDFPIGTSKLYCRPEYY GRPFISITCLDNLVWSSPKDVCRKSCKTPPDVNGMVH1T1QVGSR INYSCTTGHLIGHSSAECILSGNAAHWSTKPPICQRCGLPPTIAN GDFISTNRRENFHGSVVTYRCN1GSGGRKVVELVGEPSIYCTSNDQV GINSGPAPQCIIPNKCTPPNVENGILVSDNRSLSLNEVVEFRCQPGF VMKGRPRVKCQALNKWEPELPSCSRVCQPPPDLVHAERTQRDKDNFSP QGEVFSCEPGYDLRGAAACMRCTPQGDWSPAAPTCEVKSCDDFMGQLL NGRVLFPVNLOLGAKVDFVCDEGFLQKGSACVLAGMESLNNSVP VCEQIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTCDPHDRGTSFDL IGESTIRTSDPQNGVWSSPAPRCGILGHQAPDHFLFAKLKTQTNAS SDFFIGTSLKYECPREYYGRPFISITCLDNLVWSSPKDVCRKSCKTP DPVNGMVH1T1QVGSRINYSCTTGHLIGHSSAECILSGNAHWST KPPICQRCGLPPTIANDFISTNRRENFHGSVVTYRCN1GSGGRKV FELVGEPSIYCTSNDQVGIVWSPAPQCIIIPNKCTPPNVENGILVSDN RSLFLSNEVVEFRCQPGFMKGPRRVKCQALNKWEPELPSCSRVCQPP PEILHGEHTPSHQDNFSPGQEVFYSCEPGYDLRGAASLHCTPQGDWSP EAPRCAVSKCDDFLQGLPHGRVLFPNLQLGAKVSFVCDLEGFRKGSS VSHCVLVGMRSLWNNSVPCEHIFCPNPPA1LNGRHTGTPSGDIYVGK EISYTCDPHDRGTMFTNLIGESTIRTSDPHNGMVWSAPPCELSVR AGHCKTPEQFPFFASPTIPINDFEPPVGTSLNVECRPGYFGKMSISCL ENLVWSSVEDNCRRKSCGPPEPFNGMVHINTDTQFGSTVINYSCNEGF RLIGSPSTTCLVSGNNVTWDKKAPICEIIISCEPPPTISNGDFYSSNRT SFHNGTVVTVQCHTGPDGQELFELVGEPSIYCTSNDQVGIVWSSPPR C1STNKCTAPEVENAIRVPGNRSFFSLSITEIIRRCQPGFVWVGSHTVQ CQTNGRWGPKLPHCSRVQPPPEI1LGEHTLHSQDNFSPGQEVFYSCE PSYDLRGAASLHCTPQGDWSPAPRCTVKSCKDFLGQLPHGRVLLPLN LQLGAKVSFVCDGFRKLGRSASHCVLAGMKALWNSSVPVCEQIFCPN PPA1LNGRHTGTPFGDIPYGKEISYACDTHPDRGTMFTNLIGESSIRCT SDPQNGVWSSPAPRCELSVPAACPHPPKIQNQHGY1GGHVSLYLPGMT ISYICDCPGYLVLKGKFIFCTDQGIWSQLDHYCKEVNCSPFLFMNGISK ELEMKKVYHYGDVVTLCEDGYTLEGSFWSCQADDRWDPPFLAKCTSR THDALIVGTLSEGTIFFILLI1FLSWIILKHRKGNNAHENPEKVAIHLH SQGSSVHPRTLQTNNEENSRLV 2 Human Complement Receptor 1 CCPs 8-10 (UniProt: P17927 residues 491 to 684) Without leader sequence 3 Human Complement Receptor 1 CCPs 15-17 (UniProt: P17927 residues 941 to 1134) Without leader sequence 4 Human Complement Receptor 1; consensus sequence for CCPs 8-10, 15-17 (UniProt: P17927 residues 491 to 684; residues 941 to 1134) Without leader sequence

- continued

Sequences		
SEQ ID NO:	Description	Sequence
5	Human Complement Receptor 1 CCPs 8-10 (UniProt: P17927 residues 491 to 684) Non-glycosylated; Without leader sequence [mutated glycosylation sites underlined; substitutions in bold correspond to positions 509 and 578 of UniProt: P17927]	GHCQAPDHFLFAKLKTQT<u>Q</u>ASDFPIGTS<u>L</u>KYE<u>C</u>RPEYYGRPF<u>S</u>ITC<u>L</u>D<u>N</u>L<u>V</u>W<u>S</u>SP<u>K</u>D<u>V</u>C<u>K</u>R<u>S</u>C<u>K</u>T<u>P</u>DP<u>V</u>NG<u>M</u>V<u>H</u>IT<u>D</u>I<u>Q</u>V<u>G</u>S<u>R</u>I<u><u>Q</u>Y</u>S<u>C</u>T<u>I</u>G<u>H</u>R<u>L</u>I<u>G</u>H<u>S</u>A<u>E</u>C<u>I</u>L<u>S</u>G<u>M</u>A<u>H</u>W<u>S</u>T<u>K</u>P<u>P</u>I<u>C</u>Q<u>R</u>I<u>P</u>C<u>G</u>L<u>P</u>T<u>I</u>ANG<u>D</u>F<u>I</u>S<u>T</u>N<u>R</u>E<u>N</u>F<u>H</u>Y<u>G</u>S<u>V</u>V<u>T</u>Y<u>R</u>C<u>N</u>P<u>G</u>S<u>G</u>R<u>K</u>V<u>F</u>E<u>L</u>V<u>G</u>E<u>P</u>S<u>I</u>Y<u>C</u>T<u>S</u>N<u>D</u>D<u>Q</u>V<u>G</u>I<u>W</u>S<u>G</u>P<u>A</u>P<u>Q</u>C<u>II</u>
6	Human Complement Receptor 1 CCPs 15-17 (UniProt: P17927 residues 941 to 1134) Non-glycosylated; Without leader sequence [mutated glycosylation sites underlined; substitutions in bold correspond to positions 959 and 1028 of UniProt: P17927]	GHCQAPDHFLFAKLKTQT<u>Q</u>ASDFPIGTS<u>L</u>KYE<u>C</u>RPEYYGRPF<u>S</u>ITC<u>L</u>D<u>N</u>L<u>V</u>W<u>S</u>SP<u>K</u>D<u>V</u>C<u>K</u>R<u>S</u>C<u>K</u>T<u>P</u>DP<u>V</u>NG<u>M</u>V<u>H</u>IT<u>D</u>I<u>Q</u>V<u>G</u>S<u>R</u>I<u><u>Q</u>Y</u>S<u>C</u>T<u>I</u>G<u>H</u>R<u>L</u>I<u>G</u>H<u>S</u>A<u>E</u>C<u>I</u>L<u>S</u>G<u>M</u>T<u>A</u>H<u>W</u>S<u>T</u>K<u>P</u>P<u>I</u>C<u>Q</u>R<u>I</u>P<u>C</u>G<u>L</u>P<u>T</u>I<u>ANG</u>D<u>F</u>I<u>S</u>T<u>N</u>R<u>E</u>N<u>F</u>H<u>Y</u>G<u>S</u>V<u>V</u>T<u>Y</u>R<u>C</u>N<u>L</u>G<u>S</u>R<u>G</u>R<u>K</u>V<u>F</u>E<u>L</u>V<u>G</u>E<u>P</u>S<u>I</u>Y<u>C</u>T<u>S</u>N<u>D</u>D<u>Q</u>V<u>G</u>I<u>W</u>S<u>G</u>P<u>A</u>P<u>Q</u>C<u>II</u>
7	Native signal/leader sequence from Factor H	MRLLA <u>K</u> I <u>I</u> C <u>L</u> M <u>L</u> WA <u>I</u> C <u>V</u> A
8	Human Complement Receptor 1 CCP8/CCP15 (UniProt: P17927 residues 491-550/941-1000) 'Sequence A'	GHCQAPDHFLFAKLKTQT<u>N</u>ASDFPIGTS<u>L</u>KYE<u>C</u>RPEYYGRPF<u>S</u>ITC<u>L</u>D<u>N</u>L<u>V</u>W<u>S</u>SP<u>K</u>D<u>V</u>C<u>K</u>R
9	Human Complement Receptor 1 CCP9 (UniProt: P17927 residues 552-612)	K<u>S</u>C<u>K</u>T<u>P</u>DP<u>V</u>NG<u>M</u>V<u>H</u>IT<u>D</u>I<u>Q</u>V<u>G</u>S<u>R</u>I<u>N</u>Y<u>S</u>C<u>T</u>I<u>G</u>H<u>R</u>L<u>I</u>G<u>H</u>S<u>A</u>E<u>C</u>I<u>L</u>S<u>G</u> <u>N</u>A<u>H</u>W<u>S</u>T<u>K</u>P<u>P</u>I<u>C</u>Q
10	Human Complement Receptor 1 CCP10 (UniProt: P17927 residues 613-684)	R<u>I</u>P<u>C</u>G<u>L</u>P<u>T</u>I<u>ANG</u>D<u>F</u>I<u>S</u>T<u>N</u>R<u>E</u>N<u>F</u>H<u>Y</u>G<u>S</u>V<u>V</u>T<u>Y</u>R<u>C</u>N<u>R</u><u>G</u><u>S</u><u>G</u>R<u>K</u>V<u>F</u>E<u>L</u>V<u>G</u>E<u>P</u>S<u>I</u>Y<u>C</u>T<u>S</u>N<u>D</u>D<u>Q</u>V<u>G</u>I<u>W</u>S<u>G</u>P<u>A</u>P<u>Q</u>C<u>II</u>
11	Human Complement Receptor 1 CCP16 (UniProt: P17927 residues 1002-1062)	K<u>S</u>C<u>K</u>T<u>P</u>DP<u>V</u>NG<u>M</u>V<u>H</u>IT<u>D</u>I<u>Q</u>V<u>G</u>S<u>R</u>I<u>N</u>Y<u>S</u>C<u>T</u>I<u>G</u>H<u>R</u>L<u>I</u>G<u>H</u>S<u>A</u>E<u>C</u>I<u>L</u>S<u>G</u> <u>N</u>T<u>A</u><u>H</u>W<u>S</u>T<u>K</u>P<u>P</u>I<u>C</u>Q
12	Human Complement Receptor 1 CCP17 (UniProt: P17927 residues 1063-1134)	R<u>I</u>P<u>C</u>G<u>L</u>P<u>T</u>I<u>ANG</u>D<u>F</u>I<u>S</u>T<u>N</u>R<u>E</u>N<u>F</u>H<u>Y</u>G<u>S</u>V<u>V</u>T<u>Y</u>R<u>C</u>N<u>L</u><u>G</u><u>S</u><u>R</u><u>G</u>R<u>K</u>V<u>F</u>E<u>L</u>V<u>G</u>E<u>P</u>S<u>I</u>Y<u>C</u>T<u>S</u>N<u>D</u>D<u>Q</u>V<u>G</u>I<u>W</u>S<u>G</u>P<u>A</u>P<u>Q</u>C<u>II</u>
13	Human Complement Receptor 1 CCPs 8-10 and 15-17 (contiguous; without leader sequence) [amino acid differences between CCPs 8-10 and 15-17 indicated with wavy underline]	GHCQAPDHFLFAKLKTQT<u>N</u>ASDFPIGTS<u>L</u>KYE<u>C</u>RPEYYGRPF<u>S</u>ITC<u>L</u>D<u>N</u>L<u>V</u>W<u>S</u>SP<u>K</u>D<u>V</u>C<u>K</u>R<u>S</u>C<u>K</u>T<u>P</u>DP<u>V</u>NG<u>M</u>V<u>H</u>IT<u>D</u>I<u>Q</u>V<u>G</u>S<u>R</u>I<u>N</u>Y<u>S</u>C<u>T</u>I<u>G</u>H<u>R</u>L<u>I</u>G<u>H</u>S<u>A</u>E<u>C</u>I<u>L</u>S<u>G</u> <u>N</u>A<u>H</u>W<u>S</u>T<u>K</u>P<u>P</u>I<u>C</u>Q<u>R</u>I<u>P</u>C<u>G</u>L<u>P</u>T<u>I</u>ANG<u>D</u>F<u>I</u>S<u>T</u>N<u>R</u>E<u>N</u>F<u>H</u>Y<u>G</u>S<u>V</u>V<u>T</u>Y<u>R</u>C<u>N</u><u>L</u><u>G</u><u>S</u><u>R</u><u>G</u>R<u>K</u>V<u>F</u>E<u>L</u>V<u>G</u>E<u>P</u>S<u>I</u>Y<u>C</u>T<u>S</u>N<u>D</u>D<u>Q</u>V<u>G</u>I<u>W</u>S<u>G</u>P<u>A</u>P<u>Q</u>C<u>II</u>

- continued

Sequences		
SEQ ID NO:	Description	Sequence
14	Human Complement Receptor 1 CCPs 8-10 (linker) 15-17	<u>GHCQAPDHFLFAKLKTQTNASDFPIGTSKYECRPEYYGRPFSITCLD</u> <u>NLVWSSPKDVCKRKSCKIPPDVPNMGMVHITDIQVGSRINYSCТИGHР</u> <u>LIGHSSAECILSGNAAHWSTKPPICQRIPCGLPPTIANGDFISTNREN</u> <u>FHYGSVVTYRCNPNGSGGRKVFELVGEPSIYCTSNDQVGIWSGPAPQC</u> II-(linker)- <u>GHCQAPDHFLFAKLKTQTNASDFPIGTSKYECRPEYYGRPFSITCLD</u> <u>NLVWSSPKDVCKRKSCKIPPDVPNMGMVHITDIQVGSRINYSCТИGHР</u> <u>LIGHSSAECILSGNTAHWSTKPPICQRIPCGLPPTIANGDFISTNREN</u> <u>FHYGSVVTYRCNLGSRGKVFELVGEPSIYCTSNDQVGIWSGPAPQC</u> II
15	Human Complement Receptor 1 CCPs 8-10 and 15-17 (contiguous) Non-glycosylated; Without leader sequence [Mutated glycosylation sites underlined; substitutions in bold correspond to positions 509, 578, 959 and 1028 of UniProt: P17927]	<u>GHCQAPDHFLFAKLKTQT<u><u>QASDEPIGTSKYECRPEYYGRPFSITCLD</u></u></u> <u>NLVWSSPKDVCKRKSCKIPPDVPNMGMVHITDIQVGSR<u><u>IQY</u></u>SCТИGHР</u> <u>LIGHSSAECILSGNAAHWSTKPPICQRIPCGLPPTIANGDFISTNREN</u> <u>FHYGSVVTYRCNPNGSGGRKVFELVGEPSIYCTSNDQVGIWSGPAPQC</u> II <u>GHCQAPDHFLFAKLKTQT<u><u>QASDEPIGTSKYECRPEYYGRPFSITCLD</u></u></u> <u>NLVWSSPKDVCKRKSCKIPPDVPNMGMVHITDIQVGSR<u><u>IQY</u></u>SCТИGHР</u> <u>LIGHSSAECILSGNTAHWSTKPPICQRIPCGLPPTIANGDFISTNREN</u> <u>ENFHYGGSVVTYRCNLGSRGKVFELVGEPSIYCTSNDQVGIWSGPAPQCII</u>
16	Human Complement Receptor 1; Consensus sequence for CCPs 9, 16 (corresponding to UniProt: P17927 residues 552-612 and 1002-1062) 'Sequence B'	<u>KSCKIIPPDVPNMGMVHITDIQVGSRINYSCТИGHRLIGHSSAECILSG</u> <u>NX,AHWSTKPPICQ</u>
17	Human Complement Receptor 1; Consensus sequence for CCPs 10, 17 (corresponding to (UniProt: P17927 residues 613-684 and 1063-1134) 'Sequence C'	<u>RIPCGLPPTIANGDFISTNRENHYGSVVTYRCNX<u><u>GSX</u></u>GRKVFELVG</u> <u>EPSIYCTSNDQVGIWSGPAPQCII</u>
18	Human C3 (UniProt: P01024; Entry version 221 (20 Dec. 2017); Sequence version 2 (12 Dec. 2006)) including signal peptide	MGPTSGPSLLLLLTHPLALGSPMYSITPNILRLESEETMVLEAHDAQGDVPVIVVHDPGKKLVLSSAKTTLIPAINHMGNVIFTIPANREFKSEGRNKPFVTVQATFGTQVVEKVVLSLQSGYLFIQTDKTIYIPGSTVLYRIFTVNHKLLPVGRIVMVNIEPNPEGIPVKQDSLSSQNQLGVLPPLSDWIPELVNMGQWKIRAYENSPQQVSETEFEVKEYVLPSEVIVEPEKFYIYNEKGLEVITARFLYGGKVEGTAFVIFGIDQDGEQRISLPESLKRIPEDGSGEVVLSRKVLLDGVQNRAEDLVGKSLYVSATVILHSGSDMVAERSGIPIVISPYQIHFIKTPKYFKPGMPFDLMVFVINPDGSPA YRVPVAVQGEDTVQSLIQGDGVAKLSSINTHPSQKPLSITVRIKKQELS EAEQATRMTQALPYSTVGNSNNYHLHSVLRLTELPGETLNVNFLLRMD RAHEAKIRYYTYLIMNKGRLLKAGRQVREPQGDLLVVLPLSITTDIIPS FRLVAYYTLLIGASQOREVVAADVSVWVDVKDSCVGSLLVVKSGQSEDRQPV PGQQMTLKIEGDHGARVVLVADKGTVFVNLNNKNTQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSSGQTAQRAELQCPQPAARRRS VOLTEKRMDKVGYPKELKCCEDGMRENPMRFSQCRRTRFISLGECAKKVLFDCCNYITELRROHARASHLGLARSNLDEDIIABENIVSRSEFP ESWLNVEDLKEPPKNGIISTKLMNIFLKDSDTTWEILAVSMSDKKGIC VADPEPVTVMQDFIIDLRLPYSVRNEQVEIRAVLYNVRQNJELKVRV ELLHNPAFCSLATTKRRHQQTVTIPPKSSLSPVYVIVPLKTLQVEEV KAAYVHHFISDGVRKSLKVPEGIRMNKTVAVRTLDPERLGLREGVQKE DIPPADLSDQVPDTESETRILLQGTPVQAQMTEDAVDAERLKHLIVTPSGCGEONMIGMPTVIAVHYLDETEQWEKFGLEKRGQALELIKGYTQQ LAFRPQSSAFAAFVKRAPSTWLTYVVKVFSLAVNLTAIDSQVLCGAV KWLLKEKQRPDGVEQEDAPVIHQEMIGGLRNNNNEKDMALTAFVLISLQEAKDICEEQVNSLPGSITKAGDFLEANYMNLQRSYTVAIAGYALAQMGRLKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSYALLLQLKDFDVFPPVWRWLNEQRYYGGGYGSTQATFMVFPQALAQYQKDAPDHQELNLDV SLQLPSRSSKITHRIHWESASLLRSEETKENEGFTVTAEGKGQGTLCSVVTMYHAKAKDQLTCNKFDLKVTIKAPAPETEKRPQDAKNTMILEICTRYRGDQDATMSILDISMMTGAPDTDDLKQLANGVDRYISKYELDKAFSD RNTLIIYLDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEES

- continued

<u>Sequences</u>		
SEQ ID NO:	Description	Sequence
		CTRFYHPEKEDGKLNLKLCRDELCRCAEENCFIQLKSDDKVTLEERLDKA CEPGVDVYVKTRLVKVQLSNDFDEYIMAIEQTIKSGSDEVQVGQORTF ISPIKCREALKLEEKHHYLMWGLSSDFWGKEPNLNSYIIGKDTWVEHWP EEDECQDEENQKQCQDLSGAFTESMVMVFGCPN
19	Human C3 β chain (UniProt: P01024; Entry version 221 (20 Dec. 2017); Sequence version 2 (12 Dec. 2006); residues 23-667)	SPMYSIITPNILRLESEETMVLEAHDAQGDVPVTVHDFPGKKLVLS SEKTVLTPATNHMGNTFTIPANREFKSEKGRNKFVTQATFGTQVVE KVVLSLQSGYLFQTDKTIYTPGSTVLYRIFTVNHKLLPVGRTVMVN IENPEGIPVKQDSLSSQNQLGVPLSLWDIPELVNMQWKIRAYYENSP QQVFSTEFEVKEYVLPSPFEVIVEPTEKFYYIYNEKGLEVTITARFLY KKVGETAVFIFGIQDGEQRISLPSKRIPIEDGSGEVVLSRKVLLDG VQNPRAEALVGKSLYVSATVILHSGSDMVAERSGIPIVTSPYQIHFT KTPYFKPGMPFDLMVFVTPDGSPPAYRVPVAVQGEDTVQSLTQGDGV AKLSINTHPSQKPLSITVRTKQELSEAEQATRTMQALPYSTVGNNSNN YLHLSVLRLTRPGETLNVNFLLRMDRAHEAKI RYYTYLIMNKGRLLK AGROVREPGQDLVVLPLSITTDPIPSFLVAYYTLIGASGQREVVADS VVWDVKDSCVGSLLVKGQSEDRCVPGQQMTLKEGDHGARVVLVAV DKGVFVLNKKNLITQSKIWVVEKADIGCTPGSGKDYAGVFSAGLTF TSSSGQTAQRAELQCPQPA
20	Human C3 α' chain (UniProt: P01024; Entry version 221 (20 Dec. 2017); Sequence version 2 (12 Dec. 2006); residues 749-1663)	SNLDEDIIAEENIVSRSEFPESWLWNVEDLKEPPKNGISTKLMNIFLK DSITTWEILAVSMSDKKGICVADPFEVTVMQDFFIDLRLPYSVVRNEQ VEIRAVLYNYRQNQELKVRVELLHNPAFCSLATTKRRHQQTVTIPPKS SLSVPYVIVPLKTGLQEVEVKAAVYHHPISDGVRKSLKVVPREGIRMNK TVAVRTLDPERLGREGVQKEDIPPADLSDQVDPTESETRILLQGTPVA QMTEDAVDAERLKHЛИVTPSGCQEONMIGMTPTVIAVHYLDETEQWEK FGLERQGALELIKGGYTQQLAFRPSSAFAAFVKRAPSTWLTVYV VFSLAVNLIAIDSQVLCAGAVKWLILEKQKPDGVFQEDAPVIHQEMIGG LRNNNEKDMALTAFLVISLQEAKDICEEQVNSLPGSITKAGDFLEANY MNLQRSYTVAIAGYALAQMGRLKGPPLNKFLTTAKDKNRWEDPGKQLY NVEATSYALLLQLKDFDFFVPPVVRWLNEQRYYGGGGSTQATFMVF QALAQYQKDAPDHQELNLDVSLQLPSR KENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFDLKVTIKP TEKRPQDAKNTMILEICTRYRGDQDATMSILDISMNTGFAPDPTD DLKQLANGVDRYISKYELDKAFSDRNTLIIYLDKVSHSEDDCLAFKV QYFNVELIQPGAVKVVAYYNLEESCTRHYHPEKEDGKLNLCRDEL CAEENCFIQLKSDDKVTLEERLDKACEPGVDVYKTRLVKVQLSND YIMAIEQTIKSGSDEVQVGQORTFISPIKCREALKLEEKHHYLMWGL SDFWGKEPNLNSYIIGKDTWVEHWP EEDECQDEENQKQCQDLSGAFTE SMVMVFGCPN
21	Human C3a (UniProt: P01024; Entry version 221 (20 Dec. 2017); Sequence version 2 (12 Dec. 2006); residues 672-748)	SVQLTEKRMKVGKYPKELRKCCEDGMRENPMRFSQRTRFISLGAE CKKVFLDCCNVITELRRQHARASHLGLAR
22	Human C3 α' chain fragment 1 (UniProt: P01024; Entry version 221 (20 Dec. 2017); Sequence version 2 (12 Dec. 2006); residues 749-1303)	SNLDEDIIAEENIVSRSEFPESWLWNVEDLKEPPKNGISTKLMNIFLK DSITTWEILAVSMSDKKGICVADPFEVTVMQDFFIDLRLPYSVVRNEQ VEIRAVLYNYRQNQELKVRVELLHNPAFCSLATTKRRHQQTVTIPPKS SLSVPYVIVPLKTGLQEVEVKAAVYHHPISDGVRKSLKVVPREGIRMNK TVAVRTLDPERLGREGVQKEDIPPADLSDQVDPTESETRILLQGTPVA QMTEDAVDAERLKHЛИVTPSGCQEONMIGMTPTVIAVHYLDETEQWEK FGLERQGALELIKGGYTQQLAFRPSSAFAAFVKRAPSTWLTVYV VFSLAVNLIAIDSQVLCAGAVKWLILEKQKPDGVFQEDAPVIHQEMIGG LRNNNEKDMALTAFLVISLQEAKDICEEQVNSLPGSITKAGDFLEANY MNLQRSYTVAIAGYALAQMGRLKGPPLNKFLTTAKDKNRWEDPGKQLY NVEATSYALLLQLKDFDFFVPPVVRWLNEQRYYGGGGSTQATFMVF QALAQYQKDAPDHQELNLDVSLQLPSR SEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFDLKVTIK PAPETEKRPQDAKNTMILEICTRYRGDQDATMSILDISMNTGFAPD DLKQLANGVDRYISKYELDKAFSDRNTLIIYLDKVSHSEDDCLAFKV QYFNVELIQPGAVKVVAYYNLEESCTRHYHPEKEDGKLNLCRDEL CAEENCFIQLKSDDKVTLEERLDKACEPGVDVYKTRLVKVQLSND YIMAIEQTIKSGSDEVQVGQORTFISPIKCREALKLEEKHHYLMWGL SDFWGKEPNLNSYIIGKDTWVEHWP EEDECQDEENQKQCQDLSGAFTE SMVMVFGCPN
23	Human C3 α' chain fragment 2 (UniProt: P01024;; Entry version 221 (20 Dec. 2017); Sequence version 2 (12 Dec. 2006); residues 1321-1663) Also known as C3c fragment 2	SEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFDLKVTIK PAPETEKRPQDAKNTMILEICTRYRGDQDATMSILDISMNTGFAPD DLKQLANGVDRYISKYELDKAFSDRNTLIIYLDKVSHSEDDCLAFKV QYFNVELIQPGAVKVVAYYNLEESCTRHYHPEKEDGKLNLCRDEL CAEENCFIQLKSDDKVTLEERLDKACEPGVDVYKTRLVKVQLSND YIMAIEQTIKSGSDEVQVGQORTFISPIKCREALKLEEKHHYLMWGL SDFWGKEPNLNSYIIGKDTWVEHWP EEDECQDEENQKQCQDLSGAFTE SMVMVFGCPN

- continued

Sequences		
SEQ ID NO:	Description	Sequence
24	Human C3f (UniProt: P01024; ; Entry version 221 (20 Dec. 2017); Sequence version 2 (12 Dec. 2006); residues 1304-1320)	SSKITHRIHWESASLLR
25	Human Complement Factor I (UniProt: P05156; ; Entry version 192 (20 Dec. 2017); Sequence version 2 (11 Jan. 2011))	MKLLHVFLFLCFHLRFCKVTYTSQEDLVEKKCLAKKYTHLSCDKVFCQPWRQCEGTCVKLPYQCPCPKNGTAVCATNRRSFTPTCQQKSLECLHPGTFKLNNGTCTAEGKFSVSLKHGNNTDSEGIVEVKLVQDKTMFICKSSWSMREANACLDLGFFQQGADTQRFFKFLSDLSINSTECVLHVHCRGLETSLAECTFTKRRTMGYQDFADVVVCYTQKADSPMDDFFQCVNGKYISQMKACDGINDCGDQSDELCCKAQGKGFHCKSGVCIPSQYQCNGEVDCTITGDEVGCAGFASVTQEETEILTADMDAERRRIKSLLPKLSCGVKNRMHIRRKRVGGKRAQLGLPWPQVAIKDASGTCGGIYIGGCWILTAAHCLRA SKTHRQIWTWVWDWIHPDLKRIVIEYDRIIFHENYNAGTYQNDIALIEMKKDGDKKDCELPRSIACVPWSPYLFPQPNDTCIVSGWGREGKDNERVFSVFLQWGEVKLISNCSKFYGNRFYEKEMECAGTYDGSIDACKGDSGGPLVC LVCMDANNVTVVWGVVSWGENCGKPEFPGVYTKVANYFDWISYHVGRP FISQYNV
26	Human Complement Factor I proteolytic domain (UniProt: P05156; ; Entry version 192 (20 Dec. 2017); Sequence version 2 (11 Jan. 2011); residues 340-574)	IVGGKRAQLGLDPWPQVAIKDASGTCGGIYIGGCWILTAAHCLRASKTHRYQIWTWVWDWIHPDLKRIVIEYDRIIFHENYNAGTYQNDIALIEMKKDGDKKDCELPRSIACVPWSPYLFPQPNDTCIVSGWGREGKDNERVFSLQWGEVKLISNCSKFYGNRFYEKEMECAGTYDGSIDACKGDSGGPLVC MDANNVTVVWGVVSWGENCGKPEFPGVYTKVANYFDWISYHV
27	Consensus sequence for N-linked glycosylation	NX ₁ X ₂ wherein X ₁ = any amino acid except for P X ₂ = S or T
28	Complement Factor H isoform FHL-1 (UniProt: P08603-2) Including leader sequence (underlined) SEQ ID No. 7	<u>MRLLA</u> KII <u>CML</u> WAICVAEDCNE <u>LPPRNTE</u> ILTGSWS <u>DQTYPEGTQA</u> TYKCRPGYRSLGNVIMVCRKG <u>EVALNPLRK</u> QKR <u>PCGH</u> P <u>GDTPF</u> GTFTLTGGNVFEYGVKAVYTCNEGYQLLG <u>EIN</u> YREC <u>DTDGWTNDI</u> P <u>ICEVV</u> KCL <u>PV</u> TAP <u>ENGKIVSSAMEPDREY</u> HFGQAVRFVCNSGYKIE <u>GED</u> EMHC <u>SDDGFW</u> SK <u>EPKPKC</u> VE <u>ISCKSPDV</u> INGS <u>PISQK</u> IYK <u>ENERF</u> QYKCNMG <u>YEYSERGDAVCTESGR</u> PLPS <u>CEEKCDNPY</u> IPNGDYSPLRI <u>KHRTGD</u> EITY <u>QCRNGFYPATRGNTAKCTSTGW</u> IP <u>APRCTLKPCD</u> YP <u>DIKHGGLY</u> HENMRPY <u>PFPVAVGKYYSSY</u> CD <u>EHFETPSGSY</u> WD <u>HICTQDGWS</u> PAVP <u>CLR</u> KCY <u>FPL</u> EN <u>GYQNQYGRKFVQGK</u> SID <u>VACHPGYALPK</u> A <u>QTTVTCM</u> ENGWS <u>PTPRCIRV</u> S <u>FTL</u>
29	Human Complement Factor H (UniProt: P08603; ; Entry version 214 (25 Oct. 2017); Sequence version 4 (11 Sep. 2007)) Including leader sequence (underlined) SEQ ID No. 7	<u>MRLLA</u> KII <u>CML</u> WAICVAEDCNE <u>LPPRNTE</u> ILTGSWS <u>DQTYPEGTQA</u> TYKCRPGYRSLGNVIMVCRKG <u>EVALNPLRK</u> QKR <u>PCGH</u> P <u>GDTPF</u> GTFTLTGGNVFEYGVKAVYTCNEGYQLLG <u>EIN</u> YREC <u>DTDGWTNDI</u> P <u>ICEVV</u> KCL <u>PV</u> TAP <u>ENGKIVSSAMEPDREY</u> HFGQAVRFVCNSGYKIE <u>GED</u> EMHC <u>SDDGFW</u> SK <u>EPKPKC</u> VE <u>ISCKSPDV</u> INGS <u>PISQK</u> IYK <u>ENERF</u> QYKCNMG <u>YEYSERGDAVCTESGR</u> PLPS <u>CEEKCDNPY</u> IPNGDYSPLRI <u>KHRTGD</u> EITY <u>QCRNGFYPATRGNTAKCTSTGW</u> IP <u>APRCTLKPCD</u> YP <u>DIKHGGLY</u> HENMRPY <u>PFPVAVGKYYSSY</u> CD <u>EHFETPSGSY</u> WD <u>HICTQDGWS</u> PAVP <u>CLR</u> KCY <u>FPL</u> EN <u>GYQNQYGRKFVQGK</u> SID <u>VACHPGYALPK</u> A <u>QTTVTCM</u> ENGWS <u>PTPRCIRV</u> S <u>FTL</u> YVTADGET <u>GS</u> GT <u>CGKDGWSAQ</u> PT <u>CIK</u> SCD <u>IPVFMNART</u> KND <u>FTWFKL</u> ND <u>TLDY</u> ECH <u>DGY</u> ESNT <u>GTG</u> IV <u>CGYNGW</u> SD <u>LPI</u> CY <u>ERE</u> EL <u>PKIDV</u> HLV <u>PD</u> RKK <u>DQYKVGEVL</u> K <u>ESCKPGFT</u> IV <u>GPNSV</u> Q <u>CYHEGLSPD</u> LICK <u>EQV</u> Q <u>SCGPP</u> ELL <u>NGNV</u> KE <u>KTKEEYGH</u> SE <u>VVEY</u> C <u>NPFL</u> M <u>GPNK</u> I <u>Q</u> CV <u>DE</u> G <u>WT</u> TL <u>PVC</u> IV <u>EESTCGD</u> IP <u>EHGWA</u> QL <u>SPP</u> YY <u>GD</u> VE <u>FNCS</u> ES <u>FTM</u> H <u>NS</u> IRY <u>RCR</u> G <u>KEW</u> I <u>H</u> TC <u>INGR</u> W <u>DPEVNC</u> MA <u>QI</u> QL <u>C</u> PP <u>PP</u> Q I <u>PNSHNMTT</u> LN <u>YRD</u> GE <u>KVS</u> V <u>L</u> C <u>Q</u> EN <u>YLI</u> Q <u>E</u> GE <u>E</u> I <u>T</u> CK <u>DGR</u> W <u>Q</u> S <u>I</u> PLC VE <u>KI</u> PC <u>SQ</u> PP <u>Q</u> I <u>E</u> HT <u>IN</u> SS <u>RS</u> Q <u>E</u> SYAH <u>G</u> T <u>KL</u> S <u>Y</u> T <u>CE</u> GG <u>F</u> R <u>I</u> SE <u>E</u> NE TT <u>CYMGK</u> W <u>NSPP</u> Q <u>E</u> GL <u>PCK</u> S <u>P</u> PE <u>I</u> SH <u>G</u> V <u>V</u> A <u>H</u> M <u>D</u> S <u>Y</u> Q <u>G</u> EE <u>V</u> Y <u>K</u> CF E <u>GE</u> G <u>IDG</u> PA <u>I</u> A <u>K</u> CL <u>GEK</u> W <u>SH</u> PP <u>SC</u> I <u>K</u> T <u>D</u> CL <u>S</u> LP <u>F</u> EN <u>A</u> IP <u>M</u> GE <u>K</u> D <u>V</u> K <u>AGB</u> Q <u>U</u> T <u>Y</u> C <u>ATY</u> K <u>M</u> D <u>G</u> A <u>S</u> N <u>V</u> T <u>C</u> IN <u>S</u> R <u>WT</u> G <u>R</u> PT <u>C</u> RD <u>T</u> SC <u>V</u> N <u>P</u> TV <u>Q</u> N AY <u>IVS</u> R <u>QMS</u> K <u>YPS</u> GE <u>V</u> R <u>Y</u> Q <u>CR</u> S <u>P</u> Y <u>E</u> MF <u>G</u> DE <u>E</u> VM <u>C</u> LN <u>G</u> N <u>W</u> TE <u>P</u> PK <u>Q</u> KD ST <u>G</u> KG <u>PPP</u> PI <u>D</u> NG <u>D</u> ITS <u>F</u> PL <u>S</u> V <u>A</u> P <u>ASS</u> V <u>EY</u> Q <u>Q</u> LN <u>Y</u> Q <u>LE</u> GN <u>K</u> R <u>I</u> T <u>C</u> RNG <u>Q</u> W <u>SE</u> PP <u>K</u> CL <u>H</u> PC <u>V</u> IS <u>R</u> E <u>I</u> ME <u>NY</u> N <u>I</u> AL <u>R</u> W <u>T</u> A <u>K</u> Q <u>K</u> LY <u>S</u> RT <u>G</u> E <u>S</u> VE <u>F</u> V CK <u>RGY</u> RL <u>S</u> SR <u>S</u> HT <u>L</u> RT <u>T</u> C <u>W</u> D <u>G</u> K <u>L</u> E <u>Y</u> PT <u>C</u> AK <u>R</u>

- continued

Sequences		
SEQ ID NO:	Description	Sequence
30	Human Complement Receptor 1 (UniProt: P17927 residues 491 to 1134, from CCP8 to CCP17) [CCP domains '8-10' and '15-17' indicated by single underline, individual CCP domains 8, 9, 10, 15, 16, 17 indicated by double underline; amino acid differences between CCPs 8-10 and 15-17 indicated in bold/wavy underline]	<u>GHCQAPDH</u> <u>EFAKLKT</u> <u>QTNASDEPIGTS</u> <u>LKYECRPEYYGRPFSITCLD</u> <u>NLVWSSPKDVCKRKSC</u> <u>KTPPD</u> <u>PVNGMVH</u> <u>VIDI</u> <u>QVGSRINYSCTGHR</u> <u>LIGHSSAECILSGN</u> <u>AH</u> <u>WSTKPPICQ</u> <u>RIPCGLPTIANGDFISTNREN</u> <u>FHYGSVVTYRCN</u> <u>EGSGRKV</u> <u>FELVGEPSI</u> <u>YCTSNDQVG</u> <u>IWSGPAPQC</u> <u>IIPNKCTPPNVENG</u> <u>ILVSDNRSL</u> <u>FSLN</u> <u>EVVEFRCQPGFVMKGPRRVKC</u> <u>QALNKWEPELPSCSRVC</u> <u>COPPDVL</u> <u>H</u> <u>AERTQRDKDNE</u> <u>SPGQEVEYSCEP</u> <u>GYDLRGAASMRCTPQGD</u> <u>WSPAAPT</u> <u>CEVKSCDDFMGQ</u> <u>LLNGRVLFPVNL</u> <u>QLGAKVDFVCDEGFL</u> <u>QKGS</u> <u>SASYC</u> <u>VLAGMESLWN</u> <u>SSVPVCEQIFCPSP</u> <u>PVI</u> <u>PNGRHTGKPL</u> <u>EVFPPGKAVN</u> <u>YTC</u> <u>DPHDRGTSF</u> <u>LIGESTIRCTS</u> <u>DPQNGVWSSPAPRCG</u> <u>ILGH</u> <u>HCQAPDH</u> <u>EFAKLKT</u> <u>QTNASDEPIGTS</u> <u>LK</u> <u>YECRPEYYGRPFSITCLD</u> <u>NLVWSSPKDVCKRKSC</u> <u>KTPPD</u> <u>PVNGMVH</u> <u>VIDI</u> <u>QVGSRINYSCTGHR</u> <u>LIGHSSAECILSGN</u> <u>AH</u> <u>WSTKPPICQ</u> <u>RIPCG</u> <u>GLPPTIANGDFISTNREN</u> <u>FHYGSVVTYRCN</u> <u>EGSGRKV</u> <u>FELVGEPSIY</u> <u>CTSNDQVG</u> <u>IWSGPAPQCII</u>
31	Human Complement Receptor 1 (UniProt: P17927 residues 491 to 1134, CCP8 to CCP17) Non-glycosylated [CCP domains '8-10' and '15-17' indicated by single underline, individual CCP domains 8, 9, 10, 15, 16, 17 indicated by double underline; amino acid differences between CCPs 8-10 and 15-17 indicated in bold/wavy underline; substitutions in bold/dotted underline correspond to positions 509, 578, 959 and 1028 of UniProt: P17927]]	<u>GHCQAPDH</u> <u>EFAKLKT</u> <u>Q</u> <u>ASDEPIGTS</u> <u>LKYECRPEYYGRPFSITCLD</u> <u>NLVWSSPKDVCKRKSC</u> <u>KTPPD</u> <u>PVNGMVH</u> <u>VIDI</u> <u>QVGSR</u> <u>IQY</u> <u>SCTGHR</u> <u>LIGHSSAECILSGN</u> <u>AH</u> <u>WSTKPPICQ</u> <u>RIPCGLPTIANGDFISTNREN</u> <u>FHYGSVVTYRCN</u> <u>RGSGRKV</u> <u>FELVGEPSI</u> <u>YCTSNDQVG</u> <u>IWSGPAPQC</u> <u>IIPNKCTPPNVENG</u> <u>ILVSDNRSL</u> <u>FSLN</u> <u>EVVEFRCQPGFVMKGPRRVKC</u> <u>QALNKWEPELPSCSRVC</u> <u>COPPDVL</u> <u>H</u> <u>AERTQRDKDNE</u> <u>SPGQEVEYSCEP</u> <u>GYDLRGAASMRCTPQGD</u> <u>WSPAAPT</u> <u>CEVKSCDDFMGQ</u> <u>LLNGRVLFPVNL</u> <u>QLGAKVDFVCDEGFL</u> <u>QKGS</u> <u>SASYC</u> <u>VLAGMESLWN</u> <u>SSVPVCEQIFCPSP</u> <u>PVI</u> <u>PNGRHTGKPL</u> <u>EVFPPGKAVN</u> <u>YTC</u> <u>DPHDRGTSF</u> <u>LIGESTIRCTS</u> <u>DPQNGVWSSPAPRCG</u> <u>ILGH</u> <u>HCQAPDH</u> <u>EFAKLKT</u> <u>Q</u> <u>ASDEPIGTS</u> <u>LK</u> <u>YECRPEYYGRPFSITCLD</u> <u>NLVWSSPKDVCKRKSC</u> <u>KTPPD</u> <u>PVNGMVH</u> <u>VIDI</u> <u>QVGSR</u> <u>IQY</u> <u>SCTGHR</u> <u>LIGHSSAECILSGN</u> <u>AH</u> <u>WSTKPPICQ</u> <u>RIPCG</u> <u>GLPPTIANGDFISTNREN</u> <u>FHYGSVVTYRCN</u> <u>RGSGRKV</u> <u>FELVGEPSIY</u> <u>CTSNDQVG</u> <u>IWSGPAPQCII</u>
32	Human Complement Receptor 1 (UniProt: P17927; Entry version 181 (25 Oct. 2017), Sequence version 3 (02 Mar. 2010)) Transmembrane domain (residues 1972-1996	ALIVGTLSTGTIFFILLIIFLSWITL
33	Human Complement Receptor 1 (UniProt: P17927; Entry version 181 (25 Oct. 2017), Sequence version 3 (02 Mar. 2010)) Cytoplasmic tail (residues 1997-2039)	KHRKGNNAHENPKEVAIHLHSQGSSVHPRTLQTNNEENSRVLP
34	Tobacco Etch Virus (TEV) cleavage site	ENLYFQGS
35	Human Complement Receptor 1 CCPs 8-10 protein coding sequence Including leader sequence (double underlined) Including HindIII restriction enzyme cleavage site (AAGCTT), Kozak consensus sequence (GCCACC), stop codon (TAA) and EcoRV restriction enzyme cleavage site (GATATC)	AAGCTTGCCACCATGAGACTGCTGGCCAAGATCATTCGCTGTGCTG <u>TGGGC</u> <u>CATCTCGCTGGCGGACAT</u> <u>TGTCA</u> <u>GGCCCC</u> <u>CTGACCA</u> <u>CTTCC</u> <u>TG</u> <u>TTCCGCAAGCTGAAA</u> <u>ACCCAGACCA</u> <u>ACGCCAGCGACT</u> <u>TCCT</u> <u>ATCGGC</u> <u>ACCA</u> <u>CGCTGAAGTACG</u> <u>AGTCGAC</u> <u>ACCCGAGTA</u> <u>CTACGGCAGAC</u> <u>CTTC</u> <u>AGCATCAC</u> <u>CTGTCTGGACA</u> <u>ACCTCTGTGGT</u> <u>CTAGCCC</u> <u>AAAGAC</u> <u>GTG</u> <u>TGCAAGAGAA</u> <u>AGAC</u> <u>GTGCAAGAC</u> <u>CCCTCTGT</u> <u>GATCTGTGAAC</u> <u>GGCATG</u> <u>GTGCA</u> <u>CGCTGATCAC</u> <u>CGACATCCA</u> <u>AGTGGCAGCAGA</u> <u>ATCA</u> <u>ACTAC</u> <u>ACAGC</u> <u>TGCAC</u> <u>CCACGGCC</u> <u>ACAGACTGATGGCAGACT</u> <u>CTAGCGCC</u> <u>GAGTGT</u> <u>TATC</u> <u>CTGAGCG</u> <u>GAATGCC</u> <u>CACTG</u> <u>CTGCA</u> <u>CCAA</u> <u>GCTCT</u> <u>AACTCTG</u> <u>CCAG</u> <u>AGAATCC</u> <u>CTGCGGCTG</u> <u>CTCTACA</u> <u>ATCGCC</u> <u>AAACGGC</u> <u>GATT</u> <u>CATC</u> <u>AGCACCA</u> <u>ACAGAGA</u> <u>ACTTCC</u> <u>ACTACGGCT</u> <u>CCGTGGTC</u> <u>ACCTAC</u> <u>CAGA</u> <u>TGCA</u> <u>ATCTGGCAGCGGGC</u> <u>AGAAAGGT</u> <u>GTTCGA</u> <u>ACTTG</u> <u>GGCGAG</u> <u>CCCGCAT</u> <u>TACTGC</u> <u>ACCA</u> <u>GCAACGATGACCA</u> <u>AGTCGGC</u> <u>ATTGGAGC</u> <u>GGCC</u> <u>CTGCTC</u> <u>TCAGTGC</u> <u>CATCAT</u> <u>TAAGATATC</u>

- continued

Sequences		
SEQ ID NO:	Description	Sequence
36	Human Complement Receptor 1 CCPs 8-10 protein coding sequence (underlined) Non-glycosylated Including leader sequence (double underlined) Including HindIII restriction enzyme cleavage site (AAGCTT), Kozak consensus sequence (GCCACC), stop codon (TAA) and EcoRV restriction enzyme cleavage site (GATATC)	AAGCTTGCCACCAT <u>GAGACTGCTGGCCAAGATCATCTGCCTGATGCTG</u> <u>TGGCCATCTGCGTGGCCGACATTGTCAAGGCCCTGACACTTCCTG</u> TTCGCCAAGCT <u>GAACCCAGACACAGGCCAGCGACTTCCCTATCGC</u> ACCAGCCTGAAG <u>GTACGAGTCGAGACCCGAGTACTACGGCAGACCC</u> CTC <u>AGCAGCTGACTGGACAACCTGTGTGGCTAGGCCAAGGACGTG</u> TGCAAGAGAA <u>AGAGCTGCAAGACCCCTCTGATCTGTGAACGGCAG</u> GTGAC <u>AGATCCCTTGCGCCCTGCTTACATCGCCACAGGCCAGTT</u> AGC <u>ACAAACAGAGAGAACCTCCACTACGGCTCCGTGGTCACCTACAGA</u> TGCAAT <u>CTGGCACGGCGCAGAAAGTGTGAACTTGTGGCGAG</u> CCCAGC <u>ATCTACTGCACCAAGCAACGATGACCAAGTCGGCATTGGAGC</u> GGCC <u>CTGCTCTCAGTGCATCATCTAAGATATC</u>
37	Human Complement Receptor 1 CCPs 15-17 protein coding sequence (underlined) Including leader sequence (double underlined) Including HindIII restriction enzyme cleavage site (AAGCTT), Kozak consensus sequence (GCCACC), stop codon (TAA) and EcoRV restriction enzyme cleavage site (GATATC)	AAGCTTGCCACCAT <u>GAGACTGCTGGCCAAGATCATCTGCCTGATGCTG</u> <u>TGGCCATCTGCGTGGCCGACACTGTCAAGGCCCTGATCACTTCCTG</u> TTCGCCAAGCT <u>GAACCCAGACACAGGCCAGCGACTTCCCTATCGC</u> ACCAGCCTGAAG <u>GTACGAGTCGAGACCCGAGTACTACGGCAGACCC</u> CTC <u>AGCATTACCTGTCTGGACAACCTGTGTGGCTAGGCCAAGGACGTG</u> TGCAAGAGAA <u>AGAGCTGCAAGACCCCTCTGATCTGTGAACGGCAG</u> GTGAC <u>AGCTGATCACCGACATCCAAGTGGCAGCAGAATCCAGTACAGC</u> TG <u>GCACCCAGGGCACAGACTGATGGCCACTCTAGGCCAGTGATC</u> CTG <u>AGCGGAAACACAGCCACTGGTCCACCAAGCCTCCAAATTCAG</u> AG <u>ATCCCTTGCGCCCTGCTTACATCGCCACGGCAGTTTCATC</u> AG <u>ACCAACAGAGAGAACCTCCACTACGGCTCCGTGGTCACCTACAGA</u> TG <u>CAACCTGGCAGGGCGCAGAAAGTGTGAACTTGTGGCGAG</u> CCCAGC <u>ATCTACTGCACCAAGCAACGACGACCAAGTCGGCATTGGAGC</u> GGCC <u>CTGCTCTCAGTGCATCATCTAAGATATC</u>
38	Human Complement Receptor 1 CCPs 15-17 protein coding sequence (underlined) Non-glycosylated Including leader sequence (double underlined) Including HindIII restriction enzyme cleavage site (AAGCTT), Kozak consensus sequence (GCCACC), stop codon (TAA) and EcoRV restriction enzyme cleavage site (GATATC)	AAGCTTGCCACCAT <u>GAGACTGCTGGCCAAGATCATCTGCCTGATGCTG</u> <u>TGGCCATCTGCGTGGCCGACACTGTCAAGGCCCTGATCACTTCCTG</u> TTCGCCAAGCT <u>GAACCCAGACACAGGCCAGCGACTTCCCTATCGC</u> ACCAGCCTGAAG <u>GTACGAGTCGAGACCCGAGTACTACGGCAGACCC</u> CTC <u>AGCATTACCTGTCTGGACAACCTGTGTGGCTAGGCCAAGGACGTG</u> TGCAAGAGAA <u>AGAGCTGCAAGACCCCTCTGATCTGTGAACGGCAG</u> GTGAC <u>AGCTGATCACCGACATCCAAGTGGCAGCAGAATCCAGTACAGC</u> TG <u>GCACCCAGGGCACAGACTGATGGCCACTCTAGGCCAGTGATC</u> CTG <u>AGCGGAAACACAGCCACTGGTCCACCAAGCCTCCAAATTCAG</u> AG <u>ATCCCTTGCGCCCTGCTTACATCGCCACGGCAGTTTCATC</u> AG <u>ACCAACAGAGAGAACCTCCACTACGGCTCCGTGGTCACCTACAGA</u> TG <u>CAACCTGGCTCCAGAGGCCAGAAAGTGTGAACTTGTGGCGAG</u> CCT <u>AGCATCTACTGCACCAAGCAACGACGACCAAGTCGGCATTGGAGC</u> GGAC <u>CTGCTCTCAGTGCATCATCTAAGATATC</u>
39	Human Complement Receptor 1 CCPs 8-10 protein coding sequence (underlined) Including leader sequence (doubled underlined); His tag (dotted line); TEV cleavage site (dashed line) Including HindIII restriction enzyme cleavage site (AAGCTT), Kozak consensus sequence (GCCACC), stop codon (TAA) and EcoRV restriction enzyme cleavage site (GATATC)	AAGCTTGCCACCAT <u>GAGACTGCTGGCCAAGATCATCTGCCTGATGCTG</u> <u>TGGCCATCTGCGTGGCCACACCATCACCATCACGGCAGCAGCGAG</u> <u>AACCTGTACTCCAGGGATCTCTGGGGCCACTGTCAAGGCCCTGAT</u> CACT <u>CTCTGTTGCCAAGCTGAAAACCCAGACCAACGCCAGCGACTTC</u> CCT <u>ATCGGCACCCCTGAAGTAGCGAGTCGAGACCCGAGTACTACGGC</u> AGAC <u>CCCTTCAGCATCACCTGTCTGGACAACCTCTGTTGATCTGGCC</u> AAG <u>GGACGTGTGCAAGAGAAAGAGCTGCAAGACCCCTCTGATCTGTG</u> AAC <u>CGGCATGGTGCACGTATCACCGACATCCAAGTGGCAGCAGAAC</u> AA <u>CTACAGCTGCACCCAGGCCACAGACTGATCGGACACTCTAGCGCC</u> GAG <u>GTATCTGAGCGCAATGCCACACTGGTCCACCAAGCCTCCA</u> AT <u>CTGCCAGAGAACTCCCTTGCGGCTGCTTACATCGCCACCGC</u> GAT <u>TTCATCAGCAACAGAGAGAACCTCCACTACGGCTCCGTGGC</u> AC <u>CTACAGATGCAATCTGGCAGGGCGCAGAAAGGTGTGAACTT</u> GT <u>GGCGAGGCCAGCATCTACTGCACCAAGCAACGATGACCAAGTCGGC</u> ATT <u>GGAGCGCCCTGCTCTCAGTGCATCATCTAAGATATC</u>

-continued

Sequences		
SEQ ID NO:	Description	Sequence
40	Human Complement Receptor 1 CCPs 8-10 Including leader sequence (doubled underlined); His tag (dotted line); TEV cleavage site (dashed line)	MRLLA <u>KIICLMLWAI</u> CVAH <u>HHHHHGSS</u> ENLYFOG <u>S</u> GGHCQAPDHFLF AKLK <u>TQTNASDFP</u> IGTS <u>LKYECRPEYYGRPFSIT</u> CLDNLVWSSPKDVC KRKSCKTP <u>PPDPVN</u> GMVH <u>ITDIQV</u> GSRIN <u>YSC</u> TTGHRLIGH <u>S</u> AECIL SGNAAHWSTKPP <u>IQRIP</u> CGLPPTIANG <u>D</u> FISTNREN <u>FHYGSV</u> VTYRC NPGSGGRKV <u>FELVGE</u> PSIYCTSND <u>DQV</u> GIWSGP <u>APQCII</u>
41	Human Complement Receptor 1 CCPs 8-10 protein coding sequence (underlined) Non-glycosylated Including leader sequence (doubled underlined); His tag (dotted line); TEV cleavage site (dashed line) Including HindIII restriction enzyme cleavage site (AAGCTT), Kozak consensus sequence (GCCACC), stop codon (TAA) and EcoRV restriction enzyme cleavage site (GATATC)	AAGCTTGCCACC <u>ATGAGACTG</u> CTGGCC <u>AAAGATCAT</u> CTGCCTGATGCTG <u>TGGCC</u> CAT <u>TGCGT</u> GGCC <u>ACCACCA</u> C <u>ATCAC</u> GGCAGCAG <u>CAG</u> <u>AACTGT</u> ACT <u>TC</u> CA <u>GGATC</u> T <u>TCTG</u> GG <u>GGCC</u> ACT <u>GTC</u> AG <u>GGCCC</u> CT <u>GAT</u> <u>CACTTC</u> CT <u>TG</u> CC <u>AAAGCTG</u> AAA <u>ACCCAGAC</u> CA <u>ACGCC</u> AG <u>CGC</u> ACT <u>TC</u> <u>CCTATCG</u> GC <u>ACCACG</u> CT <u>GAAGT</u> AC <u>GAGT</u> GC <u>AGACCC</u> CG <u>GAGT</u> ACT <u>ACGGC</u> <u>AGACCC</u> TT <u>CAGCAT</u> AC <u>CTGT</u> CT <u>GGACA</u> AC <u>CTGT</u> GT <u>GGT</u> CT <u>AGCCCC</u> <u>AAGGAC</u> GT <u>TG</u> CA <u>AGAGA</u> AA <u>AGAGCTG</u> CA <u>AGACCC</u> CT <u>CTGAT</u> CC <u>TGTG</u> <u>AACCGC</u> AT <u>TGGT</u> GC <u>ACGT</u> AT <u>CACCG</u> AC <u>ATCCA</u> AG <u>TGGC</u> AG <u>CGAGA</u> ATC <u>AACTAC</u> AG <u>TG</u> CA <u>CCACCG</u> CC <u>ACAGACT</u> G <u>ATCGG</u> AC <u>ACTCTAGCGC</u> <u>GAGT</u> GT <u>TCTG</u> GA <u>GCGG</u> CA <u>ATGCGC</u> AC <u>ACTGTG</u> CC <u>ACCAAGCC</u> CT <u>TCCA</u> <u>ATCTGCC</u> CA <u>GAGAA</u> AT <u>CCCTTG</u> GG <u>CCCTG</u> CC <u>CTTACA</u> AT <u>CGCC</u> AC <u>CGGC</u> <u>GATT</u> TC <u>CATCAG</u> CA <u>ACAGAGA</u> AA <u>AGACTTCC</u> AC <u>TACGG</u> CT <u>CCGG</u> <u>ACCTAC</u> AG <u>TG</u> CA <u>ACCTG</u> GC <u>AGCGG</u> CC <u>AGAAAGGTG</u> TT <u>CGA</u> ACTT <u>GTGGC</u> G <u>AGGCC</u> AG <u>CACAT</u> CT <u>ACTG</u> AC <u>CAAGCA</u> CG <u>ATGACCA</u> AG <u>TCGG</u> <u>ATTGGAGCGGCC</u> CT <u>GCT</u> CC <u>TCA</u> GT <u>GCAT</u> CAT <u>CC</u> TA <u>AGAT</u> ATC
42	Human Complement Receptor 1 CCPs 8-10 Non-glycosylated Including leader sequence (doubled underlined); His tag (dotted line); TEV cleavage site (dashed line)	MRLLA <u>KIICLMLWAI</u> CVAH <u>HHHHHGSS</u> ENLYFOG <u>S</u> GGHCQAPDHFLF AKLK <u>TQTNASDFP</u> IGTS <u>LKYECRPEYYGRPFSIT</u> CLDNLVWSSPKDVC KRKSCKTP <u>PPDPVN</u> GMVH <u>ITDIQV</u> GSRIN <u>YSC</u> TTGHRLIGH <u>S</u> AECIL SGNAAHWSTKPP <u>IQRIP</u> CGLPPTIANG <u>D</u> FISTNREN <u>FHYGSV</u> VTYRC NPGSGGRKV <u>FELVGE</u> PSIYCTSND <u>DQV</u> GIWSGP <u>APQCII</u>
43	Human Complement Receptor 1 CCPs 15-17 protein coding sequence (underlined) Including leader sequence (doubled underlined); His tag (dotted line); TEV cleavage site (dashed line) Including HindIII restriction enzyme cleavage site (AAGCTT), Kozak consensus sequence (GCCACC), stop codon (TAA) and EcoRV restriction enzyme cleavage site (GATATC)	AAGCTTGCCACC <u>ATGAGACTG</u> CTGGCC <u>AAAGATCAT</u> CTGCCTGATGCTG <u>TGGCC</u> CAT <u>TGCGT</u> GGCC <u>ACCACCA</u> C <u>ATCAC</u> GGCAGCAG <u>CAG</u> <u>AACTGT</u> ACT <u>TC</u> CA <u>GGATC</u> T <u>TCTG</u> GG <u>GGCC</u> ACT <u>GTC</u> AG <u>GGCCC</u> CT <u>GAT</u> <u>CACTTC</u> CT <u>TG</u> CC <u>AAAGCTG</u> AAA <u>ACCCAGAC</u> CA <u>ACGCC</u> AG <u>CGC</u> ACT <u>TC</u> <u>CCTATCG</u> GC <u>ACCACG</u> CT <u>GAAGT</u> AC <u>GAGT</u> GC <u>AGACCC</u> CG <u>GAGT</u> ACT <u>ACGGC</u> <u>AGACCC</u> TT <u>CAGCAT</u> AC <u>CTGT</u> CT <u>GGACA</u> AC <u>CTGT</u> GT <u>GGT</u> CT <u>AGCCCC</u> <u>AAGGAC</u> GT <u>TG</u> CA <u>AGAGA</u> AA <u>AGAGCTG</u> CA <u>AGACCC</u> CT <u>CTGAT</u> CC <u>TGTG</u> <u>AACCGC</u> AT <u>TGGT</u> GC <u>ACGT</u> AT <u>CACCG</u> AC <u>ATCCA</u> AG <u>TGGC</u> AG <u>CGAGA</u> ATC <u>AACTAC</u> AG <u>TG</u> CA <u>CCACCG</u> CC <u>ACAGACT</u> G <u>ATCGG</u> AC <u>ACTCTAGCGC</u> <u>GAGT</u> GT <u>TCTG</u> GA <u>GCGG</u> CA <u>ATGCGC</u> AC <u>ACTGTG</u> CC <u>ACCAAGCC</u> CT <u>TCCA</u> <u>ATCTGCC</u> CA <u>GAGAA</u> AT <u>CCCTTG</u> GG <u>CCCTG</u> CC <u>CTTACA</u> AT <u>CGCC</u> AC <u>CGGC</u> <u>GATT</u> TC <u>CATCAG</u> CA <u>ACAGAGA</u> AA <u>AGACTTCC</u> AC <u>TACGG</u> CT <u>CCGG</u> <u>ACCTAC</u> AG <u>TG</u> CA <u>ACCTG</u> GC <u>AGCGG</u> CC <u>AGAAAGGTG</u> TT <u>CGA</u> ACTT <u>GTGGC</u> G <u>AGGCC</u> AG <u>CACAT</u> CT <u>ACTG</u> AC <u>CAAGCA</u> CG <u>ATGACCA</u> AG <u>TCGG</u> <u>ATTGGAGCGGCC</u> CT <u>GCT</u> CC <u>TCA</u> GT <u>GCAT</u> CAT <u>CC</u> TA <u>AGAT</u> ATC
44	Human Complement Receptor 1 CCPs 15-17 Including leader sequence (doubled underlined); His tag (dotted line); TEV cleavage site (dashed line)	MRLLA <u>KIICLMLWAI</u> CVAH <u>HHHHHGSS</u> ENLYFOG <u>S</u> GGHCQAPDHFLF AKLK <u>TQTNASDFP</u> IGTS <u>LKYECRPEYYGRPFSIT</u> CLDNLVWSSPKDVC KRKSCKTP <u>PPDPVN</u> GMVH <u>ITDIQV</u> GSRIN <u>YSC</u> TTGHRLIGH <u>S</u> AECIL SGNTAHWSTKPP <u>IQRIP</u> CGLPPTIANG <u>D</u> FISTNREN <u>FHYGSV</u> VTYRC NLGS <u>RGRKV</u> FELV <u>GE</u> PSIYCTSND <u>DQV</u> GIWSGP <u>APQCII</u>
45	Human Complement Receptor 1 CCPs 15-17 protein coding sequence (underlined) Non-glycosylated Including leader	AAGCTTGCCACC <u>ATGAGACTG</u> CTGGCC <u>AAAGATCAT</u> CTGCCTGATGCTG <u>TGGCC</u> CAT <u>TGCGT</u> GGCC <u>ACCACCA</u> C <u>ATCAC</u> GGCAGCAG <u>CAG</u> <u>AACTGT</u> ACT <u>TC</u> CA <u>GGATC</u> T <u>TCTG</u> GG <u>GGCC</u> ACT <u>GTC</u> AG <u>GGCCC</u> CT <u>GAT</u> <u>CACTTC</u> CT <u>TG</u> CC <u>AAAGCTG</u> AAA <u>ACCCAGAC</u> CA <u>ACGCC</u> AG <u>CGC</u> ACT <u>TC</u> <u>CCTATCG</u> GC <u>ACCACG</u> CT <u>GAAGT</u> AC <u>GAGT</u> GC <u>AGACCC</u> CG <u>GAGT</u> ACT <u>ACGGC</u> <u>AGACCC</u> TT <u>CAGCAT</u> AC <u>CTGT</u> CT <u>GGACA</u> AC <u>CTGT</u> GT <u>GGT</u> CT <u>AGCCCC</u>

- continued

Sequences		
SEQ ID NO:	Description	Sequence
	sequence (doubled underlined); His tag (dotted line); TEV cleavage site (dashed line) Including HindIII restriction enzyme cleavage site (AAGCTT), Kozak consensus sequence (GCCACC), stop codon (TAA) and EcoRV restriction enzyme cleavage site (GATATC)	AAGGACGTGTGCAAGAGAAAAGAGCTGCAAGACCCCTCTGATCTGTG AACCGCATGGTCACGTGATCACCGACATCCAAGTGGCAGCAGAAATC CAGTACAGTGACCCACAGGCCACAGACTGATGGCCACTCTAGCGCC GAGGTGATCTGAGCGGAAACACAGCCACTGGTCCACCAAGCTTCCA ATCTGCCAGAGAAATCCCTGCGGCCCTGCCTCCATACTGCCAACCGGC GATTTCATCAGCACCAACAGAGAGAACTTCCACTACGGCTCCGGTGGC ACCTACAGATGCAACCTGGGCTCCAGAGGCCGAAGGTGTTCGAACTT GTGGCGAGCCTAGCATCTACTGCAACAGCAACGAGACCAAGTCGGC ATTGGAGCGGACCTGCTCCAGTCATCATCCTAAGATATC
46	Human Complement Receptor 1 CCPs 15-17 Non-glycosylated Including leader sequence (doubled underline); His tag (dotted line); TEV cleavage site (dashed line)	MRLLAKIICLMLWAICVA <u>HHHHHHGSS</u> ENLYFOG <u>SSGGHCQADPHFLF</u> AKLKTQTQASDFPIGTS <u>LKYECRPEYYGRPFSITCLDNLVWSSPKDVKR</u> KSC <u>TKPPDPVNGMVH</u> C KRKSCKTPPD <u>PVNGMVH</u> VITDIQVGSRI <u>QYSCTGHRLIGHSSAE</u> CILSGNAAHWSTKPPIC <u>QRI</u> PC SGNAAHWSTKPPIC <u>QRI</u> PC <u>G</u> LPPTIANGDFISTNREN <u>FHYGSVVTYRCNP</u> PGSGRKVFELVGEPSIY NPGSGGRKV <u>FELVGEPSIY</u> CTSNDQVGIWSGPAPQCII
47	CR1a Including leader sequence (double underline); CCPs 8-10	MRLLAKIICLMLWAICVA <u>GAGHCQADPHFLFAKLKTQTQASDFPIGTS</u> LK YECP <u>EYGRPFSITCLDNLVWSSPKDVKR</u> KSC <u>TKPPDPVNGMVH</u> VI TDI <u>QVGSRINYSCTGHRLIGHSSAE</u> CILSGNAAHWSTKPPIC <u>QRI</u> PC GLPPTIANGDFISTNREN <u>FHYGSVVTYRCNP</u> PGSGRKVFELVGEPSIY CTSNDQVGIWSGPAPQCII
48	nCR1a Including leader sequence (double underline); CCPs 8-10; N509Q and N578Q (numbered according to UniProt: P17927)	MRLLAKIICLMLWAICVA <u>GAGHCQADPHFLFAKLKTQTQASDFPIGTS</u> LK YECP <u>EYGRPFSITCLDNLVWSSPKDVKR</u> KSC <u>TKPPDPVNGMVH</u> VI TDI <u>QVGSRINYSCTGHRLIGHSSAE</u> CILSGNAAHWSTKPPIC <u>QRI</u> PC GLPPTIANGDFISTNREN <u>FHYGSVVTYRCNP</u> PGSGRKVFELVGEPSIY CTSNDQVGIWSGPAPQCII
49	CR1b Including leader sequence (double underline); CCPs 15-17	MRLLAKIICLMLWAICVA <u>GAGHCQADPHFLFAKLKTQTQASDFPIGTS</u> LK YECP <u>EYGRPFSITCLDNLVWSSPKDVKR</u> KSC <u>TKPPDPVNGMVH</u> VI TDI <u>QVGSRINYSCTGHRLIGHSSAE</u> CILSGNTAHWSTKPPIC <u>QRI</u> PC GLPPTIANGDFISTNREN <u>FHYGSVVTYRCNLGSRGRK</u> V <u>FELVGEPSI</u> Y CTSNDQVGIWSGPAPQCII
50	nCR1b Including leader sequence (double underline); CCPs 15-17; N959Q and N1028Q (numbered according to UniProt: P17927)	MRLLAKIICLMLWAICVA <u>GAGHCQADPHFLFAKLKTQTQASDFPIGTS</u> LK YECP <u>EYGRPFSITCLDNLVWSSPKDVKR</u> KSC <u>TKPPDPVNGMVH</u> VI TDI <u>QVGSRINYSCTGHRLIGHSSAE</u> CILSGNTAHWSTKPPIC <u>QRI</u> PC GLPPTIANGDFISTNREN <u>FHYGSVVTYRCNP</u> PGSGRKVFELVGEPSIY CTSNDQVGIWSGPAPQCII
51	CR1a + CR1b Including leader sequence (double underline); CCPs 8-10 + 15-17	MRLLAKIICLMLWAICVA <u>GAGHCQADPHFLFAKLKTQTQASDFPIGTS</u> LK YECP <u>EYGRPFSITCLDNLVWSSPKDVKR</u> KSC <u>TKPPDPVNGMVH</u> VI TDI <u>QVGSRINYSCTGHRLIGHSSAE</u> CILSGNTAHWSTKPPIC <u>QRI</u> PC GLPPTIANGDFISTNREN <u>FHYGSVVTYRCNP</u> PGSGRKVFELVGEPSIY CTSNDQVGIWSGPAPQCII <u>G</u> LPPTIANGDFISTNREN <u>FHYGSVVTYRCNLGSRGRK</u> V <u>FELVGEPSI</u> Y VITDI <u>QVGSRINYSCTGHRLIGHSSAE</u> CILSGNTAHWSTKPPIC <u>QRI</u> PC PCGLPPTIANGDFISTNREN <u>FHYGSVVTYRCNLGSRGRK</u> V <u>FELVGEPSI</u> Y IYCTSNDQVGIWSGPAPQCII
52	nCR1a + nCR1b Including leader sequence (double underline); CCPs 8-10 + 15-17; N509Q and N578Q , N959Q and N1028Q (numbered according to UniProt: P17927)	MRLLAKIICLMLWAICVA <u>GAGHCQADPHFLFAKLKTQTQASDFPIGTS</u> LK YECP <u>EYGRPFSITCLDNLVWSSPKDVKR</u> KSC <u>TKPPDPVNGMVH</u> VI TDI <u>QVGSRINYSCTGHRLIGHSSAE</u> CILSGNAAHWSTKPPIC <u>QRI</u> PC GLPPTIANGDFISTNREN <u>FHYGSVVTYRCNP</u> PGSGRKVFELVGEPSIY CTSNDQVGIWSGPAPQCII <u>G</u> LPPTIANGDFISTNREN <u>FHYGSVVTYRCNLGSRGRK</u> V <u>FELVGEPSI</u> Y VITDI <u>QVGSRINYSCTGHRLIGHSSAE</u> CILSGNTAHWSTKPPIC <u>QRI</u> PC PCGLPPTIANGDFISTNREN <u>FHYGSVVTYRCNLGSRGRK</u> V <u>FELVGEPSI</u> Y IYCTSNDQVGIWSGPAPQCII

- continued

Sequences		
SEQ ID NO:	Description	Sequence
53	CR1a + nCR1b Including leader sequence (double underline); CCPs 8-10 + 15-17; N959Q and N1028Q (numbered according to UniProt: P17927)	MRLLA KI I CLML WA I CVAGHCQAPDHFLFAKLKTQTNA SDFPIGTS LK YECP EY YGRPFSITC LDNL VWSSPKDVCKRK SCKTPPD PVN G MVHVI TDI QVG S RIN YSCTIGHRLIGHSSAECILSGNAAH WSTKPP ICQRI PC GLPPTIANGDFISTNREN FHYGSV T YRCNP GS GRK V FELV GEPSIY CTSNDQVGIWSGP A PQCII GHC QAPDHFLFAKLKTQASDFPIGTS LKYECRPEYYGRPFSITC LDNL VWSSPKDVCKRK SCKTPPD PVN G MVH VITDIQVGSR I QYSCTIGHRLIGHSSAECILSGNTAH WSTKPP ICQRI PC PCGLPPTIANGDFISTNREN FHYGSV T YRCNP GS GRK V FELV GEPS IYCTSNDQVGIWSGP A PQCII
54	nCR1a + CR1b Including leader sequence (double underline); CCPs 8-10 + 15-17; N509Q and N578Q (numbered according to UniProt: P17927)	MRLLA KI I CLML WA I CVAGHCQAPDHFLFAKLKTQTNA SDFPIGTS LK YECP EY YGRPFSITC LDNL VWSSPKDVCKRK SCKTPPD PVN G MVHVI TDI QVG S RIN YSCTIGHRLIGHSSAECILSGNAAH WSTKPP ICQRI PC GLPPTIANGDFISTNREN FHYGSV T YRCNP GS GRK V FELV GEPSIY CTSNDQVGIWSGP A PQCII GHC QAPDHFLFAKLKTQASDFPIGTS LKYECRPEYYGRPFSITC LDNL VWSSPKDVCKRK SCKTPPD PVN G MVH VITDIQVGSR I QYSCTIGHRLIGHSSAECILSGNTAH WSTKPP ICQRI PC PCGLPPTIANGDFISTNREN FHYGSV T YRCNP GS GRK V FELV GEPS IYCTSNDQVGIWSGP A PQCII

Functional Properties of the Polypeptide

The polypeptide according to the present invention may be characterised by reference to one or more functional properties.

In particular, a polypeptide according to the present invention may possess one or more of the following properties (as determined by analysis in an appropriate assay for said property):

Binds to C3b;

Binds to C3b with an affinity of binding which is similar to the affinity of binding to C3b displayed by a co-factor for Complement Factor I (or a fragment thereof);

Binds to C3b with an affinity of binding which is higher than the affinity of binding to C3b displayed by a co-factor for Complement Factor I (or a fragment thereof);

Binds to C3b with an affinity of binding which is similar to the affinity of binding to C3b displayed by Complement Receptor 1 (or a fragment thereof);

Binds to C3b in the region of C3b bound by a co-factor for Complement Factor I (or a fragment thereof);

Binds to C3b in the region of C3b bound by Complement Receptor 1 (or a fragment thereof);

Binds to C3b in the region of C3b bound by Complement Receptor 1 CCP domains 8-10 and/or 15-17 (or a fragment thereof);

Acts as a co-factor to enable Complement Factor I-mediated inactivation of C3b;

Acts as a co-factor to enable Complement Factor I-mediated reduction/prevention of the formation of a functional C3bBb-type C3 convertase;

Acts as a co-factor to enable Complement Factor I-mediated reduction/prevention of the formation of a functional C4b2a3b-type C5 convertase;

Acts as a co-factor to enable Complement Factor I-mediated reduction/prevention of the formation of a functional C4b2a3b-type C5 convertase;

Acts as a co-factor to enable Complement Factor I-mediated reduction of C3bBb-type C3 convertase activity;

Acts as a co-factor to enable Complement Factor I-mediated reduction of C3bBb3b-type C5 convertase activity;

25 Acts as a co-factor to enable Complement Factor I-mediated reduction of C4b2a3b-type C5 convertase activity;

30 Acts as a co-factor to enable Complement Factor I-mediated reduction of the amount of C3bBb-type C3 convertase;

35 Acts as a co-factor to enable Complement Factor I-mediated reduction of the amount of C3bBb3b-type C5 convertase;

40 Acts as a co-factor to enable Complement Factor I-mediated reduction of the amount of C4b2a3b-type C5 convertase;

Reduces the amount of C3b via Complement Factor I; Increases the amount of iC3b via Complement Factor I; Increases the amount of C3dg via Complement Factor I; Increases the amount of C3d via Complement Factor I; Increases the amount of C3f via Complement Factor I; Reduces the amount of C5b via Complement Factor I; Reduces the amount of C5a via Complement Factor I; Decreases the amount of iC3b via Complement Factor I compared to the amount of iC3b

45 produced by FH and/or FHL-1 via Complement Factor I; Increases the ratio of C3dg to iC3b via Complement Factor I;

50 Is capable of inhibiting complement activation;

Diffuses through Bruch's membrane (BrM);

Displays superior ability to diffuse through BrM compared to Complement Factor I;

Displays superior ability to diffuse through BrM compared to a co-factor for Complement Factor I (or a fragment thereof);

Displays similar ability to diffuse through BrM compared to a co-factor for Complement Factor I (or a fragment thereof);

Displays superior ability to diffuse through BrM compared to Complement Factor H;

Displays similar ability to diffuse through BrM compared to Complement Factor H isoform FHL-1;

Displays superior ability to diffuse through BrM compared to Complement Factor H isoform FHL-1;

Displays similar ability to diffuse through BrM compared to soluble Complement Receptor 1;

55

60

65

Displays superior ability to diffuse through BrM compared to soluble Complement Receptor 1.

Whether a given polypeptide possesses the functional properties referred to in the previous paragraph can be analysed, for example, as described herein.

A polypeptide according to the present invention may be capable of binding to C3b. In some embodiments, a polypeptide according to the present invention may comprise or consist of a C3b binding region. In some embodiments, the C3b binding region of the polypeptide according to the present invention comprises or consists of a C3b binding region of CR1, e.g. CCP domains 8-10 and/or 15-17.

As used herein, a "C3b binding region" refers to a region capable of binding to C3b. In some embodiments, the C3b binding region is capable of specific binding to C3b. Binding to C3b may be mediated by non-covalent interactions such as Van der Waals forces, electrostatic interactions, hydrogen bonding, and hydrophobic interactions formed between the C3b binding region and C3b. In some embodiments, the C3b binding region binds to C3b with greater affinity, and/or with greater duration than it binds to molecules other than C3b.

The ability of a polypeptide according to the present invention or a putative C3b binding region to bind to C3b can be analysed using techniques well known to the person skilled in the art, including ELISA, Surface Plasmon Resonance (SPR; see e.g. Hearty et al., Methods Mol Biol (2012) 907:411-442; or Rich et al., Anal Biochem. 2008 Feb. 1; 373(1):112-20), Bio-Layer Interferometry (see e.g. Lad et al., (2015) J Biomol Screen 20(4): 498-507; or Concepcion et al., Comb Chem High Throughput Screen. 2009 September; 12(8):791-800), MicroScale Thermophoresis (MST) analysis (see e.g. Jerabek-Willemsen et al., Assay Drug Dev Technol. 2011 August; 9(4): 342-353), or by a radiolabelled antigen binding assay (RIA). Through such analysis binding to a given target can be determined and quantified. In some embodiments, the binding may be the response detected in a given assay.

In some embodiments, a polypeptide according to the present invention displays binding to C3b in such an assay which is greater than 1 times, e.g. one of >1.01, >1.02, >1.03, >1.04, >1.05, >1.06, >1.07, >1.08, >1.09, >1.1, >1.2, >1.3, >1.4, >1.5, >1.6, >1.7, >1.8, >1.9, >2, >3, >4, >5, >6, >7, >8, >9, >10, >15, >20, >25, >30, >35, >40, >45, >50, >60, >70, >80, >90, or >100 times the level of binding signal detected in such an assay to a negative control molecule to which the C3b binding region does not bind.

In some embodiments, a polypeptide according to the present invention is capable of binding to C3b with an affinity of binding which is similar to the affinity of binding to C3b displayed by CR1 or another co-factor for Complement Factor I (or a fragment thereof) in a given assay. An affinity of binding which is similar to a reference affinity of binding can be e.g. $\pm 40\%$ of the level of binding, e.g. one of $\pm 35\%$, $\pm 30\%$, 25% , 20% , 15% , 10% or $\pm 5\%$ of the level of binding to C3b displayed by reference CR1 or the reference co-factor for Complement Factor I in a comparable assay.

In some embodiments a polypeptide according to the present invention is capable of binding to C3b with an affinity of binding which is higher than the affinity of binding to C3b displayed by a co-factor for Complement Factor I (or a fragment thereof) in a given assay. In some embodiments a polypeptide according to the present invention is capable of binding to C3b with an affinity of binding which is 1.5 times, 2 times, 2.5 times, 3 times, 3.5 times, 4 times, 4.5 times, 5 times, 5.5 times, 6 times, 6.5 times, 7 times, 7.5 times, 8 times, 8.5 times, 9 times, 9.5 times, 10 times, 15 times, 20 times, 25 times, 30 times, 35 times, 40

times, 45 times, 50 times, 75 times, 100 times, 150 times, 200 times, 250 times, 300 times, 350 times, 400 times, 450 times, 500 times, 550 times, 600 times, 650 times, 700 times, 750 times, 800 times, 850 times, 900 times, 950 times, or 1000 times the affinity of binding to C3b displayed by a co-factor for Complement Factor I (or a fragment thereof) in a given assay. In some embodiments a polypeptide according to the present invention is capable of binding to C3b with an affinity of binding which is 10000 times, 100000 times, or 1000000 times the affinity of binding to C3b displayed by a co-factor for Complement Factor I (or a fragment thereof) in a given assay. In some embodiments a polypeptide according to the present invention is capable of binding to C3b with an affinity of binding which is 2, 3, 4, 5, 6, 7, 8, 9 or 10 order(s) of magnitude greater than the affinity of binding to C3b displayed by a co-factor for Complement Factor I (or a fragment thereof) in a given assay. In some embodiments the co-factor for Complement Factor I is Complement Factor H or truncated FH isoform FHL-1. The co-factor for Complement Factor I may be CR1.

In some embodiments, a polypeptide according to the present invention is capable of binding to C3b in the region of C3b that is bound by a co-factor for Complement Factor I (i.e. binds to the same region or an overlapping region). In some embodiments, the polypeptide is capable of binding to C3b in the region bound by CR1 (or a fragment thereof). In some embodiments, the polypeptide is capable of binding to C3b in the region bound by CR1 CCP domains 8-10 and/or 15-17. In some embodiments, the polypeptide is capable of binding to C3b in the region bound by one or more of Complement Factor I co-factors Complement Factor H, CD46, CD55, C4BP, SPICE, VCP, or MOPICE (or fragments thereof).

Whether a polypeptide according to the present invention binds to C3b in the region of C3b bound by a given co-factor for Complement Factor I (or a fragment thereof) can be determined by various methods known to the skilled person, including ELISA, and surface plasmon resonance (SPR) analysis. An example of a suitable assay to determine whether a C3b binding region binds to C3b in the region bound by a given co-factor for Complement Factor I (or a fragment thereof) is a competition ELISA assay.

For example, whether a polypeptide according to the present invention binds to C3b in the region of C3b bound by a given co-factor for Complement Factor I (or a fragment thereof) can be determined by analysis of interaction of the co-factor/fragment with C3b in the presence of, or following incubation of one or both of the co-factor/fragment and C3b with the polypeptide according to the present invention. A C3b binding region which binds to C3b in the region of C3b bound by a given co-factor/fragment is identified by the observation of a reduction/decrease in the level of interaction between the co-factor/fragment and C3b in the presence of—or following incubation of one or both of the interaction partners with—the polypeptide according to the present invention, as compared to the level of interaction in the absence of the polypeptide according to the present invention (or in the presence of an appropriate control peptide/polypeptide). Suitable analysis can be performed in vitro, e.g. using recombinant interaction partners. For the purposes of such assays, one or both of the interaction partners and/or the polypeptide according to the present invention may be labelled, or used in conjunction with a detectable entity for the purposes of detecting and/or measuring the level of interaction.

In some embodiments, the polypeptide according to the present invention acts as a co-factor for Complement Factor

I. For example, the polypeptide may potentiate cleavage of C3b by Complement Factor I, and/or present C3b in a favourable orientation for proteolytic cleavage by Complement Factor I. The polypeptide preferably does not inhibit proteolytic cleavage of C3b by Complement Factor I. In some embodiments, the Complement Factor I is endogenous. In some embodiments, the Complement Factor I is exogenous.

As used herein, an ‘endogenous’ protein/peptide refers to a protein/peptide which is encoded/expressed by the relevant cell type, tissue, or subject (prior to treatment with a polypeptide, nucleic acid, vector, cell or pharmaceutical composition according to the present invention). A ‘non-endogenous’ or ‘exogenous’ protein/peptide refers to a protein/peptide which is not encoded/expressed by, the relevant cell type, tissue, or subject (prior to treatment with a polypeptide, nucleic acid, vector, cell or pharmaceutical composition according to the present invention).

A polypeptide according to the present invention that acts as a co-factor for Complement Factor I can be determined by e.g. by analysis of the level or rate of proteolytic cleavage of C3b by Complement Factor I in a suitable assay in the presence of (or after incubation with) a polypeptide according to the present invention as compared to the level or rate of proteolytic cleavage of C3b by Complement Factor I in the absence of the polypeptide according to the present invention (or in the presence of an appropriate control peptide/polypeptide). A C3b binding region which acts as a co-factor for Complement Factor I is identified by the detection of an increased level or rate of proteolytic cleavage of C3b by Complement Factor I in the presence of (or after incubation with) a polypeptide according to the present invention as compared to the level or rate of proteolytic cleavage of C3b by Complement Factor I in the absence of the polypeptide according to the present invention (or in the presence of an appropriate control peptide/polypeptide). The level or rate of proteolytic cleavage of C3b by Complement Factor I can be determined e.g. by detection of one or more products of cleavage of C3b by Complement Factor I, e.g. iC3b, C3dg, C3d or C3f. The level or rate of proteolytic cleavage of C3b by Complement Factor I can be determined e.g. by detection of a reduction in the presence of C3b. In some embodiments a polypeptide according to the present invention that acts as a co-factor for Complement Factor I produces a smaller amount of iC3b overall compared to an amount of iC3b produced by FH/FHL-1 via Complement Factor I. In some embodiments a polypeptide according to the present invention that acts as a co-factor for Complement Factor I increases the ratio of C3dg to iC3b via Complement Factor I compared to the ratio of C3dg to iC3b produced by FH/FHL-1 via Complement Factor I. For example, the polypeptide may not increase the overall amount of iC3b, but may instead increase the amount of C3dg, C3f and/or C3d via Complement Factor I.

In some embodiments a polypeptide according to the present invention is capable of inhibiting or reducing complement activation. A polypeptide may inhibit or reduce complement over-activation. The level of complement activation/over-activation may be determined by the assays described herein, e.g. abnormal levels of complement components, or by tests/assays that are known by one skilled in the art, e.g. as described in Shih and Murali Am. J. Hematol. 2015, 90: 1180-1186; Kirschfink and Mollnes, Clin Diagn Lab Immunol. 2003, 10(6): 982-989; Nilsson and Ekdahl, Clinical and Developmental Immunology, 2012, Article ID 962702; which are hereby incorporated by reference in their entirety.

In some embodiments, the polypeptide according to the present invention possesses the ability to diffuse through, i.e. pass through, Bruch’s Membrane (BrM), as determined by analysis in an appropriate assay for said property.

- 5 The ability of a given polypeptide to diffuse through BrM can be analysed e.g. in vitro, e.g. as described in Clark et al J. Immunol (2014) 193, 4962-4970. Briefly, BrM can be isolated from donor eyes as described in McHarg et al., J Vis Exp (2015) 1-7, and the macular area can be mounted in an 10 Ussing chamber. Once mounted, the 5 mm diameter macular area is the only barrier between two identical compartments. Both sides of BrM can be washed with PBS, and human serum can be diluted 1:1 with PBS and added to the Ussing compartment on one side of the BrM (the sample chamber). 15 The polypeptide to be analysed can be added to the sample chamber in PBS, and PBS alone can be added to the compartment on the other side of the BrM (the diffusate chamber), and the Ussing chamber can be incubated at room temperature for 24 hours with gentle stirring in both the 20 sample and diffusate chambers. Samples from each chamber can subsequently be analysed for the presence of the polypeptide, e.g. using antibody based detection methods such as ELISA analysis or western blot. Detection of the polypeptide in the diffusate chamber indicates that the polypeptide 25 is capable of diffusing through BrM. Suitable positive and negative control proteins known to be able to/not to be able to diffuse through BrM can be included in such experiments.

In some embodiments, a polypeptide according to the present invention displays superior ability to diffuse through 30 BrM than Complement Factor I. In some embodiments, a polypeptide according to the present invention displays superior ability to diffuse through BrM than Complement Factor H. FH, consisting of 20 CCP domains, is a large molecule and does not pass through BrM. In some embodiments, a polypeptide according to the present invention displays similar ability to diffuse through BrM as compared to the truncated Complement Factor H isoform FHL-1 (UniProt: P08603-2; SEQ ID NO:28). In some embodiments, a polypeptide according to the present invention displays superior ability to diffuse through BrM as compared to Complement Factor H isoform FHL-1. In some embodiments, a polypeptide according to the present invention displays similar ability to diffuse through BrM as compared to full length soluble CR1 (30 CCP domains; SEQ ID NO:1 45 lacking SEQ ID NO:32 and 33). In some embodiments, a polypeptide according to the present invention displays superior ability to diffuse through BrM as compared to full length soluble CR1. A polypeptide according to the present invention that is able to diffuse though BrM preferably remains functionally active, i.e. acts as a cofactor for Complement Factor I, after diffusing though BrM.

A polypeptide of the present invention displaying superior ability to diffuse through BrM as compared to a given reference polypeptide can be identified by analysing diffusion through BrM as described above. The diffusion through BrM may be detected by measuring the rate of diffusion through to the diffusate chamber and/or detecting the proportion of polypeptide present in the diffusate chamber at the end of the experiment. A polypeptide of the present invention displaying similar ability to diffuse through BrM as compared to a given reference polypeptide can be identified by analysing diffusion through BrM as described above. A similar ability to diffuse through BrM may be indicated by detecting a rate of diffusion through to the diffusate chamber 60 which is within 30%, e.g. within one of 25%, 20%, 15%, or 10% of the rate of diffusion for a reference polypeptide, and/or by detecting a proportion of the polypeptide of the 65

41

present invention present in the diffusate chamber at the end of the experiment that is within 30%, e.g. within one of 25%, 20%, 15%, or 10% of the proportion of a reference polypeptide present in the diffusate chamber.

As a result of the ability of the polypeptides to diffuse through the BrM, the polypeptides of the present invention are also able to diffuse away from the site of C3b inactivation if/once they have performed their cofactor role with FI. In other words, the polypeptides of the present invention may be capable of being present transiently at areas of complement activation. This is advantageous because accumulation of complement-related debris is undesirable, particularly in the context of macular degeneration where cellular debris accumulation can lead to the formation of drusen.

A polypeptide of the present invention may be capable of being expressed in a cell, e.g. a cell as described herein. A polypeptide of the present invention may be capable of being secreted by a cell, e.g. a cell as described herein. In some embodiments the cell is an ocular cell, e.g. an RPE cell, as described herein.

Nucleic Acids, Cells, Compositions and Kits

The present invention provides a nucleic acid encoding a polypeptide according to the present invention. In some embodiments, the nucleic acid is purified or isolated, e.g. from other nucleic acid, or naturally-occurring biological material. In some embodiments the nucleic acid(s) comprise or consist of DNA and/or RNA.

Provided herein are nucleic acid sequences that encode a polypeptide comprising or consisting of SEQ ID NO:2, 3, 5, 6, 13, 14, 15, 30, 31, 40, 42, 44, 46, 47, 48, 49, 50, 51, 52, 53 or 54. The encoded polypeptide may be produced with or without a leader sequence e.g. a secretory pathway sequence.

The encoded polypeptide may be produced together with a leader sequence which is then subsequently removed from said polypeptide.

In some embodiments, a nucleic acid according to the present invention comprises, or consists of, one or more of SEQ ID NO:35, 36, 37, 38, 39, 41, 43, and/or 45, or equivalent nucleic acid sequences thereof which would be translated into the same respective polypeptides due to codon degeneracy.

The present invention also provides a vector comprising nucleic acid encoding a polypeptide according to the present invention.

The nucleotide sequence may be contained in a vector, e.g. an expression vector. A “vector” as used herein is a nucleic acid molecule used as a vehicle to transfer exogenous nucleic acid into a cell. The vector may be a vector for expression of the nucleic acid in the cell. Such vectors may include a promoter sequence operably linked to the nucleotide sequence encoding the sequence to be expressed. A vector may also include a termination codon and expression enhancers. A vector may include regulatory elements, such as a polyadenylation site. Any suitable vectors, promoters, enhancers and termination codons known in the art may be used to express a peptide or polypeptide from a vector according to the invention. Nucleic acid sequences described herein may be codon optimised for optimised expression in a desired cell or organism.

The term “operably linked” may include the situation where a selected nucleic acid sequence and regulatory nucleic acid sequence (e.g. promoter and/or enhancer) are covalently linked in such a way as to place the expression of nucleic acid sequence under the influence or control of the regulatory sequence (thereby forming an expression cas-

42

sette). Thus a regulatory sequence is operably linked to the selected nucleic acid sequence if the regulatory sequence is capable of effecting transcription of the nucleic acid sequence. The resulting transcript(s) may then be translated into a desired peptide(s)/polypeptide(s).

The nucleic acid and/or vector according to the present invention is preferably provided for introduction into a cell, e.g. a human cell. Suitable vectors include plasmids, binary vectors, DNA vectors, mRNA vectors, viral vectors (e.g. gammaretroviral vectors (e.g. murine Leukemia virus (MLV)-derived vectors), lentiviral vectors, retroviral vectors, adenovirus vectors, adeno-associated virus (AAV) vectors, vaccinia virus vectors and herpesvirus vectors, e.g. Herpes Simplex Virus vectors), transposon-based vectors, and artificial chromosomes (e.g. yeast artificial chromosomes), e.g. as described in Maus et al., *Annu Rev Immunol* (2014) 32:189-225 or Morgan and Boyerinas, *Biomedicines* 2016 4, 9, which are both hereby incorporated by reference in its entirety. In some embodiments, the lentiviral vector may be pELNS, or may be derived from pELNS. In some embodiments, the vector may be a vector encoding CRISPR/Cas9. In some embodiments, the adeno-associated virus (AAV) vector is selected from AAV serotype 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11, or hybrids and/or mutants thereof. In some embodiments, the AAV vector is an AAV serotype 2 (AAV-2) vector, or a hybrid and/or mutant thereof. Viral and non-viral delivery systems for introducing genetic material into cells are reviewed, for example, in Nayerossadat et al., *Adv Biomed Res.* 2012; 1: 27; MacLaren et al. *Ophthalmology.* 2016, 123(10 Suppl): S98-S106; Petit and Punzo, *Discov Med.* 2016, 22(121): 221-229; Aguirre, *Invest Ophthalmol Vis Sci.* 2017, 58(12): 5399-5411; Lundstrom, *Diseases.* 2018, 6(2): 42; which are hereby incorporated by reference in their entirety. Any suitable nucleotide or vector delivery method can be used in the context of the present invention.

In some embodiments the expression of a nucleic acid or a nucleic acid contained in a vector, according to the present invention, is driven by a promoter that drives expression in a specific retinal cell type, e.g. rods, cones, RPE, or ganglion cells, as described for example in Beltran W A, et al. *Gene Ther.* 2010; 17:1162-74 and Boye S E, et al. *Hum Gene Ther.* 2012; 23:1101-15, which are hereby incorporated by reference in their entirety.

In some embodiments, the expression of a nucleic acid or a nucleic acid contained in a vector, according to the present invention, is driven by a promoter that drives expression of that nucleic acid in retinal pigment epithelial (RPE) cells. In some embodiments, the promoter is an RPE65 or VMD2 promoter, or modified version thereof. In some embodiments the promoter is a chicken β actin promoter.

In some embodiments, the vector may be a eukaryotic vector, e.g. a vector comprising the elements necessary for expression of protein from the vector in a eukaryotic cell. In some embodiments, the vector may be a mammalian vector, e.g. comprising a cytomegalovirus (CMV) or SV40 promoter to drive protein expression.

In some embodiments the vector comprises an inducible promoter, i.e. gene expression is activated by the promoter only in the presence or absence of a particular molecule. Suitable inducible promoters will be known to the skilled person. Examples of inducible promoters are described in e.g. Le et al. *Invest Ophthalmol Vis Sci.* 2008, 49(3): 1248-1253 and McGee Sanftner et al. *Mol Ther.* 2001, 3(5): 688-696; which are hereby incorporated by reference in their entirety.

In some embodiments, the nucleic acid comprises, or consists of, a nucleic acid sequence encoding a polypeptide having an amino acid sequence having at least 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO:2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 30, 31, 40, 42, 44, 46, 47, 48, 49, 50, 51, 52, 53, or 54, or any combination of sequences A, B, and/or C as described hereinabove.

The present invention also provides a cell comprising or expressing a polypeptide according to the present invention. Also provided is a cell comprising or expressing a nucleic acid or vector according to the invention. The cell comprising or expressing polypeptide, nucleic acid or vector according to the present invention may secrete a polypeptide according to the present invention. That is, expression of the polypeptide, nucleic acid or vector by the cell may result in the soluble production of a polypeptide according to the present invention from the cell.

The cell may be a eukaryotic cell, e.g. a mammalian cell. The mammal may be a human, or a non-human mammal (e.g. rabbit, guinea pig, rat, mouse or other rodent (including any animal in the order Rodentia), cat, dog, pig, sheep, goat, cattle (including cows, e.g. dairy cows, or any animal in the order Bos), horse (including any animal in the order Equidae), donkey, and non-human primate). In some embodiments, the cell may be from, or may have been obtained from, a human subject.

In some embodiments, the cell is a cell of the eye. In some embodiments, the cell is a cell of the neurosensory retina, retinal pigment epithelium (RPE), choroid or macula. In some embodiments, the cell is a retinal cell. In some embodiments, the cell is a retinal pigment epithelial cell. In some embodiments, the cell is a human retinal pigment epithelial cell (RPE). In some embodiments the cell is a photoreceptor cell.

The present invention also provides a method for producing a cell comprising a nucleic acid or vector according to the present invention, comprising introducing a nucleic acid or vector according to the present invention into a cell. In some embodiments, introducing an isolated nucleic acid(s) or vector(s) according to the invention into a cell comprises transformation, transfection, electroporation or transduction (e.g. retroviral transduction). Methods for producing a cell according to the present invention may be performed according to methods known to the skilled person for the production of cells comprising nucleic acid/vectors.

The present invention also provides a method for producing a cell comprising or expressing a polypeptide according to the present invention, comprising introducing a nucleic acid or vector according to the present invention into a cell. In some embodiments, the methods additionally comprise culturing the cell under conditions suitable for expression of the nucleic acid or vector by the cell. In some embodiments, the methods are performed *in vitro* or *ex vivo*. In some embodiments, the methods are performed *in vivo*.

The present invention also provides cells obtained or obtainable by the methods according to the present invention.

The present invention also provides compositions comprising a polypeptide, nucleic acid, vector, or cell according to the present invention.

Polypeptides, nucleic acids, vectors and cells according to the present invention may be formulated as pharmaceutical compositions for clinical use and may comprise a pharmaceutically acceptable carrier, diluent, excipient or adjuvant.

In accordance with the present invention methods are also provided for the production of pharmaceutically useful compositions, such methods of production may comprise one or more steps selected from: isolating a polypeptide, cell, nucleic acid or vector as described herein; and/or mixing a polypeptide, cell, nucleic acid or vector as described herein with a pharmaceutically acceptable carrier, adjuvant, excipient or diluent.

A kit of parts is also provided. In some embodiments the kit may have at least one container having a predetermined quantity of a polypeptide, nucleic acid, vector, cell, and/or composition according to the present invention.

The kit may provide the polypeptide, nucleic acid, vector, cell or composition together with instructions for administration to a subject in order to treat a specified disease/condition. The polypeptide, nucleic acid, vector, cell or composition may be formulated so as to be suitable for injection or infusion. In some embodiments, the polypeptide, nucleic acid, vector, cell or composition may be formulated so as to be suitable for intravenous, intraocular, sub-retinal, suprachoroidal or intraconjunctival injection, administration as an eye drop (i.e. ophthalmic administration), or oral administration.

In some embodiments the kit may comprise materials for producing a cell according to the present invention. For example, the kit may comprise materials for modifying a cell to express or comprise a polypeptide, nucleic acid or vector according to the present invention, or materials for introducing into a cell the nucleic acid or vector according to the present invention.

In some embodiments the kit may further comprise at least one container having a predetermined quantity of another therapeutic agent (e.g. a therapeutic agent for the treatment of AMD). In such embodiments, the kit may also comprise a second medicament or pharmaceutical composition such that the two medicaments or pharmaceutical compositions may be administered simultaneously or separately such that they provide a combined treatment for the specific disease or condition. In some embodiments, the second medicament or pharmaceutical composition comprises Complement Factor I.

Producing Polypeptides

The present invention also provides a method for producing a polypeptide according to the present invention, the method comprising introducing into a cell a nucleic acid or vector according to the present invention, and culturing the cell under conditions suitable for expression of the polypeptide.

The polypeptide may be a fusion protein. The polypeptide may be subsequently isolated and/or substantially purified.

Polypeptides according to the present invention may be prepared according to methods for the production of polypeptides known to the skilled person.

Polypeptides may be prepared by chemical synthesis, e.g. liquid or solid phase synthesis. For example, peptides/polypeptides can be synthesised using the methods described in, for example, Chandru et al., *Molecules* (2013), 18: 4373-4388, which is hereby incorporated by reference in its entirety.

Alternatively, polypeptides may be produced by recombinant expression. Molecular biology techniques suitable for recombinant production of polypeptides are well known in the art, such as those set out in Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th Edition), Cold Spring Harbor Press, 2012, and in *Nat Methods*. (2008); 5(2): 135-146 both of which are hereby incorporated by reference in their entirety.

For recombinant production according to the invention, any cell suitable for the expression of polypeptides may be used. The cell may be a prokaryote or eukaryote. In some embodiments the cell is a prokaryotic cell, such as a cell of archaea or bacteria. In some embodiments the bacteria may be Gram-negative bacteria such as bacteria of the family Enterobacteriaceae, for example *Escherichia coli*. In some embodiments, the cell is a eukaryotic cell such as a yeast cell, a plant cell, insect cell or a mammalian cell, e.g. CHO, HEK (e.g. HEK293), HeLa or COS cells.

In some cases the cell is not a prokaryotic cell because some prokaryotic cells do not allow for the same folding or post-translational modifications as eukaryotic cells. In addition, very high expression levels are possible in eukaryotes and proteins can be easier to purify from eukaryotes using appropriate tags. Specific plasmids may also be utilised which enhance secretion of the protein into the media.

In some embodiments polypeptides may be prepared by cell-free-protein synthesis (CFPS), e.g. according using a system described in Zemella et al. *Chembiochem* (2015) 16(17): 2420-2431, which is hereby incorporated by reference in its entirety.

Production may involve culture or fermentation of a eukaryotic cell modified to express the polypeptide(s) of interest. The culture or fermentation may be performed in a bioreactor provided with an appropriate supply of nutrients, air/oxygen and/or growth factors. Secreted proteins can be collected by partitioning culture media/fermentation broth from the cells, extracting the protein content, and separating individual proteins to isolate secreted polypeptide(s). Culture, fermentation and separation techniques are well known to those of skill in the art, and are described, for example, in Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th Edition; incorporated by reference herein above).

Bioreactors include one or more vessels in which cells may be cultured. Culture in the bioreactor may occur continuously, with a continuous flow of reactants into, and a continuous flow of cultured cells from, the reactor. Alternatively, the culture may occur in batches. The bioreactor monitors and controls environmental conditions such as pH, oxygen, flow rates into and out of, and agitation within the vessel such that optimum conditions are provided for the cells being cultured.

Following culturing the cells that express the antigen-binding molecule/polypeptide(s), the polypeptide(s) of interest may be isolated. Any suitable method for separating proteins from cells known in the art may be used. In order to isolate the polypeptide it may be necessary to separate the cells from nutrient medium. If the polypeptide(s) are secreted from the cells, the cells may be separated by centrifugation from the culture media that contains the secreted polypeptide(s) of interest. If the polypeptide(s) of interest collect within the cell, protein isolation may comprise centrifugation to separate cells from cell culture medium, treatment of the cell pellet with a lysis buffer, and cell disruption e.g. by sonication, rapid freeze-thaw or osmotic lysis.

It may then be desirable to isolate the polypeptide(s) of interest from the supernatant or culture medium, which may contain other protein and non-protein components. A common approach to separating protein components from a supernatant or culture medium is by precipitation. Proteins of different solubilities are precipitated at different concentrations of precipitating agent such as ammonium sulfate. For example, at low concentrations of precipitating agent, water soluble proteins are extracted. Thus, by adding dif-

ferent increasing concentrations of precipitating agent, proteins of different solubilities may be distinguished. Dialysis may be subsequently used to remove ammonium sulfate from the separated proteins.

Other methods for distinguishing different proteins are known in the art, for example ion exchange chromatography and size chromatography. The polypeptide may also be affinity-purified using an appropriate binding partner for a molecular tag on the polypeptide (e.g. a His, FLAG, Myc, GST, MBP, HA, E, or Biotin tag). These may be used as an alternative to precipitation, or may be performed subsequently to precipitation.

In some cases it may further be desired to process the polypeptide, e.g. to remove a sequence of amino acids, molecular tag, moiety, etc.

In some embodiments, treatment is with an appropriate endopeptidase for the cleavage and removal of an amino acid sequence.

In some embodiments, treatment is with an enzyme to remove the moiety of interest. In some embodiments, the polypeptide is treated to remove glycans (i.e. the polypeptide is deglycosylated), e.g. by treatment with a glycosidase such as with a Peptide:N-glycosidase (PNGase).

Once the polypeptide(s) of interest have been isolated from culture it may be desired or necessary to concentrate the polypeptide(s). A number of methods for concentrating proteins are known in the art, such as ultrafiltration or lyophilisation.

In some embodiments, the production of the polypeptide occurs in vivo, e.g. after introduction to the host of a cell comprising a nucleic acid or vector encoding a polypeptide of the present invention, or following introduction into a cell of the host of a nucleic acid or vector encoding a polypeptide of the present invention. In such embodiments, the polypeptide is transcribed, translated and post-translationally processed to the mature polypeptide. In some embodiments, the polypeptide is produced in situ at the desired location in the host. In some embodiments, the desired location is the eye, e.g. in a cell of the retina, choroid, retinal pigment epithelium (RPE) or macula. In some embodiments, the desired location is at or in a retinal cell. In some embodiments, the desired location is at or in a RPE cell.

Therapeutic Applications

Any of the polypeptides, nucleic acids, vectors, cells and pharmaceutical compositions according to the present invention find use in therapeutic and prophylactic methods.

The present invention provides a polypeptide, nucleic acid, vector, cell, or pharmaceutical composition according to the present invention, for use in a method of medical treatment or prophylaxis. The present invention also provides the use of a polypeptide, nucleic acid, vector, cell or pharmaceutical composition according to the present invention in the manufacture of a medicament for treating or preventing a disease or condition. The present invention also provides a method of treating or preventing a disease or condition, comprising administering to a subject a therapeutically or prophylactically effective amount of a polypeptide, nucleic acid, vector, cell or pharmaceutical composition according to the present invention.

In particular, the polypeptides, nucleic acids, vectors, cells and pharmaceutical compositions according to the present invention find use to treat or prevent diseases/conditions associated with complement dysregulation, in particular overactive complement response. In some embodiments, the overactive complement response is linked to the presence of C3b. In some embodiments the disease/condition to be treated or prevented is a complement-related

disease. In some embodiments the disease/condition to be treated or prevented is pathologically associated with complement activation. In some embodiments the disease/condition to be treated or prevented is pathologically associated with complement over-activation. In some embodiments the disease/condition to be treated or prevented is driven by complement activation or over-activation. In some embodiments the disease/condition is complement activation or over-activation.

The polypeptides, nucleic acids, vectors, cells and pharmaceutical compositions find use to treat or prevent diseases/conditions which would benefit from one or more of: a reduction in the level or activity of C3bBb-type C3 convertase, C3bBb3b-type C5 convertase or C4b2a3b-type C5 convertase; a reduction in the level of C3b, C5b or C5a; an increase in the level of iC3b, C3f, C3dg or C3d; or a reduction in the level or activity of iC3b and an increase in the level of C3f, C3dg or C3d.

'Treatment' may, for example, be reduction in the development or progression of a disease/condition, alleviation of the symptoms of a disease/condition or reduction in the pathology of a disease/condition. Treatment or alleviation of a disease/condition may be effective to prevent progression of the disease/condition, e.g. to prevent worsening of the condition or to slow the rate of development. In some embodiments treatment or alleviation may lead to an improvement in the disease/condition, e.g. a reduction in the symptoms of the disease/condition or reduction in some other correlate of the severity/activity of the disease/condition. Prevention/prophylaxis of a disease/condition may refer to prevention of a worsening of the condition or prevention of the development of the disease/condition, e.g. preventing an early stage disease/condition developing to a later, chronic, stage.

In some embodiments, the disease or condition to be treated or prevented may be a disease/condition associated with C3b or a C3b-containing complex, an activity/response associated with C3b or a C3b-containing complex, or a product of an activity/response associated with C3b or a C3b-containing complex. That is, in some embodiments, the disease or condition to be treated or prevented is a disease/condition in which C3b, a C3b-containing complex, an activity/response associated with C3b or a C3b-containing complex, or the product of said activity/response is pathologically implicated. In some embodiments, the disease/condition may be associated with an increased level of C3b or a C3b-containing complex, an increased level of an activity/response associated with C3b or a C3b-containing complex, or increased level of a product of an activity/response associated with C3b or a C3b-containing complex as compared to the control state.

The treatment may be aimed at reducing the level of C3b or a C3b-containing complex, an activity/response associated with C3b or a C3b-containing complex, or a product of an activity/response associated with C3b or a C3b-containing complex. In some embodiments, the treatment is aimed at: reducing the level or activity of C3bBb-type C3 convertase, C3bBb3b-type C5 convertase or C4b2a3b-type C5 convertase; reducing the level of C3b, C5b or C5a; increasing the level of iC3b, C3f, C3dg or C3d, or reducing the level of iC3b and increasing the level of C3f, C3dg or C3d.

Administration of the polypeptides, nucleic acids, vectors, cells and compositions of the present invention may cause a reduction in the level of C3b or a C3b-containing complex, an activity/response associated with C3b or a C3b-contain-

ing complex, or a product of an activity/response associated with C3b or a C3b-containing complex through cleavage of C3b.

In some embodiments, the treatment may be aimed at 5 reducing the level of C3b or a C3b-containing complex, an activity/response associated with C3b or a C3b-containing complex, or a product of an activity/response associated with C3b or a C3b-containing complex in a subject, e.g. at a particular location, in a particular organ, tissue, structure or 10 cell type. In some embodiments, the treatment may be aimed at reducing the level of C3b or a C3b-containing complex, an activity/response associated with C3b or a C3b-containing complex, or a product of an activity/response associated with C3b or a C3b-containing complex in the eye, e.g. in the 15 retina, choroid, RPE, macula and/or at the BrM/RPE interface.

In some embodiments, the treatment may comprise modifying a cell or population of cells to comprise/express a polypeptide, nucleic acid or vector of the present invention.

20 In some embodiments, the treatment may comprise modification of the cell/population in vivo, for in situ production of the polypeptide of the invention. In some embodiments the cell/population of cells is/are an ocular cell/cells. In some embodiments, the cell/population of cells is a RPE cell and/or population of RPE cells. In some embodiments the cell/population of cells is a photoreceptor cell and/or population of photoreceptor cells.

In some embodiments, the present invention provides a nucleic acid or vector of the present invention for use in gene

30 therapy. In some embodiments, the treatment comprises administering the nucleic acid and/or vector to a subject. In some embodiments, the treatment comprises introducing the nucleic acid and/or vector into a cell of a subject, using techniques described herein or well known in the art, see e.g.

35 MacLaren et al. Ophthalmology. 2016, 123(10 Suppl): S98-S106; Aguirre, Invest Ophthalmol Vis Sci. 2017, 58(12): 5399-5411; Lundstrom, Diseases. 2018, 6(2): 42; which are hereby incorporated by reference in their entirety. In some embodiments the cell is an ocular cell or cells. In some

40 embodiments, the cell is an RPE cell or cells. In some embodiments the cell is a photoreceptor cell or cells.

In some embodiments, the treatment may comprise administering to a subject a cell or population of cells modified to comprise/express a polypeptide, nucleic acid or vector of the present invention. In some embodiments, the treatment may comprise modification of the cell/population

45 ex vivo or in vitro.

In some embodiments, the treatment is aimed at providing the subject with a cell or population of cells which produce and/or will produce the polypeptide of the invention, e.g. by administering a cell according to the present invention, or by generating a cell according to the present invention.

In some embodiments, the cell referred to herein is a cell of the eye i.e. an ocular cell. In some embodiments, the cell 55 is a cell of the retina, choroid, retinal pigment epithelium (RPE) or macula. In some embodiments, the cell is a retinal cell. In some embodiments, the cell is an RPE cell. In some embodiments the cell is a photoreceptor cell.

The present invention provides a method of treating or 60 preventing a disease or condition in a subject, the method comprising modifying at least one cell to express or comprise a polypeptide, nucleic acid or vector according to the present invention. In some embodiments the at least one cell is an ocular cell. In some embodiments, the at least one cell is an RPE cell.

The at least one cell modified according to the present invention can be modified according to methods well known

to the skilled person. The modification may comprise nucleic acid transfer for permanent or transient expression of the transferred nucleic acid. Any suitable genetic engineering platform may be used to modify a cell according to the present invention. Suitable methods for modifying a cell include the use of genetic engineering platforms such as gammaretroviral vectors, lentiviral vectors, adenovirus vectors, adeno-associated virus (AAV) vectors, DNA transfection, transposon-based gene delivery and RNA transfection, for example as described in Maus et al., *Annu Rev Immunol* (2014) 32:189-225, incorporated by reference hereinabove.

The subject to be treated may be any animal or human. The subject is preferably mammalian, more preferably human. The subject may be a non-human mammal, but is more preferably human. The subject may be male or female. The subject may be a patient. A subject may have been diagnosed with a disease or condition requiring treatment, or be suspected of having such a disease or condition.

The subject to be treated may display an elevated level of C3b or a C3b-containing complex, an activity/response associated with C3b or a C3b-containing complex, or a product of an activity/response associated with C3b or a C3b-containing complex, e.g. as determined by analysis of the subject, or a sample (e.g. a cell, tissue, blood sample) obtained from the subject, using an appropriate assay.

The subject may have an increased level of expression or activity of a positive regulator/effector of C3b or a C3b-containing complex or of an activity/response associated with C3b or a C3b-containing complex, or may have an increased level of expression or activity of a product of an activity/response associated with C3b or a C3b-containing complex. The subject may have an increased level of an activity upregulated by C3b or a C3b-containing complex.

The subject may have a reduced level of expression or activity of a negative regulator of C3b or a C3b-containing complex or of an activity/response associated with C3b or a C3b-containing complex, or may have a reduced level of expression or activity a factor downregulated by C3b or a C3b-containing complex. The subject may have a reduced level of an activity downregulated by C3b or a C3b-containing complex.

The increase/reduction may be relative to the level of expression/activity in the absence of the relevant disease/condition, e.g. the level of expression/activity in a healthy control subject or sample obtained from a healthy control subject.

In some embodiments, the subject may be at risk of developing/contracting a disease or condition. In some embodiments, the subject may possess one or more predisposing factors increasing risk of developing/contracting a disease or condition.

In some embodiments, the subject may possess one or more risk factors for Age-related Macular Degeneration (AMD). In some embodiments, the subject may possess one or more of AMD-associated genetic variants. AMD-associated genetic variants are described e.g. in Clark et al., *J Clin Med* (2015) 4(1):18-31, which is hereby incorporated by reference in its entirety. In some embodiments, the subject may possess one or more of the following AMD-associated genetic variants (or a variant having $LD=r^2 \geq 0.8$ with such variant): Y402H in CFH (i.e. rs1061170^C), rs1410996^C, 162V in CFH, R53C in CFH, D90G in CFH, R1210C in CFH, or rs6685931^T in CFHR4.

In some embodiments the subject may possess one or more risk factors for early-onset macular degeneration (EOMD). EOMD is thought to be caused by monogenic inheritance of rare variants of the CFH gene (see e.g. Boon

C J et al. *Am J Hum Genet* 2008; 82(2):516-23; van de Ven J P, et al. *Arch Ophthalmol* 2012; 130(8):1038-47; Yu Y et al. *Hum Mol Genet* 2014; 23(19):5283-93; Duvvari M R, et al. *Mol Vis* 2015; 21:285-92; Hughes A E, et al. *Acta Ophthalmol* 2016; 94(3):e247-8; Wagner et al. *Sci Rep* 2016; 6:31531). In some embodiments, the subject may possess one or more of EOMD-associated genetic variants. EOMD-associated genetic variants are described in e.g. Servais A et al. *Kidney Int*, 2012; 82(4):454-64 and Dragon-Durey M A, et al. *J Am Soc Nephrol* 2004; 15(3):787-95; which are hereby incorporated by reference in their entirety. In some embodiments, the subject may possess one or more of the following EOMD-associated genetic variants: CFH c.1243del, p.(Ala415Profs*39) het; CFH c.350+1G>T het; CFH c.619+1G>A het; CFH c.380G>A, p.(Arg127His); CFH c.694C>T, p.(Arg232Ter); or CFH c.1291T>A, p.(Cys431Ser).

In some embodiments, the subject is selected for therapeutic or prophylactic treatment with the polypeptide, nucleic acid, vector, cell or composition of the present invention based on their being determined to possess one or more risk factors for AMD and/or EOMD, e.g. one or more AMD/EOMD-associated genetic variants. In some embodiments, the subject has been determined to have one or more such risk factors. In some embodiments, the methods of the present invention involving determining whether a subject possesses one or more such risk factors.

In some embodiments, the disease or condition to be treated or prevented may be an ocular disease/condition. In some embodiments, the disease or condition to be treated or prevented is a complement-related ocular disease. In some embodiments, the disease or condition to be treated or prevented is macular degeneration. In some embodiments, the disease or condition to be treated or prevented is age-related macular degeneration (AMD). AMD is commonly-defined as causing vision loss in subjects age 50 and older.

In some embodiments, the disease or condition to be treated or prevented is selected from age-related macular degeneration (AMD), early AMD, intermediate AMD, late AMD, geographic atrophy ('dry' (i.e. non-exudative) AMD), 'wet' (neovascular or exudative) AMD, choroidal neovascularisation (CNV), glaucoma, autoimmune uveitis, and diabetic retinopathy. In some embodiments, the disease or condition to be treated or prevented is AMD. In some embodiments, the disease or condition to be treated or prevented is geographic atrophy ('dry' AMD). In some embodiments, the disease or condition to be treated or prevented is 'wet' AMD. In some embodiments the disease or condition to be treated or prevented is a combination of the diseases/conditions above, e.g. 'dry' and 'wet' AMD. In some embodiments the disease or condition to be treated or prevented is not 'wet' AMD or choroidal neovascularisation. In some embodiments a subject to be treated is age 50 or older, i.e. is at least 50 years old.

As used herein "early AMD" refers to a stage of AMD characterised by the presence of medium-sized drusen, commonly having a width of up to ~200 µm, within the Bruch's membrane adjacent to the RPE layer. Subjects with early AMD typically do not present significant vision loss. As used herein "intermediate AMD" refers to a stage of AMD characterised by large drusen and/or pigment changes in the retina. Intermediate AMD may be accompanied by some vision loss. As used herein "late AMD" refers to a stage of AMD characterised by the presence of drusen and vision loss due to damage to the macula. In all stages of AMD, 'reticular pseudodrusen' (RPD) or 'reticular drusen' may be present, referring to the accumulation of extracellular mate-

rial in the subretinal space between the neurosensory retina and RPE. "Late AMD" encompasses 'dry' and 'wet' AMD. In 'dry' AMD (also known as geographic atrophy), there is a gradual breakdown of the light-sensitive cells in the macula that convey visual information to the brain and of the supporting tissue beneath the macula. In 'wet' AMD (also known as choroidal neovascularization and exudative AMD), abnormal blood vessels grow underneath and into the retina. These vessels can leak fluid and blood which can lead to swelling and damage of the macula and subsequent scar formation. The damage may be rapid and severe.

In some embodiments the disease or condition to be treated or prevented is early-onset macular degeneration (EOMD). As used herein "EOMD" refers to a phenotypically severe sub-type of macular degeneration that demonstrates a much earlier age of onset than classical AMD and results in many more years of substantial visual loss. The EOMD subset is described in e.g. Boon C J et al. Am J Hum Genet 2008; 82(2):516-23 and van de Ven J P, et al. Arch Ophthalmol 2012; 130(8):1038-47. In some embodiments a subject to be treated is age 49 or younger. In some embodiments a subject to be treated is between ages 15 and 49, i.e. is between 15 and 49 years old.

In some embodiments, the disease or condition to be treated or prevented is a disease/condition driven by complement over-activation. In some embodiments, the disease or condition to be treated or prevented may be selected from atypical Haemolytic Uremic Syndrome (aHUS), Membranoproliferative Glomerulonephritis Type II (MPGN II), sepsis, and Paroxysmal nocturnal hemoglobinuria (PNH).

Methods of medical treatment may also involve *in vivo*, *ex vivo*, and adoptive immunotherapies, including those using autologous and/or heterologous cells or immortalized cell lines.

Administration of a polypeptide described herein is preferably in a "therapeutically effective amount", this being sufficient to show benefit to the individual. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of the disease being treated. Prescription of treatment, e.g. decisions on dosage etc., is within the responsibility of general practitioners and other medical doctors, and typically takes account of the condition to be treated, the condition of the individual patient, the site of delivery, the method of administration and other factors known to practitioners. Examples of the techniques and protocols mentioned above can be found in Remington's Pharmaceutical Sciences, 20th Edition, 2000, pub. Lippincott, Williams & Wilkins.

Polypeptides, nucleic acids, vectors and cells according to the present invention may be formulated as pharmaceutical compositions or medicaments for clinical use and may comprise a pharmaceutically acceptable carrier, diluent, excipient or adjuvant. The composition may be formulated for topical, parenteral, systemic, intracavitary, intravenous, intra-arterial, intramuscular, intrathecal, intraocular, intrconjunctival, subretinal, suprachoroidal, subcutaneous, intra-dermal, intrathecal, oral or transdermal routes of administration which may include injection or infusion, or administration as an eye drop (i.e. ophthalmic administration). Suitable formulations may comprise the polypeptide, nucleic acid, vector, or cell in a sterile or isotonic medium. Medicaments and pharmaceutical compositions may be formulated in fluid, including gel, form. Fluid formulations may be formulated for administration by injection or infusion (e.g. via catheter) to a selected organ or region of the human or animal body. In some embodiments the polypeptides, nucleic acids, vector, cells and compositions of the

invention are formulated for intravitreal routes of administration e.g. by intravitreal injection. In some embodiments polypeptides, nucleic acids, vector, cells and compositions of the invention are formulated for submacular delivery i.e. placing the therapeutic molecules in direct contact with the target cell layers.

The particular mode and/or site of administration may be selected in accordance with the location where the C3b inactivation is desired. In some embodiments, the polypeptides, nucleic acids, vectors, or pharmaceutical compositions of the present invention are formulated for administration and/or administered into the subretinal space between the photoreceptor cells and the retinal pigment epithelium (RPE) in the eye. In some embodiments, the polypeptides, nucleic acids, vectors, or pharmaceutical compositions of the present invention are formulated for administration and/or administered into the retinal pigment epithelium (RPE).

In accordance with the present invention methods are also provided for the production of pharmaceutically useful compositions, such methods of production may comprise one or more steps selected from: isolating a polypeptide, nucleic acid, vector, or cell as described herein; and/or mixing a polypeptide, nucleic acid, vector, or cell as described herein with a pharmaceutically acceptable carrier, adjuvant, excipient or diluent.

For example, a further aspect of the present invention relates to a method of formulating or producing a medicament or pharmaceutical composition for use in a method of medical treatment, the method comprising formulating a pharmaceutical composition or medicament by mixing polypeptide, nucleic acid, vector, or cell as described herein with a pharmaceutically acceptable carrier, adjuvant, excipient or diluent.

Administration may be alone or in combination with other treatments (e.g. other therapeutic or prophylactic intervention), either simultaneously or sequentially dependent upon the condition to be treated. The polypeptide, nucleic acid, vector, cell or composition according to the present invention and a therapeutic agent may be administered simultaneously or sequentially.

Simultaneous administration refers to administration of the polypeptide, nucleic acid, vector, cell or composition and therapeutic agent together, for example as a pharmaceutical composition containing both agents (combined preparation), or immediately after each other and optionally via the same route of administration, e.g. to the same tissue, artery, vein or other blood vessel. Sequential administration refers to administration of one of the polypeptide, nucleic acid, vector, cell or composition or therapeutic agent followed after a given time interval by separate administration of the other agent. It is not required that the two agents are administered by the same route, although this is the case in some embodiments. The time interval may be any time interval. In some embodiments, the polypeptide, nucleic acid, vector, cell or composition and therapeutic agent are administered separately, simultaneously or sequentially to the eye.

In some embodiments, the other treatment/therapeutic agent is a therapeutically effective amount of Complement Factor I. In some embodiments, Complement Factor I is administered to the subject simultaneously or sequentially with administration of a polypeptide, nucleic acid, vector, cell, or pharmaceutical composition according to the present invention. In some embodiments, the treatment may comprise modifying a cell or population of cells *in vitro*, *ex vivo* or *in vivo* to express and/or secrete Complement Factor I.

The cell or population of cells may be the same cell or population of cells as a cell or population of cells modified to comprise/express a polypeptide, nucleic acid or vector according to the present invention, for example the treatment may comprise modifying a cell or population of cells *in vitro*, *ex vivo* or *in vivo* to express and/or secrete a polypeptide, nucleic acid or vector according to the present invention, and Complement Factor I. In some embodiments, Complement Factor I is administered to a subject, wherein the subject comprises a cell or population of cells modified to comprise/express a polypeptide, nucleic acid or vector of the present invention. In some embodiments, Complement Factor I is administered to a subject wherein the subject has expressed *in situ* or is expressing *in situ* a polypeptide, nucleic acid or vector of the present invention.

Complement Factor I, or a composition comprising Complement Factor I, may be formulated for topical, parenteral, systemic, intracavitary, intravenous, intravitreal, intra-arterial, intramuscular, intrathecal, intraocular, intrconjunctival, subretinal, suprachoroidal, subcutaneous, intra-dermal, intrathecal, oral or transdermal routes of administration which may include injection or infusion, or administration as an eye drop (i.e. ophthalmic administration). Suitable formulations may comprise a sterile or isotonic medium. Medicaments and pharmaceutical compositions may be formulated in fluid, including gel, form. Fluid formulations may be formulated for administration by injection or infusion (e.g. via catheter) to a selected organ or region of the human or animal body.

In some embodiments, the other treatment/therapeutic agent is a therapeutically effective amount of an anti-VEGF therapy (e.g. ranibizumab (Lucentis; Genentech/Novartis), bevacizumab (off label Avastin; Genentech), afiblivercept (Eylea/VEGF Trap-Eye; Regeneron/Bayer)), pegaptanib (Macugen[®]), laser photocoagulation, or photodynamic therapy (PDT) e.g. with VisudyneTM (verteporfin).

Multiple doses of the polypeptide, nucleic acid, vector, cell or composition may be provided. One or more, or each, of the doses may be accompanied by simultaneous or sequential administration of another therapeutic agent.

Multiple doses may be separated by a predetermined time interval, which may be selected to be one of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or 31 days, or 1, 2, 3, 4, 5, or 6 months. By way of example, doses may be given once every 7, 14, 21 or 28 days (plus or minus 3, 2, or 1 days).

A polypeptide, nucleic acid, vector or composition according to the present invention may be formulated in a sustained release delivery system, in order to release the polypeptide, nucleic acid, vector or composition at a predetermined rate. Sustained release delivery systems may maintain a constant drug/therapeutic concentration for a specified period of time. In some embodiments, a polypeptide, nucleic acid, vector or composition according to the present invention is formulated in a liposome, gel, implant, device, or drug-polymer conjugate e.g. hydrogel.

Genetic Factors in Age-Related Macular Degeneration (AMD)

Complement Factor H (encoded by the CFH gene) is another co-factor for Complement Factor I. Complement Factor H structure and function is reviewed e.g. in Wu et al., Nat Immunol (2009) 10(7): 728-733, which is hereby incorporated by reference in its entirety. Human Complement Factor H (UniProt: P08603; SEQ ID NO:29) has a 1,233 amino acid sequence (including an N-terminal, 18 amino acid signal peptide), and comprises 20 complement control protein (CCP) domains. The first four CCP domains (i.e.

CCP1 to CCP4) of Complement Factor H are necessary for Complement Factor I co-factor activity for cleavage of C3b to iC3b. CCPs 19 to 20 have also been shown to engage with C3b and C3d (Morgan et al., Nat Struct Mol Biol (2011) 18(4): 463-470), whilst CCP7 and CCPs 19 to 20 bind to glycosaminoglycans (GAGs) and sialic acid, and are involved in discrimination between self and non-self (Schmidt et al., J Immunol (2008) 181(4): 2610-2619; Kajander et al., PNAS (2011) 108(7): 2897-2902).

One of the major SNPs associated with genetic risk of developing AMD is found in the CFH gene and leads to the Y402H polymorphism in Complement Factor H (see e.g. Haines et al., Science (2005) 308:419-21), and it's alternative splice variant factor H-like protein 1 (FHL-1). Around 30% of individuals of white European heritage have at least one copy of this polymorphism, whilst being a heterozygote increases the risk of AMD by ~3-fold (Sofat et al., Int J Epidemiol (2012) 41:250-262). The Y402H polymorphism, which manifests in the seventh complement control protein (CCP) domain, reduces the binding of FH/FHL-1 to BrM, leading to perturbations in the binding of these blood-borne complement regulators and dampened complement regulation on this surface (Clark et al., J Biol Chem (2010) 285:30192-202).

The binding of FH/FHL-1 to BrM is mediated by sulphated sugars including the glycosaminoglycans (GAGs) heparan sulphate (HS) and dermatan sulphate (DS). The family of GAG sequences found in BrM appears to have greater tissue specificity than previously thought, as they are able to recruit FH/FHL-1 through their CCP7 domains and not FH's secondary anchoring site found in CCPs19-20 (Clark et al., J Immunol (2013) 190:2049-2057). This is likely to be an evolutionary twist, as it has been discovered that the main regulator of complement within BrM is the truncated FHL-1 protein (Clark et al. J. Immunol (2014) 193, 4962-4970), which only has the one surface anchoring site in CCP7 and lacks CCPs19-20. In contrast, the Y402H polymorphism is not associated with kidney disease where the CCP19-20 domain of FH is known to be the main GAG-mediated anchoring site (Clark et al., J Immunol (2013) 190:2049-2057). Age-related changes in the BrM expression levels of HS and DS, themselves considered part of the normal ageing process, have also been associated with AMD, and may go some way as to explain the age-related nature of the genetically driven AMD.

A rare mutation (R1210C) in the C-terminal CCP19-20 region of FH, which does not bind to BrM, has a very high level of association with AMD, and FH protein carrying this mutation is found covalently bound to albumin (Sánchez-Corral et al., Am J Hum Genet (2002) 71:1285-1295) preventing the FH protein from leaving the circulation and entering eye tissue. Some research suggests that the large confluent drusen that precede geographic atrophy and the associated pigmentary changes in the RPE indicate that dry AMD results firstly from dysfunction of the RPE with secondary effects within the choroid (Bhutto and Lutty Mol Aspects Med (2012) 33:295-317). In contrast, Whitmore et al. reported changes in the choriocapillaris preceding all forms of late-stage AMD including the deposition of the terminal complement membrane attack complex (MAC), and argue that excessive complement activation in the choriocapillaris is the primary event with RPE atrophy being secondary (Whitmore et al., Prog Retin Eye Res (2015) 45:1-29). These data imply that a genetic predisposition conferred by alterations in complement genes is tolerated until changes in both BrM and the underlying choriocapillaris come to the fore. Whether these changes are age-

related, driven by oxidative stress or a result of RPE cell dysfunction remains to be seen, but naturally occurring changes in these structures are known to be age-related.

C3 and C3b

Complement component 3 (C3) is an immune system protein having a central role in innate immunity and the complement system. Processing of C3 is described, for example, in Foley et al. J Thromb Haemostasis (2015) 13: 610-618, which is hereby incorporated by reference in its entirety. Human C3 (UniProt: P01024; SEQ ID NO:18) comprises a 1,663 amino acid sequence (including an N-terminal, 22 amino acid signal peptide). Amino acids 23 to 667 encode C3 β chain (SEQ ID NO:19), and amino acids 749 to 1,663 encode C3 α' chain (SEQ ID NO:20). C3 β chain and C3 α' chain associate through interchain disulphide bonds (formed between cysteine 559 of C3 β chain, and cysteine 816 of the C3 α' chain) to form C3b. C3a is a 77 amino acid fragment corresponding to amino acid positions 672 to 748 of C3 (SEQ ID NO:21), generated by proteolytic cleavage of C3 following activation through the classical complement pathway and the lectin pathways.

C3b is a potent opsonin, targeting pathogens, antibody antigen immune complexes and apoptotic cells for phagocytosis by phagocytes and NK cells. C3b is also involved in the formation of convertase enzyme complexes for activating and amplifying complement responses. C3b associates with Factor B to form the C3bBb-type C3 convertase (alternative complement pathway), and can associate with C4b and C2a to form the C4b2a3b-type C5 convertase (classical pathway), or with C3bBb to form the C3bBb3b-type C5 convertase (alternative pathway).

Processing of C3b to the form iC3b, which is proteolytically inactive and which cannot itself promote further complement amplification, involves proteolytic cleavage of the C3b α' chain at amino acid positions 1303 and 1320 to form an α' chain fragment 1 (corresponding to amino acid positions 672 to 748 of C3; SEQ ID NO:22), an α' chain fragment 2 (corresponding to amino acid positions 1321 to 1,663 of C3; SEQ ID NO:23). Thus, iC3b comprises the C3 β chain, C3 α' chain fragment 1 and C3 α' chain fragment 2 (associated via disulphide bonds). Cleavage of the α' chain also liberates C3f, which corresponds to amino acid positions 1304 to 1320 of C3 (SEQ ID NO:24).

As used herein "C3" refers to C3 from any species and include isoforms, fragments, variants or homologues of C3 from any species. In some embodiments, the C3 is mammalian C3 (e.g. cynomolgous, human and/or rodent (e.g. rat and/or murine) C3). Isoforms, fragments, variants or homologues of C3 may optionally be characterised as having at least 70%, preferably one of 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity to the amino acid sequence of immature or mature C3 from a given species, e.g. human C3 (SEQ ID NO:18).

As used herein "C3b" refers to and includes isoforms, fragments, variants or homologues of C3b from any species. In some embodiments, the C3b is mammalian C3b (e.g. cynomolgous, human and/or rodent (e.g. rat and/or murine) C3b).

Isoforms, fragments, variants or homologues of C3b may optionally be characterised as comprising a C3 α' chain fragment 1, C3 α' chain fragment 2 and a C3 β having at least 70%, preferably one of 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity to the amino acid sequences of the respective polypeptides from a given species, e.g. human. That is, the C3b may comprise: a C3 α' chain fragment 1 having at

least 70%, preferably one of 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity to SEQ ID NO:22; a C3 α' chain fragment 2 having at least 70%, preferably one of 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity to SEQ ID NO:23; and a C3 β chain having at least 70%, preferably one of 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity to SEQ ID NO:19.

10 Isoforms, fragments, variants or homologues of C3b may optionally be functional isoforms, fragments, variants or homologues, e.g. having a functional property/activity of the reference C3b, as determined by analysis by a suitable assay for the functional property/activity. For example, Isoforms, 15 fragments, variants or homologues of C3b may be characterised by the ability to act as an opsonin, and/or to form functional C3/C5 convertase.

Complement Factor I

Processing of C3b to iC3b is performed by Complement 20 Factor I (encoded in humans by the gene CF). Human Complement Factor I (UniProt: P05156; SEQ ID NO:25) has a 583 amino acid sequence (including an N-terminal, 18 amino acid signal peptide). The precursor polypeptide is cleaved by furin to yield the mature Complement Factor I, 25 comprising a heavy chain (amino acids 19 to 335), and light chain (amino acids 340 to 583) linked by interchain disulphide bonds. Amino acids 340 to 574 of the light chain encode the proteolytic domain of Complement Factor I (SEQ ID NO:26), which is a serine protease containing the catalytic triad responsible for cleaving C3b to produce iC3b (Ekdahl et al., J Immunol (1990) 144 (11): 4269-74).

As used herein "Complement Factor I (FI)" refers to Complement Factor I from any species and includes isoforms, fragments, variants or homologues of Complement 35 Factor I from any species. In some embodiments, the Complement Factor I is mammalian Complement Factor I (e.g. cynomolgous, human and/or rodent (e.g. rat and/or murine) Complement Factor I).

Isoforms, fragments, variants or homologues of Complement 40 Factor I may optionally be characterised as having at least 70%, preferably one of 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity to the amino acid sequence of immature or mature Complement Factor I from a given species, e.g. 45 human Complement Factor I (SEQ ID NO:25). Isoforms, fragments, variants or homologues of Complement Factor I may optionally be functional isoforms, fragments, variants or homologues, e.g. having a functional property/activity of the reference Complement Factor I (e.g. full-length human Complement Factor I), as determined by analysis by a suitable assay for the functional property/activity. For example, an isoform, fragment, variant or homologue of Complement Factor I may display serine protease activity and/or may be capable of inactivating C3b.

50 Proteolytic cleavage of C3b by Complement Factor I to yield iC3b is facilitated by co-factors for Complement Factor I. Co-factors for Complement Factor I typically bind to C3b and/or Complement Factor I, and potentiate processing of C3b to iC3b by Complement Factor I.

Complement Receptor 1

Complement Receptor 1 (CR1) acts as a cofactor for Complement Factor I, enabling cleavage of C3b to iC3b and downstream products.

iC3b does not amplify or activate the complement system, 55 but it can still act as an opsonin to target pathogens for phagocytosis. iC3b production therefore results in local immune system activation and inflammatory effects. This

can have negative consequences for sufferers of complement-related disease and may contribute to the development or worsening of an existing complement-related disease/condition.

Factor H (FH) and truncated FH isoform FHL-1 act as co-factors for FI to produce iC3b, but they cannot promote further degradation of iC3b which can lead to undesirable iC3b accumulation. In addition, the accumulation of iC3b may contribute to further debris in the affected area, leading to e.g. the (further) development of drusen in macular degeneration.

In contrast, CR1 and the polypeptides according to the present invention, see e.g. FIGS. 2, 3B, 7A, can act in combination with FI to promote further breakdown of iC3b into advantageous downstream products such as C3c, C3dg and C3b. These molecules are not opsonins and thus avoid recruiting immune system components. Their presence in an affected area is preferable to iC3b accumulation.

As used herein, "Complement Receptor 1 (CR1)" refers to CR1 from any species and includes isoforms, fragments, variants or homologues of CR1 from any species. In some embodiments, the CR1 is mammalian CR1 (e.g. cynomolgous, human and/or rodent (e.g. rat and/or murine) CR1).

Aspects and embodiments of the present invention will now be illustrated, by way of example, with reference to the accompanying figures. Further aspects and embodiments will be apparent to those skilled in the art. All documents mentioned in this text are incorporated herein by reference.

The invention includes the combination of the aspects and preferred features described except where such a combination is clearly impermissible or expressly avoided.

The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

The features disclosed in the foregoing description, or in the following claims, or in the accompanying drawings, expressed in their specific forms or in terms of a means for performing the disclosed function, or a method or process for obtaining the disclosed results, as appropriate, may, separately, or in any combination of such features, be utilised for realising the invention in diverse forms thereof.

While the invention has been described in conjunction with the exemplary embodiments described above, many equivalent modifications and variations will be apparent to those skilled in the art when given this disclosure. Accordingly, the exemplary embodiments of the invention set forth above are considered to be illustrative and not limiting. Various changes to the described embodiments may be made without departing from the spirit and scope of the invention.

For the avoidance of any doubt, any theoretical explanations provided herein are provided for the purposes of improving the understanding of a reader. The inventors do not wish to be bound by any of these theoretical explanations.

Throughout this specification, including the claims which follow, unless the context requires otherwise, the word "comprise," and variations such as "comprises" and "comprising," will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

It must be noted that, as used in the specification and the appended claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise. Ranges may be expressed herein as from "about" one particular value, and/or to "about" another particular value. When such a range is expressed, another embodiment

includes from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, by the use of the antecedent "about," it will be understood that the particular value forms another embodiment.

Where a nucleic acid sequence is disclosed the reverse complement thereof is also expressly contemplated.

The following numbered paragraphs (paras) describe particular aspects and embodiments of the present invention:

1. A polypeptide having at least 80% sequence identity to SEQ ID NO:4, wherein the polypeptide has a length of 700 amino acids or fewer.
2. The polypeptide according to para 1, wherein the polypeptide has a length of 50 to 700 amino acids.
3. The polypeptide according to para 1 or 2, wherein X₁ is A or T, X₂ is P or L, and/or X₃ is G or R.
4. The polypeptide according to any one of paras 1 to 3, comprising SEQ ID NO:2, SEQ ID NO:3 or SEQ ID NO:13.
5. The polypeptide according to any one of paras 1 to 4, consisting of SEQ ID NO:2, SEQ ID NO:3, or SEQ ID NO:13.
6. The polypeptide according to any one of paras 1 to 5, which is capable of binding to C3b.
7. The polypeptide according to any one of paras 1 to 6, which binds to C3b in the region bound by a co-factor for Complement Factor I.
8. The polypeptide according to any one of paras 1 to 7, which binds to C3b in the region bound by Complement Receptor 1 (CR1).
9. The polypeptide according to any one of paras 1 to 8, which acts as a co-factor for Complement Factor I.
10. The polypeptide according to any one of paras 1 to 9, which is capable of diffusing across Bruch's membrane (BrM).
11. The polypeptide according to any one of paras 1 to 10, which is not glycosylated or is partially glycosylated.
12. The polypeptide according to any one of paras 1 to 11, wherein the amino acid sequence comprises one or more amino acid substitutions at position 509, 578, 959 and/or 1028 (numbered according to Uniprot: P17927).
13. The polypeptide according to para 12, wherein the one or more amino acid substitutions are selected from N509Q, N578Q, N959Q and/or N1028Q (numbered according to Uniprot: P17927).
14. The polypeptide according to any one of paras 1 to 13, comprising, or consisting, of SEQ ID NO:5, SEQ ID NO:6, and/or SEQ ID NO:15.
15. The polypeptide according to any one of paras 1 to 14, additionally comprising a secretory pathway sequence.
16. The polypeptide according to para 15, wherein the secretory pathway sequence comprises SEQ ID NO:7.
17. The polypeptide according to para 15 or para 16, wherein the polypeptide additionally comprises a cleavage site for removing the secretory pathway sequence.
18. A nucleic acid encoding the polypeptide according to any one of paras 1 to 17.
19. A vector comprising the nucleic acid of para 18.
20. A cell comprising the polypeptide according to any one of paras 1 to 17, the nucleic acid according to para 18, or the vector according to para 19.
21. A method for producing a polypeptide, comprising introducing into a cell a nucleic acid according to para 18 or a vector according to para 19, and culturing the cell under conditions suitable for expression of the polypeptide.
22. A cell which is obtained or obtainable by the method according to para 21.

23. A pharmaceutical composition comprising the polypeptide according to any one of paras 1 to 17, the nucleic acid according to para 18, the vector according to para 19, or the cell according to para 20 or 22, optionally comprising a pharmaceutically acceptable carrier, adjuvant, excipient, or diluent.

24. The polypeptide according to any one of paras 1 to 17, the nucleic acid according to para 18, the vector according to para 19, the cell according to para 20 or 22, or the pharmaceutical composition according to para 23, for use in a method of treating or preventing a disease or condition.

25. Use of the polypeptide according to any one of paras 1 to 17, the nucleic acid according to para 18, the vector according to para 19, the cell according to para 20 or 22, or the pharmaceutical composition according to para 23, in the manufacture of a medicament for treating or preventing a disease or condition.

26. A method of treating or preventing a disease or condition, comprising administering to a subject the polypeptide according to any one of paras 1 to 17, the nucleic acid according to para 18, the vector according to para 19, the cell according to para 20 or 22, or the pharmaceutical composition according to para 23.

27. A method of treating or preventing a disease or condition in a subject, comprising modifying at least one cell of the subject to express or comprise a polypeptide according to any one of paras 1 to 17, a nucleic acid according to para 18, or a vector according to para 19.

28. The polypeptide, nucleic acid, vector, cell, or pharmaceutical composition for use according to para 24, the use according to para 25, or the method according to para 26 or para 27, wherein the disease or condition is a disease or condition in which C3b or a C3b-containing complex, an activity/response associated with C3b or a C3b-containing complex, or a product of an activity/response associated with C3b or a C3b-containing complex is pathologically implicated.

29. The polypeptide, nucleic acid, vector, cell, or pharmaceutical composition for use, the use, or the method according to any one of paras 24 to 28, wherein the disease or condition is age-related macular degeneration (AMD).

30. The polypeptide, nucleic acid, vector, cell, or pharmaceutical composition for use, the use, or the method according to any one of paras 24 to 29, wherein the method for treating or preventing a disease or condition comprises modifying at least one retinal pigment epithelial (RPE) cell of the subject to express or comprise a polypeptide according to any one of paras 1 to 17, a nucleic acid according to para 18, or a vector according to para 19.

31. A kit of parts comprising a predetermined quantity of the polypeptide according to any one of paras 1 to 17, the nucleic acid according to para 18, the vector according to para 19, the cell according to para 20 or para 22 or the pharmaceutical composition according to para 23.

BRIEF DESCRIPTION OF THE FIGURES

Embodiments and experiments illustrating the principles of the invention will now be discussed with reference to the accompanying figures.

FIGS. 1A and 1B. CR1a and nCR1a protein expression from human HEK293 cells. (1A) CR1a is expressed as two glyco-forms that are reduced to one lower MW band after deglycosylation treatment with PNGase. (1B) nCR1a is expressed as single form of same MW as deglycosylated CR1a.

FIG. 2. The ability of CR1a and nCR1a expressed and secreted by human cells to act as cofactors for Factor I-mediated breakdown of C3b. CR1a/nCR1a+FI+C3b reactions produced iC3b and C3dg (lanes 7 and 11).

5 FIGS. 3A to 3C. (3A) Presence of CR1a in the diffusion chamber of an Ussing chamber after 24 hours, showing ability to diffuse through Bruch's membrane (BrM). (3B, 3C) The ability of (3B) CR1a and (3C) nCR1a to act as cofactors for Factor I to break down C3b after polypeptide diffusion through Bruch's membrane (BrM).

10 FIG. 4. Schematic of an Ussing chamber used in diffusion experiments, in which (a) is enriched Bruch's membrane from human donor eyes, covering a 5 mm aperture and representing the only passage of liquid from one chamber to another; (b) are sampling access points; and (c) are magnetic stirrer bars.

15 FIG. 5. Binding kinetics of CR1a, FH and FHL-1 for C3b measured by Biolayer interferometry (BLI).

20 FIG. 6. Detection of secreted CR1a (AAV-CR1a media) from the tissue culture media of AAV-CR1a transduced ARPE-19 cells using anti-CR1a antibody. Recombinant CR1a protein was used as a positive control. Media alone and media from culture of AAV-GFP transduced ARPE-19 cells were used as negative controls.

25 FIGS. 7A and 7B. Ability of secreted CR1a from human ARPE-19 cells to act as a cofactor for Factor I-mediated breakdown of C3b. (7A) Secreted CR1a acts as FI cofactor to produce iC3b (product e) and C3dg (product f). CR1a secreted from human HEK293 cells was provided as a control. (7B) Cultured AAV-CR1a transduced RPE cells showed increased ability to break down C3b (product a) to iC3b (product b) compared to media containing non-transduced cells.

30 FIG. 8. Schematic of the macular region of the eye, showing photoreceptors, retinal pigment epithelium, Bruch's membrane, and the choriocapillaris and intercapillary septa in the choroid.

35 FIG. 9. Representative fluorescence microscopy image of 40 retinal tissue from Forest et al. (2015) *Dis. Mod. Mech.* 8, 421-427 (FIG. 1). Shown are drusen deposits underlying degraded RPE cells and areas of complement activation. Scale bar 20 µm.

45 FIG. 10. Schematic of the macular region of the eye showing key stages of localised expression of effective complement therapeutic.

50 FIG. 11. Example of an expression vector comprising nucleic acid encoding a polypeptide according to the present invention (e.g. CR1a), promoter elements, replication elements and selection elements.

EXAMPLES

In the following Example(s), the inventors describe the 55 design of recombinant CR1 proteins comprising the C3b binding co-factor regions of Complement Receptor 1. Also described is the ability of these proteins to be expressed by human cells, diffuse through Bruch's membrane enriched from human donor eyes, and confer regulatory activity, i.e. facilitate the FI-mediated breakdown of C3b into iC3b and further breakdown products.

Example 1

65 DNA inserts encoding the amino acid sequences shown in SEQ ID NOs:2 and 5 were prepared by recombinant DNA techniques, and cloned into a vector to generate constructs

61

for recombinant expression of CR1 peptides. The amino acid sequences and features thereof are shown below:

CR1a
~~(SEQ ID NO: 47)~~
MRLLA^{KI}I^CLMLWA^ICVAGH^CQAPDHFLFAKLKTQTNASDFPIG^TSLKY
ECRPEYYGRPFSITCLDNLVWSSPKDVCKRKSC^KT^PDPVNGMVHVITD
I^QVGSRINY^SCTTG^HR^LI^GHSSAECILSGNAAH^WSTKPPICQRIPCGLP
PTIANGDFISTNREN^FHYG^SVVTYRCNP^GSGGRKV^FELVGEPSIYCTS^N
DDQVG^IW^SGP^APQCII

Signal peptide (SEQ ID NO:7); CCPs8-10 of CR1 (SEQ ID NO:2) (UniProt: P17927, residues 491-685)

nCR1a
~~(SEQ ID NO: 48)~~
MRLLA^{KI}I^CLMLWA^ICVAGH^CQAPDHFLFAKLKTQASDFPIG^TSLKY
ECRPEYYGRPFSITCLDNLVWSSPKDVCKRKSC^KT^PDPVNGMVHVITD
I^QVGSRIQYSCTTG^HR^LI^GHSSAECILSGNAAH^WSTKPPICQRIPCGLP
PTIANGDFISTNREN^FHYG^SVVTYRCNP^GSGGRKV^FELVGEPSIYCTS^N
DDQVG^IW^SGP^APQCII

Signal peptide (SEQ ID NO:7); CCPs8-10 of CR1 (SEQ ID NO:5) (UniProt: P17927, residues 491-685) comprising substitutions N509Q and N578Q

The 18-amino acid signal peptide is designed to be cleaved from the polypeptides upon secretion.

In some experiments, HIS-tagged CR1a and nCR1a were used, e.g. as shown in SEQ ID NO:40 and 21, respectively.
 Expression of Protein from Human Cells

HEK 293T cells (7×10^6 cell per plate) were grown in 15 cm culture plates overnight in 17 ml of Dulbecco's Modified Eagle's Medium with high glucose (DMEM, Sigma, catalogue number D469) supplemented with 10% Foetal Bovine Serum (FBS, Sigma, catalogue number F9665) in 5% CO₂ incubator at 37° C. Once the cells reached 60% confluence they were transiently transfected with 14.4 µg plasmid expressing either CR1a (SEQ ID NO:2) or nCR1a (SEQ ID NO:5) linked to the signal peptide (SEQ ID NO:7) with 86.4 µl of 7.5 mM Polyethylenimine (PEI, Polysciences, catalogue number 24765-2) 150 mM NaCl (Fisher Scientific UK Ltd, catalogue number 1073592). For the negative control 14.4 µl of Tris-EDTA buffer was used instead of the plasmid DNA. Five hours after transfection, the transfection medium containing 10% FBS was replaced with 17.5 ml fresh DMEM with high glucose supplemented with 2% FBS (referred to henceforth as expression media). Expression media was collected 24, 48, 72 and 140 hours after transfection. 80 µl of 0.5 M Phenylmethanesulfonyl fluoride (PMSF, Sigma, Catalogue number P7626) was added to every 100 ml of expression media collected and stored at 4° C. Expression media collected after 24 hrs was used for diffusion and function studies described below.

Characterisation of Secreted Protein from Human Cells

Protein expression from human cells is shown in FIGS. 1A and 1B. Recombinantly expressed and purified CR1a protein was found to be secreted from the cells as two glyco-forms. Treatment with PNGaseF, an enzyme that removes glycosylation, reduced the two bands to one band of lower apparent molecular weight (1A). A non-glycosylated form of the protein (nCR1a) was expressed. Western

62

blotting demonstrated that nCR1a produces a single band that migrates to the same position as the enzymatically de-glycosylated protein (1B).

C3b Breakdown Activity

CR1a and nCR1a expressed and secreted from human HEK293 cells were tested for their ability to act as cofactors for Factor I-mediated breakdown of C3b.

1 µg of recombinant CR1a or nCR1a protein was mixed with 2 µg of pure C3b protein and 0.04 µg of pure complement Factor I (FI; VWR International, catalogue no. 341280) for 15 minutes at 37° C. FHL-1 was provided as a co-factor control for FI. CR1a/nCR1a, CR1a/nCR1a+FI, CR1a/nCR1a+C3b and C3b alone were also provided as controls. Reactions were stopped with the addition of 4×SDS-loading buffer and heating to 100° C. for 5 minutes. Samples were then run on 4-12% NuPAGE Bis-Tris gels, run at 200V for 60 minutes. Samples were transferred onto nitrocellulose membranes at 80 mA for 1.5 hours using semi-dry transfer apparatus in transfer buffer (25 mM Tris, 192 mM glycine, 10% (v/v) Methanol). The membranes were blocked in PBS, 10% (w/v) milk, 0.2% (w/v) BSA for 16 hours at 4° C. before the addition of anti-C3b antibody (HycultBiotech, catalogue no. HM2287), at 100 µg/ml, in PBS, 0.2% (v/v) Tween-20 (PBS-T) for 1 hours at room temperature. Membranes were washed 2×30 min in PBS-T before the addition of a 1:2500 dilution of HRP conjugated goat anti-mouse for 1 hour at room temperature, protected from light. Membranes were washed 2×30 min in PBS-T before the addition of SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific, catalogue no. 34080) for 3 min at room temperature. Reactive bands were detected by exposing Super RX-N X-ray film (FujiFilm, catalogue no. PPB5080) to the treated membrane for 2 min at room temperature, and developed on an automated X-ray film developer.

The results are shown in FIG. 2. Both secreted proteins were found to be able to act as cofactors for Factor I, leading to the breakdown of C3b to firstly iC3b (a₁), and further to C3dg (a₁₋₁). The C3b breakdown using CR1a/nCR1a continues further than the normal native breakdown of C3b observed using FI+FHL-1 (second lane), which only produces iC3b (a₁).

Ussing Chamber Diffusion Experiments and C3b Breakdown Activity

The macular region of enriched Bruch's membrane isolated from donor eyes was mounted in an Ussing chamber (Harvard Apparatus, Hamden, CT) as described in McHarg et al., J Vis Exp (2015) 1-7, supra. Once mounted, the 5-mm-diameter macular area was the only barrier between two identical compartments, i.e. liquid must pass through Bruch's membrane (FIG. 4). Both sides of Bruch's membrane were washed with 2 ml PBS for 5 min at room temperature. The structural integrity of Bruch's membrane was tested prior to the experiments by its ability to retain 2 ml of liquid in one chamber without leaking into the next. 2 ml of expression media containing recombinant protein (CR1a and/or nCR1a, see above) was added to a chamber (henceforth designated the sample chamber) and 2 ml of fresh PBS was added to the other chamber (henceforth referred to as the diffusate/diffusion chamber). The Ussing chamber was left at room temperature for 24 hours with gentle stirring in each compartment with magnetic stirrer bars to avoid generating gradients of diffusing protein.

CR1a was added to the sample chamber. After 24 hours, samples from each chamber (original sample chamber and diffusion chamber) were analysed for the presence of CR1a.

63

The results are shown in FIG. 3A. CR1a was found to be present in the diffusion chamber after 24 hours.

Separately, CR1a and nCR1a were added to sample chambers and samples from both chambers were analysed for C3b-breakdown activity. After 24 hours, 18.6 µl samples were taken from each chamber and mixed with 1 µl (1 µg) of pure C3b protein and 0.4 µl of pure complement factor I (0.04 µg, VWR International, catalogue no. 341280) for one of 15, 30, or 60 minutes at 37° C. Reactions were stopped with the addition of 4xSDS-loading buffer and heating to 100° C. for 5 minutes. Samples were then run on 4-12% NuPAGE Bis-Tris gels, run at 200V for 60 minutes.

The results are shown in FIGS. 3B and 3C. Breakdown of C3b in the diffusion chamber demonstrates that both glycosylated CR1a (3B) and non-glycosylated nCR1a (3C) were able to cross Bruch's membrane from the sample chamber and remain functionally active. CR1a and nCR1a from both chambers were found to act successfully as co-factors for FI to break down C3b into proteolytically-inactive C3b (iC3b) and further products, as demonstrated by the presence of bands representing C3b breakdown products iC3b and C3dg. Thus, CR1a and nCR1a are able to diffuse through BrM and retain the ability to confer C3b breakdown in the presence of Factor I.

Example 2

The expression levels of polypeptides according to the present invention are compared to assess if there is an optimal formulation, i.e. whether glycosylated polypeptides express at higher levels than non-glycosylated polypeptides. It is anticipated that the glycosylation state of the polypeptides will have a minimal effect on level of expression.

CR1a was found to be expressed well by human cells. nCR1a was found to be expressed at lower levels than CR1a but both polypeptides were found to be functionally active, see e.g. FIG. 2.

Example 3

The binding kinetics of purified polypeptides according to the present invention for C3b is tested using Surface Plasmon Resonance (SPR), whereby C3b is immobilised onto SPR chips and the polypeptides used in the fluid phase. The association and dissociation constants are directly measured and a kD value for the interaction inferred.

Whilst no differences are observed in the binding of the CR1a (CCPs 8-10) or CR1 b (CCPs 15-17) species to C3b, glycosylated polypeptides bind more strongly to C3b than non-glycosylated polypeptides.

Binding Kinetics Measured by Bio-Layer Interferometry (BLI)

The affinity of CR1a for C3b protein was measured using OctetRed96 System (ForteBio, Pall Corp., USA). Biotinylated C3b protein was diluted in 0.2% PBST to final concentration of 0.4 µg/mL and loaded onto High Precision Streptavidin (SAX) Biosensors (ForteBio, Pall Corp., USA) for 600 s, previously hydrated with the same buffer for 20 min. C3b-loaded sensors were then washed with 0.2% PBST for 150 s (baseline) and dipped into wells containing CR1a in different concentrations ranging from 30.0 µg/mL to 2.6 µg/mL for 600 s (association) followed by wash with 0.2% PBST for 600 s (dissociation). Association and dissociation profiles were recorded and analysed with ForteBio Data Analysis v9 (ForteBio, Pall Corp., USA). The negative control i.e. 0.2% PBST containing well was used in parallel to subtract binding resulting from nonspecific interactions

64

with the sensors. Experiments were performed using the kinetics mode, at 25° C. and sample plates were pre-agitated for 3 min. The binding profiles were globally fitted to 1:1 (one analyte in solution to one binding site on the surface).

5 The KD was determined using data of the association (from 0 s to 600 s) and dissociation (from 0 s to 100 s) phases from four the lowest available analyte concentration by steady-state analysis. Binding affinities were also determined for native soluble C3b-binding complement regulators Factor H (FH) and FHL-1 to immobilised C3b.

10 The results are shown in FIG. 5. The binding affinity of CR1a for C3b was found to be 21 nM. For comparison, the binding affinity for C3b of FH was 580 nM and of FHL-1 was 1.2 mM. CR1a therefore binds significantly more strongly to C3b than either native soluble complement regulator. The strong binding affinity of CR1a for C3b means that CR1a is a more effective agent to promote C3b break down than agents based on FH or FHL-1.

15 A strong binding affinity also enables CR1a to promote degradation of iC3b into desirable further downstream products e.g. C3dg (see FIG. 2). In contrast, FH and FHL-1 cannot cause degradation of C3b beyond iC3b. iC3b is a pro-inflammatory molecule which acts to recruit immune cells to the site of complement activation, which in turn cause negative inflammatory effects. Further breakdown of iC3b to C3dg by CR1a is thus advantageous and avoids further damage caused by the immune system.

Example 4

The diffusion rates of polypeptides according to the present invention are compared using Ussing chamber experiments (described above) to see if any differences arise due to the formulation of the protein.

20 Experiments include using Bruch's membranes derived from donors with AMD to determine whether the material deposited in the Bruch's membrane in this condition, including drusen and basal linear deposits, compromises the ability of the polypeptide to cross Bruch's membrane. It is anticipated that the optimal polypeptide will cross Bruch's membrane efficiently even in the presence of AMD changes.

Example 5

25 Nucleic acid according to the present invention, optionally encoding the preferred signal peptide, a polypeptide having at least 80% sequence identity to SEQ ID NO:4 and a termination codon, is inserted into an AAV vector. The resultant expression vector is used to transfect cultured RPE cells. The expression and secretion of the encoded polypeptide is evaluated. It is anticipated that the polypeptide having at least 80% sequence identity to SEQ ID NO:4 is secreted by the RPE cells and that the signal peptide has been cleaved from the secreted polypeptide.

30 Adeno-associated virus (AAV) transduction Nucleic acid encoding the CR1a polypeptide described in Example 1 was transfected into human APRE-19 cells (ATCC, USA) derived from retinal pigment epithelium.

35 AAV2 serotype viral particles were pre-packed with CR1a plasmid. ARPE-19 cells were seeded on six well cell culture plate (Corning) in 2 ml of DMEM/F12 growth medium (ATCC, USA) supplemented with 10% (v/v) fetal bovine serum (ATCC, USA) at the density of 300,000 cells per well. Cells were then left for incubation at 37° C. in humidified atmosphere of 5% CO₂ for 24 hrs. After incubation, cells were washed twice with 2 ml serum free DMEM/F12 growth medium. AAV2-CR1a at multiplicity of infection

(MOI) 100,000 in serum free DMEM/F12 growth medium in total volume of 1 ml was added. AAV2-CR1a containing medium was incubated with cells for 24 hrs (37° C., 5% CO₂) followed by replacement with 2 ml of fresh serum free DMEM/F12 growth medium the next day. Control cells transduced with AAV-GFP were grown in parallel. Transduction efficiency was assessed 14 days post-infection. Secretion of CR1a by ARPE-19 cells was detected by dot blot: conditioned media from CR1a-transduced RPE cells was contacted with in-house polyclonal anti-CR1a antibody.

The results are shown in FIG. 6. Immunoreactivity with the anti-CR1a antibody was seen in both samples tested (AAV-CR1a media), compared to media from RPE cells transduced with AAV-GFP (Ctrl media) used as a negative control. Purified recombinant CR1a protein was included as a positive control.

Recombinant CR1a polypeptides secreted from human APRE-19 and HEK293 cells were assessed for their ability to break down C3b to iC3b and C3dg.

The results are shown in FIG. 7A. Functional CR1a polypeptide secreted from human APRE-19 cells was found to act as a cofactor for Factor I leading to breakdown of C3b into iC3b (product e) and C3dg (product f). Reaction containing FHL-1+FI+C3b provided a MW control for C3b and product iC3b.

C3b break down was assessed in tissue culture media of human RPE cells (ARPE-19) transduced with AAV-delivered CR1a. AAV-GFP transduced RPE cells were used as a negative control. Transduced cells were supplemented with purified C3b and Factor I 14 days after transduction. C3b and iC3b levels in tissue culture media were detected by Western blot.

The results are shown in FIG. 7B. RPE cells produce their own complement components and were found to have a low native turnover of C3b (product a) into iC3b (product b; lane 2). However, RPE cells expressing AAV-delivered CR1a were found to have increased C3b breakdown capacity to iC3b (product b; lane 3) compared to the native turnover rate.

This indicates that CR1a is secreted successfully from eye cells, e.g. retinal pigment epithelium cells, is functionally active as a FI cofactor, enhances C3b break down to downstream products, is capable of complement regulation, and would provide therapeutic benefit for conditions involving over-activation of complement e.g. an excess of C3b.

Example 6

The polypeptides of the present invention find use in methods of treatment or prevention of complement-related disorders. One example of a complement related disorder is macular degeneration in the eye, e.g. AMD.

FIG. 8 provides a schematic of the eye macular region. The retinal pigment epithelium (RPE) is a continuous monolayer of cuboidal/columnar epithelial cells between the neurosensory retina and the vascular choroid. The cells have physical, optical, metabolic/biochemical and transport functions and play a critical role in the normal visual process. The RPE is separated from the choroid by Bruch's membrane (BrM): a thin (2-4 µm), acellular, five-layered, extracellular matrix. The BrM serves two major functions: the substratum of the RPE and a blood vessel wall. Immediately adjacent to the BrM, and within the choroid, is a layer of capillaries termed the choriocapillaris. Complement activation centres on the extracellular matrix of the choriocapillaris, termed the intercapillary septa.

The hallmark lesions of AMD, drusen, are formed from the accumulation of lipids and cellular debris, including many complement activation products. Drusen develop within the BrM adjacent to the RPE layer and disrupt the flow of nutrients from the choroid to the RPE, leading to cell dysfunction and death. The death of RPE cells also causes dysfunction of photoreceptor cells and subsequent loss of visual acuity.

A representative fluorescence microscopy image of drusen deposits is provided in FIG. 9, taken from Forest et al. (2015) *Dis. Mod. Mech.* 8, 421-427 (FIG. 1). FIG. 9 shows retinal tissue from an 82-year-old female with AMD. A cell membrane marker shows degraded RPE cells overlying drusen. Areas of complement activation within the drusen and around the blood vessels are indicated by the terminal complement complex marker C5b-9. Nuclei are stained. Scale bar 20 µm.

Systemic administration of a complement-regulating molecule would require a high dose and carries a substantial risk of detrimental off-target effects on functioning complement systems. Localised administration, e.g. expression from RPE cells, is a safer and more effective delivery method. However, for an RPE-expressed molecule to be effective in treating/preventing a complement-related disorder in the intercapillary septa of the choriocapillaris, i.e. the site of complement over-activation, the molecule will have to cross Bruch's membrane (see FIG. 10). That is, an effective complement-inhibiting therapy must meet the requirements of all three stages in FIG. 10.

As demonstrated herein, CR1a is:

1. Secreted from human RPE cells (see e.g. FIG. 6);
2. Able to passively diffuse across human BrM (see e.g. FIG. 3); and
3. Able to mediate the breakdown of C3b into iC3b, and further desirable break down products, in the presence of complement factor I (see e.g. FIGS. 2, 3, 7).

Therefore, CR1a is able to be expressed by RPE cells, is able to reach the areas of complement activation where C3b regulation is necessary, and can act as an effective therapeutic agent to treat complement over-activation in the eye, e.g. in AMD.

Example 7

45 Laser-Induced Choroidal Neovascularisation Model

This model applies laser burns to the retinas of mice or rats to induce choroidal neovascularization, for example as described in Schnabolk et al, *Mol Ther Methods Clin Dev.* 2018; 9: 1-11. The size of the choroidal neovascular complexes can be measured by fluorescein angiography or by histology.

Rodents are given a subretinal injection of AAV vector, e.g. AAV2, containing CR1a cDNA, or empty AAV vector as a control. Rodents receiving AAV vector containing CR1a cDNA secrete CR1a protein from their retinal pigment epithelial cells. When maximal CR1a protein excretion is achieved, laser burns are applied to the rodent retinas and the size of the choroidal neovascular complexes is measured at a pre-specified time after the laser burns. The complement inhibitory effect of CR1a polypeptide is found to decrease the size of the choroidal neovascular complexes compared to rodents administered the empty AAV vector.

Sodium Iodide Induced Retinal Degeneration

Mice are injected intravenously with sodium iodate which induces a degeneration of the retinal pigment epithelium that is partially dependent upon complement activation, for example as described in Katschke et al., *Sci Rep.* 2018;

8(1):7348. CR1a is delivered by intravitreal injection as a recombinant protein or by subretinal delivery using AAV vector, e.g. AAV2, containing CR1a cDNA. Empty AAV

vector is administered as a control. CR1a treated mice are found to have less retinal degeneration than the control mice that did not receive CR1a treatment.

SEQUENCE LISTING

```

Sequence total quantity: 54
SEQ ID NO: 1      moltype = AA  length = 2039
FEATURE          Location/Qualifiers
source           1..2039
                 mol_type = protein
                 organism = Homo sapiens
SEQUENCE: 1
MGASSPRSP E PVGPPAPGLP FCCGGSSLAV VVLLALPVAV GQCNAPEWLP FARPTNLTDE 60
PEFPPIGYTLN YECRPGYSGR PFSIICLKNS VWTGAKDRCR RKSCRNPPDP VNGMVHV1KG 120
IQFGSQIKYS CTKGYRLIGS SSATCIISGD TVIWDNETPI CDRIPCGLPP TITNGDFIST 180
NRENFHGSV VTYRCNPMSG GRKVFELVGE PSIYCTSND QVGIIWSGPAP QCIIIPNKCTP 240
PNVENGILVS DNRSLFSLNE DNRSLFSPGQ VMKGPRRVKC QALNKWEPEL PSCSRVCQPP 300
PDVLHAERTQ RDKDNFSPGQ EVFYSCCEPGY DLRGAAISMRC TPQGDWSPAA PTCEVKSCDD 360
FMGQLLNNGRV LFPVNQLQGA KVDVFCDEGF QLKGSASSASYC VLAGMESLWN SSVPVCEQIF 420
CPSPPVIPNG RHTGKPLEV PFGKTVNYTC DPHPDRGTSF DLIGESTIRC TSDPQGNGVW 480
SSPAKRCMRC GHQCAPDHFL FAKLKTQTNM DFSPIGTSLK YECRPEYYGR PFSITCLDN 540
VWSSPKDVCK RKSCKTPPDV NRENFHGSV VTYRCNPMSG GRKVFELVGE 600
AAHWSTKPP1 CQRIPCGLPP TIANGDFIST NRENFHGSV VTYRCNPMSG GRKVFELVGE 660
PSIYCTSND QVGIIWSGPAP QCIIIPNKCTP PNVENGLVS DNRSLFSLNE VVEFRCQPGF 720
VMKGPRRVCK QALNKWEPEL PSCSRVCQPP PDVLHAERTQ RDKDNFSPGQ EVFYSCCEPGY 780
DLRGAAISMRC TPQGDWSPAA PTCEVKSCDD FMGQLLNNGRV LFPVNQLQGA KVDVFCDEGF 840
QLKGSSASYC VLAGMESLWN SSVPVCEQIF CPSPPVIPNG RHTGKPLEV PFGKAVNYTC 900
DHPHDRGTSF DLIGESTIRC TSDPQGNGVW SSPAPRGCIL GHQCAPDHFL FAKLKTQTNM 960
SDFPIGTSLK YECRPEYYGR PFSITCLDN VVSSPKDVCK RKSCKTPPDV VNGMVHV1TD 1020
IQVGSRINYS CTTGHLIGH SSAECILSGN TAHWSTKPP1 CQRIPCGLPP TIANGDFIST 1080
NRENFHGSV VTYRCNLGSR GRKVFELVGE PSIYCTSND QVGIIWSGPAP QCIIIPNKCTP 1140
PNVENGILVS DNRSLFSLNE VVEFRCQPGF VMKGPRRVKC QALNKWEPEL PSCSRVCQPP 1200
PEILHGHEART SHQDNFSPGQ EVFYSCCEPGY DLRGAAASLHC TPQGDWSPAA PRCAVKSCDD 1260
FLGQLPHGRV LFPLNLQGA KVFVCDDEGF RLKGSSVSHC VLVGMRSLWN NSVPVCEHIF 1320
CPNPPA1LNQ RHTGTPSGDI PYGKEISYTC DPHPDRGTMF NLIGESTIRC TSDPQGNGVW 1380
SSPAPRCELS VRAGHCKTPE QFFPASPTIP INDFEFPVGT SLNYECRPGY FGKMFISCL 1440
ENLVWSSVED NCRRKSCGPP PEPFNGMVHI NTDTQFGSTV NYSCMEGFRRL IGPSTTCLV 1500
SGNNVTWDKK APICEIIISCE PPPTISNGDF YSNNRTSFHN GTVVTYQCHT GDPGEQLFEL 1560
VGERSIYCTS KDDQVGVWSS PPRCISTNK TAPEVENAI RVPGNRSFFS LTEIIRFRQ 1620
PGFVMVGSHT VQCQTNGRWG PKLPHCSRC QPPPEILHGE HTLSHQDNFS PGQEVSFSCE 1680
PSYDLRGAES LHCTPQGDS PEAPRCTVKS CDDFLCQLPH GRVLLPLNLQ LGAKVSVFCD 1740
EGFRLKGRSA SHCVELAGMKW LWNSSPVPC QIFCPNPPAI LNGRHTGTPF GDIPYKGIES 1800
YACDTHPDRC MTFNLIQESS IRCTSDPQGN GWSSPAPRCE ELSVPAACPH PPKIQNGHYI 1860
GGHVSILYLPG MTISYICDPG YLLVGKGFIF CTDQGIWSQL DHYCKEVNCS FPLFMNGISK 1920
ELEMKKVYHY GDYVTLKCED GYTFLEGSPWS QCQADDRWDP PLAKCTSRTH DALIVGTLSG 1980
TIFFILLIIF LSWIILKHRK GNNAHENPKE VAIHLHSQGG SSVHPRTLQ NEENSRLVP 2039

SEQ ID NO: 2      moltype = AA  length = 194
FEATURE          Location/Qualifiers
source           1..194
                 mol_type = protein
                 organism = Homo sapiens
SEQUENCE: 2
GHQCAPDHFL FAKLKTQTNM SDFPIGTSLK YECRPEYYGR PFSITCLDN VVSSPKDVCK 60
RKSCKTPPDV VNGMVHV1TD IQVGSRINYS CTTGHLIGH SSAECILSGN AHWSTKPP1 120
CQRIPCGLPP TIANGDFIST NRENFHGSV VTYRCNPMSG GRKVFELVGE PSIYCTSND 180
QVGIIWSGPAP QCII 194

SEQ ID NO: 3      moltype = AA  length = 194
FEATURE          Location/Qualifiers
source           1..194
                 mol_type = protein
                 organism = Homo sapiens
SEQUENCE: 3
GHQCAPDHFL FAKLKTQTNM SDFPIGTSLK YECRPEYYGR PFSITCLDN VVSSPKDVCK 60
RKSCKTPPDV VNGMVHV1TD IQVGSRINYS CTTGHLIGH SSAECILSGN TAHWSTKPP1 120
CQRIPCGLPP TIANGDFIST NRENFHGSV VTYRCNLGSR GRKVFELVGE PSIYCTSND 180
QVGIIWSGPAP QCII 194

SEQ ID NO: 4      moltype = AA  length = 194
FEATURE          Location/Qualifiers
SITE             111
                 note = MISC_FEATURE - X = A or T
SITE             157
                 note = MISC_FEATURE - X = P or L
SITE             160
                 note = MISC_FEATURE - X = G or R
source          1..194

```

-continued

```

mol_type = protein
organism = Homo sapiens

SEQUENCE: 4
GHCQAPDHFL FAKLKTQTNA SDFFIGTSLK YECRPEYYGR PFSITCLDNL VWSSPKDVCK 60
RKSCCKTPDP VNGMVHVITD IQVGSRINYA CTTGHRLIGH SSAECILSGN XAHWSTKPP 120
CQRIPCGLPP TIANGDFIST NRENPHYGSV VTYRCNXGSX GRKVFELVGE PSIYCTSND 180
QVGIWSGPAP QCII 194

SEQ ID NO: 5      moltype = AA length = 194
FEATURE          Location/Qualifiers
source           1..194
mol_type = protein
organism = Homo sapiens

SEQUENCE: 5
GHCQAPDHFL FAKLKTQTQA SDFFIGTSLK YECRPEYYGR PFSITCLDNL VWSSPKDVCK 60
RKSCCKTPDP VNGMVHVITD IQVGSRIQYS CTTGHRLIGH SSAECILSGN AAHWSTKPP 120
CQRIPCGLPP TIANGDFIST NRENPHYGSV VTYRCNPMSG GRKVFELVGE PSIYCTSND 180
QVGIWSGPAP QCII 194

SEQ ID NO: 6      moltype = AA length = 194
FEATURE          Location/Qualifiers
source           1..194
mol_type = protein
organism = Homo sapiens

SEQUENCE: 6
GHCQAPDHFL FAKLKTQTQA SDFFIGTSLK YECRPEYYGR PFSITCLDNL VWSSPKDVCK 60
RKSCCKTPDP VNGMVHVITD IQVGSRIQYS CTTGHRLIGH SSAECILSGN TAHWSTKPP 120
CQRIPCGLPP TIANGDFIST NRENPHYGSV VTYRCNLGSR GRKVFELVGE PSIYCTSND 180
QVGIWSGPAP QCII 194

SEQ ID NO: 7      moltype = AA length = 18
FEATURE          Location/Qualifiers
REGION          1..18
source           note = Synthetic construct
                1..18
mol_type = protein
organism = synthetic construct

SEQUENCE: 7
MRLLAIIICL MLWAICVA 18

SEQ ID NO: 8      moltype = AA length = 61
FEATURE          Location/Qualifiers
source           1..61
mol_type = protein
organism = Homo sapiens

SEQUENCE: 8
GHCQAPDHFL FAKLKTQTNA SDFFIGTSLK YECRPEYYGR PFSITCLDNL VWSSPKDVCK 60
R 61

SEQ ID NO: 9      moltype = AA length = 61
FEATURE          Location/Qualifiers
source           1..61
mol_type = protein
organism = Homo sapiens

SEQUENCE: 9
KSCKTPPPDV NGMVHVITDI QVGSRINYS TTGHRLIGH SSAECILSGNA AHWSTKPPIC 60
Q 61

SEQ ID NO: 10     moltype = AA length = 72
FEATURE          Location/Qualifiers
source           1..72
mol_type = protein
organism = Homo sapiens

SEQUENCE: 10
RIPCGLPPTI ANGDFISTNR ENFHYGSVVT YRCNPMSGGR KVFEVLGEPS IYCTSNDQV 60
GIWSGPAPQC II 72

SEQ ID NO: 11     moltype = AA length = 61
FEATURE          Location/Qualifiers
source           1..61
mol_type = protein
organism = Homo sapiens

SEQUENCE: 11
KSCKTPPPDV NGMVHVITDI QVGSRINYS TTGHRLIGH SSAECILSGNT AHWSTKPPIC 60
Q 61

SEQ ID NO: 12     moltype = AA length = 72
FEATURE          Location/Qualifiers
source           1..72

```

-continued

```

mol_type = protein
organism = Homo sapiens

SEQUENCE: 12
RIPCGLPPTI ANGDFISTNR ENFHYGSVVT YRCNLGSRGR KVFEVLGEPS IYCTSNDQV 60
GIWSGPAPQC II 72

SEQ ID NO: 13      moltype = AA length = 388
FEATURE          Location/Qualifiers
source           1..388
mol_type = protein
organism = Homo sapiens

SEQUENCE: 13
GHQCQAPDHFL FAKLKTQTNA SDFPIGTSLK YECRPEYYGR PFSITCLDNL VWSSPKDVCK 60
RKSCCKTPDPN VNGMVHVITD IQVGSRINYS CTTGHLIGH SSAECILSGN AAHWSTKPP 120
CQRIPCGLPP TIANGDFIST NRENPHYGSV VTYRCNPNGSG GRKVFEVLGE PSIYCTSND 180
QVGIWSPGP QCIIGHCQAP DHFLFAKLKT QTNASDFPIG TSLKYECRPE YYGRPFSTIC 240
LDNLVWSSPK DVCKRKSCKT PPDPVNGMVH VITDIQVGSR INYSCTTGHRLIGHSSAECI 300
LSGNTAHWST KPPICQRIPC GLPPTIANGD FISTNRENFH YGSVVTYRCN LGSRGRKVFE 360
LVGEPSIYCT SNDDQVGIWS GPAPQCII 388

SEQ ID NO: 14      moltype = AA length = 423
FEATURE          Location/Qualifiers
SITE             195
note = misc_feature - Xaa can be any naturally occurring
amino acid
REGION           196..229
note = MISC_FEATURE - Xaa can be any naturally occurring
amino acid or absent
source            1..423
mol_type = protein
organism = Homo sapiens

SEQUENCE: 14
GHQCQAPDHFL FAKLKTQTNA SDFPIGTSLK YECRPEYYGR PFSITCLDNL VWSSPKDVCK 60
RKSCCKTPDPN VNGMVHVITD IQVGSRINYS CTTGHLIGH SSAECILSGN AAHWSTKPP 120
CQRIPCGLPP TIANGDFIST NRENPHYGSV VTYRCNPNGSG GRKVFEVLGE PSIYCTSND 180
QVGIWSPGP QCIIXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX HCQAPDHFL 240
AKLKTQTNAS DFPPIGTSKY ECRPEYYGRP FSITCLDNLW WSSPKDVCKR KSCCKTPDPV 300
NGMVHVITDI QVGSRINYSC TTGHLIGHSAECLSGNT AHWSTKPPIC QRIPCGLPP 360
IANGDFISTN RENPHYGSVV TYRCNLGSRG RKVFELVGEPS SIYCTSNDQ VGIWSPGPQ 420
CII 423

SEQ ID NO: 15      moltype = AA length = 388
FEATURE          Location/Qualifiers
source           1..388
mol_type = protein
organism = Homo sapiens

SEQUENCE: 15
GHQCQAPDHFL FAKLKTQTQA SDFPIGTSLK YECRPEYYGR PFSITCLDNL VWSSPKDVCK 60
RKSCCKTPDPN VNGMVHVITD IQVGSRIQYS CTTGHLIGH SSAECILSGN AAHWSTKPP 120
CQRIPCGLPP TIANGDFIST NRENPHYGSV VTYRCNPNGSG GRKVFEVLGE PSIYCTSND 180
QVGIWSPGP QCIIGHCQAP DHFLFAKLKT QTQASDFPIG TSLKYECRPE YYGRPFSTIC 240
LDNLVWSSPK DVCKRKSCKT PPDPVNGMVH VITDIQVGSR IQYSCTTGHRLIGHSSAECI 300
LSGNTAHWST KPPICQRIPC GLPPTIANGD FISTNRENFH YGSVVTYRCN LGSRGRKVFE 360
LVGEPSIYCT SNDDQVGIWS GPAPQCII 388

SEQ ID NO: 16      moltype = AA length = 61
FEATURE          Location/Qualifiers
SITE             50
note = MISC_FEATURE - Xaa = Ala or Thr
source            1..61
mol_type = protein
organism = Homo sapiens

SEQUENCE: 16
KSCKTPPPDV NGMVHVITDI QVGSRINYSC TTGHLIGHSAECLSGNX AHWSTKPPIC 60
Q 61

SEQ ID NO: 17      moltype = AA length = 72
FEATURE          Location/Qualifiers
SITE             35
note = MISC_FEATURE - Xaa = Pro or Leu
SITE             38
note = MISC_FEATURE - Xaa = Gly or Arg
source            1..72
mol_type = protein
organism = Homo sapiens

SEQUENCE: 17
RIPCGLPPTI ANGDFISTNR ENFHYGSVVT YRCNXGSXGR KVFEVLGEPS IYCTSNDQV 60
GIWSGPAPQC II 72

```

-continued

SEQ ID NO: 18	moltype = AA length = 1663
FEATURE	Location/Qualifiers
source	1..1663
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 18	
MGPTSGPSLL LLLLTHLPLA LGSPMYSIIT PNILRLESEE TMVLEAHDAQ GDVPVTVTVH 60	
DFPGKKLVLS SEKTVLTPAT NHMGNTFTI PANREFKSEK GRNKFVTVQA TFGTQVVEKV 120	
VILSLSQSGYL FIQTDTKTIYI PGSTVLYRIF TVNHKLLPVG RTVMVNIENP EGIPVKQDSL 180	
SSQNQLGVLPN LSWDPELNVN MGQWKIRAYY ENSPQQVFP EFEVKEYVLP SFEVIVEPTE 240	
KFYIYIYNEKG LEVTITARFL YGKKVEGTAF VIFGIQDGEQ RISLSPESLKR IPIEDGSGEV 300	
VLSRKVLLDG VQNPRADELV GKSLYVSATV ILHSGSDMVQ AERSGIPIVT SPYQIHFTKT 360	
PKYFKPGMPF DLMVFVTPNP GSPAYRVPVA VQGEDTVQSL TQGDGVAKLS INTHPSQKPL 420	
SITVRTTKQE LSEAFAQTRT MQALPYSTKNE NSNNYHLHSLV LRTELRLPGET LNWNFLLRMD 480	
RAHEAKIRYY TYLIMMKGRU LKAGRQVRREP QGDLVVPLPS ITTDFIPSFR LVAYYTLIGA 540	
SGQREVVADS VVVDVKDCSV GSLVVKSGQS EDRQPVPQGQ MTLKIEGDHG ARVVLVAVDK 600	
GVFVLNKKNQ LTQSKIWDVV EKADGCTPG SGKDYGAVFS DAGLTFTSSS QQQTQRAEL 660	
QCQPQPAARRR RSVQLEKRM DKVGKYPKEL RKCCEDGMRE NPMRFSQRR TRFISLGEAC 720	
KKVFLDCCNY JTELRLRQHAR ASHGLLARSN LDDEDIIABEN IVSRSEFPEES WLWNVEDLKE 780	
PPKNGISTKL MNIFLKSIT TWEILAVSMS DKKGICVADP FEVTVMQDFP IDLRLPYSVV 840	
RNBQVEIRAV LYNYRQNQEL KVRVLLHNP AFCSLATTK RHQQTVTIIP KSSLSPVYVI 900	
VPLKTGLQEV EKAAVYHHF ISDGVRKSLK VVPEGIRMNK TAVAVRDLPE RLREGVQKE 960	
DIPADLSDQ VPDTESETR LLQGTPVAMQ TEDAVDAERL KHLIVTPSGC GEQNMIGMTP 1020	
TVIAVHYLDE TEQWEKFGLR KRQGALELIK KGTYQQLAER QPSSAFAAV KRAPSTWLTA 1080	
YVVKVFSLAV NLIAIDSQVL CGAVKWLILE KQKPDGVFQE DAPVIHQEMI GGLRNNEKND 1140	
MALTAFLVIS LQEAKDICEE VNQLSPGSIT KAGDFLEANY MNLQSYTVA IAGYALAQMG 1200	
RLKGPLLNUKF LTTAKDKNRW EDPGKQLYNV EATSYALLAL LQLQKDFVFP PVWRWLNEQR 1260	
YYGGGGSTQ ATFMVFQALA QYQKDAPDHQ ELNLDVSQL PSRSSKITHR IHWESASLLR 1320	
SEETKENEGRF TVTAEGKGQG TLSVVTMYHA KADQLTCKN FDLKVTKIPA PETEKRPQDA 1380	
KNTMILECT YRGDGQDATM SILDISMIMTQ FAPDTDDLKQ LANGVDRYIS KYELDKAFSD 1440	
RNTLIIYLDK VSHSEDDCLA FKVHQYFNVE LIQPGAVKVY AYYNLEESTC RFYHPEKEDG 1500	
KLNKLCRDEL CRCAEENCFI QKSDDKVTLE ERLDKACEPG VDYVYKTRLV KVQLSNDFDE 1560	
YIMAIEQTIK SGSDEVQVGQ QRTFISPIKC REALKLEEKH HYLMWGLSSD FWGEKPNLSY 1620	
IIGKDTWVEH WPEEDECQDE ENQKQCDLG AFTESTMVFG CPN 1663	
SEQ ID NO: 19	moltype = AA length = 645
FEATURE	Location/Qualifiers
source	1..645
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 19	
SPMYSIITPN ILRLESEETM VLEAHDAQD VPVTVTVHDF PGKKLVLSS KTVLTPATNH 60	
MGNVTFTIPA NREFKSEKGR NKFVTVQATF GTQVVEKVL VSLQSGYLFQ QTDTKIYTPG 120	
STVLYRIFTV NHKLLPVGRT VMVNIENPEG IPVKQDSLSS QNOLGVPLS WDIPELVNMG 180	
QWKIRAYEN SPQQVFSTEF EVKEYVLPSPF EVIVEPTEKF YYIYNEKGLE VTITARFLY 240	
KKVEGTAFVI FGIQDGEQRI SLPESLKRIP IEDGSGEVVL SRKVLLDGQV NPRAEFLVKG 300	
SLYVSATVIL HSGSDMVQAE RSGIPIVTSP YQIHFTKTPK YFKPGMPFDL MVFVTNPDS 360	
PAYRVPVAVQ QEDTVQSLTQ GDGVAKLSSIN THPSQKPLSI TVRTKQELM DEEQATRMTQ 420	
ALPYSTVGNS NNYLHLSVLR TELRPGETLN VNFLLRMDRA HEAKIRYTYT LIMNKGRLLK 480	
AGRQVREPGQ DLVVLPLSIT TDFIPSFLRV AYTLIGASQ QREVVADSVW VDVKDSCVGS 540	
LVVKVSGQSED RQPVPGQMT LKIEGDHGAR VVLAQVADKVG FVLNKKNKLQ QSKIWDVVEK 600	
ADIGCTPGSG KDYAGVFSDA GLTFTSSSGQ QTAQRAELQC QOPAA 645	
SEQ ID NO: 20	moltype = AA length = 915
FEATURE	Location/Qualifiers
source	1..915
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 20	
SNLDDEDIIAE ENIVSRSEFP ESWLWNVEDL KEPPKNGIST KLMNIFLKSIT ITTWEILAVS 60	
MSDKKGICVA DPFEVTVMQD FFIDLRLPYS VVRNEQVEIR AVLYNYRQNQ ELKVRVELLH 120	
NPAFCSLATT KRRHQQTVTI PPKSSLSPY VIVPLKTCGLQ EVEVKAAVYH HFISDGVRKS 180	
LKVPPEGIRN MKTVAVARTL PERLGREGVQ KEDIPPADLS DQVDPDTESET RILLQGTPVA 240	
QMTEDAVDAE RLKHLLVTPS GCGEQONMIGM TPTVIAVHDE DEEQATRMTQ LEKQGALEL 300	
IKKGYTQQLA FRQPSSAFAA FVKRAPSTWL TAYVVKVFSL AVNLIAIDSQ VLCGAVKWL 360	
LEKQKPDGVF QEDAPVIHQE MIGGLRNNNE KDMALTAFLV ISLQEAKDIC EEQVNSLPGS 420	
ITKAGDFLEA NYMLNQRSYI VAIAGYALAQ MGRLKGPQLN KFLTTAKDKK RWEDPGKQLY 480	
NVEATSYALL ALQLQDFDF VPPVYRNLNE QRYYGGGGYGS TQATFMVFOA LAQYQKDAPD 540	
HQBLNLDVSL QLPSRSKIIH RHIHWESASL LRSEETKENE GFTVTAEGKG QGTLSSVVTM 600	
HAKAKDQLTC NKFDLKVTKI PAPETEKRQ DAKNTMILEI CTRYRGDQDA TMSILDISM 660	
TGFAPDTDL KQLANGVDRY ISKYELDKAF SDRNTLIIYL DVSHSED DC LAFKVHQYFN 720	
VELIQPGAVK VYAYYNLEES CTRFYHPEKE DGKLNLKCRD ELCRCAEENC FIQKSDDKVT 780	
LEERLDKACE PGVDYVYKTR LVKVLQLSND DEYIMAIEQT IKSGSDEVQV GQQRTFISPI 840	
KCREALKLEE KKHLYMWGLS SDFWGKEPNL SYIIGKDTWV EHWPEDECQ DEENQKQCDN 900	
LGAFTESMVV FGCPN 915	
SEQ ID NO: 21	moltype = AA length = 77
FEATURE	Location/Qualifiers
source	1..77

-continued

```

mol_type = protein
organism = Homo sapiens

SEQUENCE: 21
SVQLTEKMRD KVGKYPKELR KCCEGMREN PMRFSCQRRT RFISLGEACK KVFLDCCNYI 60
TELRRQHARA SHLGLAR 77

SEQ ID NO: 22      moltype = AA length = 555
FEATURE           Location/Qualifiers
source            1..555
mol_type = protein
organism = Homo sapiens

SEQUENCE: 22
SNLDDEDIAE ENIVSRSEFP ESWLNWEDL KEPPKNGIST KLMNIFLKD ITTWEILAVS 60
MSDKKGICVA DPFEVTVMQD FFDLRLPYS VVRNEQVEIR AVLYNYRQNQ ELKVRVELLH 120
NPAPFCSLATT KRRHQQTVTI PPKSSLSPVY VIVPLKTGLQ EVEVKAAVYH HFISDGVRKS 180
LKVPPEGIRM NKTVAVRTLD PERLGREGVQ KEDIPPADLS DQVPDTESET RILLQGTPVA 240
QMTEDADVAE RLKHLLIVTPE GCGEONMIGM TPTVIAVHLY DETEQWEKFG LEKROGAEL 300
IKKGYTQQLA FRQPSSAFAA FVKRAPSTWL TAYVVKFVSL AVNLIAIDSQ VLCGAVKWLI 360
LEKQKPDGVF QEDAPVIHQE MIGGLRNNNE KDMALTAFVL ISLQEAKDIC EEQVNSLPGS 420
ITKAGDFLEA NYMNQLQRSYT VAIAGYALAQ MGRLKGPLL KFLTTAKDKN RWEDPGKQLY 480
NVEATSYALL ALLQLKDFDF VPPVVRWLNE QRYYGGGYGS TQATFMFQQA LAQYQKDAPD 540
HQELNLDVSL QLPSR 555

SEQ ID NO: 23      moltype = AA length = 343
FEATURE           Location/Qualifiers
source            1..343
mol_type = protein
organism = Homo sapiens

SEQUENCE: 23
SEETKENEGF TVTAEGKGQG TLSVVTMYHA KAKDQLTCNK FDLKVTIKPA PETEKRPQDA 60
KNTMILEICT RYRGDQDATM SILDISMNTG FADPTDDLKQ LANGVDRYIS KYELDKAFSD 120
RNTLIIYLDK VSHSEDDCLA FKVHQYFNVE LIQPGAVKVY AYYNLEEST RFYHPEKEKG 180
KLNKLCRDEL CRCAEENCFI QKSDDKVTL ERLDKACEPG VDYVYKTRLV KVQLSNDFDE 240
YIMAIEQTIK SGSDEVQVGQ QRTFISPIKC REALKLEEKK HYLMWGLSSD FWGEKPNLSSY 300
IIGKDTWVEH WPEEDECQDE ENQKQCQDLG AFTESTMVFG CPN 343

SEQ ID NO: 24      moltype = AA length = 17
FEATURE           Location/Qualifiers
source            1..17
mol_type = protein
organism = Homo sapiens

SEQUENCE: 24
SSKITHRIHW ESASLLR 17

SEQ ID NO: 25      moltype = AA length = 583
FEATURE           Location/Qualifiers
source            1..583
mol_type = protein
organism = Homo sapiens

SEQUENCE: 25
MKLLHVFLLE LCFHLRFCKV TYTSQEDLVE KKCLAKKYTH LSCDKVFCQP WQRCIEGTCV 60
CKLPYQCPKN CTAVCATNRS SFPTYCQQKS LECLHPGTKF LNNGTCTAEG KFSVSLKHGN 120
TDSEGIIVEVK LVDQDKTMFI CKSSWSMREA NVACLDLGFQ QGADTQRFRK LSDLSINST 180
CLHVHCRGLE TSLAECTFTV RRTMGYQDFA DVVCYTKAD SPMDFFQCV NGKYISQMKA 240
CDGINDCGDQ SDELCCKACQ GKGFHKSGV CIPSYQQCNG EVDCITGEDE VGCAGFASVT 300
QEETEILITAD MDAERRRIKS LLPKLSCGVK NRHMIRRRI VGGKRAQLGD LPWQVAIKDA 360
SGITCGGIYI GGCWILTAAH CLRASKTHR QIWTVVWDI HPDLKRIVIE YVDRIIFHEN 420
YNAGTYQNDI ALIEMKKDGN KKDCELPRS PACVPWSPYL FQPNDTCIVS GWGREKDNER 480
VFSLQWGEVK LISNCSKFYG NRYFEKEMEC AGTYDGSIDA CKGDSGGPLV CMDANNVTYV 540
WGVSWSGENC GKPEFPGVYT KVANYFDWIS YHVGRPFISQ YNV 583

SEQ ID NO: 26      moltype = AA length = 235
FEATURE           Location/Qualifiers
source            1..235
mol_type = protein
organism = Homo sapiens

SEQUENCE: 26
IVGGKRAQLG DLWPWQVAIKD ASGITCGGIY IGGCWILTAH HCLRASKTHR YQIWTVVWDW 60
IHPDLKRIVI EYVDRIFHE NYNACTYQND IALIEMKKDGN NKKDCELPRS IPACVPWSPY 120
LFQPNNDTCIV SGWGREKDNE RVFSLQWGEV KLISNC SKFY GNRFYEKEME CAGTYDG SID 180
ACKGDGGPL VMCDANNVTYV WGVSWSGEN CGKPEFPGVY TKVANYFDWI SYHVG 235

SEQ ID NO: 27      moltype = length =
SEQUENCE: 27

SEQ ID NO: 28      moltype = AA length = 449
FEATURE           Location/Qualifiers
REGION            1..449

```

-continued

source	note = Synthetic construct 1..449
	mol_type = protein
	organism = synthetic construct
SEQUENCE: 28	
MRLLAIIICL MLWAICVAED CNELPPRRNT EILTGSWSDDQ TYPEGTQAIY KCRPGYRSLG 60	
NVIMVCRKGE WVALNPLRKC QKRPCGHGPQD TPFGTFTLTG GNVFEYGVKA VYTCNEGYQL 120	
LGEINYRECD TDGWTNDIPI CEVVKCLPVT APENGKIVSS AMEPDREYHF GQAVRFVCNS 180	
GYKIEGDEEM HCSDDGFWSK EKPKCVEISC KSPDVINGSP ISQKIIYKEN ERFQYKCNMG 240	
YEYSERGDAV CTESGWRLPLP SCEEKSCDNP YIPNGDYSP RIKHRTGDEI TYQCRNGFYP 300	
ATRGNTAKCT STGWIPAPRC TLKPCDYPDI KHGGLYHENM RRPYFPVAVG KYYSYCCDEH 360	
FETPSGSYWD HIHCTQDGWS PAVPCLRKY FPYLENGYNQ NYGRKFVQGK SIDVACHPGY 420	
ALPKAQTTVT CMENGWSPTP RCIRVKTCRKY FPYLENGYNQ NYGRKFVQGK SIDVACHPGY 449	
SEQ ID NO: 29	moltype = AA length = 1231
FEATURE	Location/Qualifiers
source	1..1231
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 29	
MRLLAIIICL MLWAICVAED CNELPPRRNT EILTGSWSDDQ TYPEGTQAIY KCRPGYRSLG 60	
NVIMVCRKGE WVALNPLRKC QKRPCGHGPQD TPFGTFTLTG GNVFEYGVKA VYTCNEGYQL 120	
LGEINYRECD TDGWTNDIPI CEVVKCLPVT APENGKIVSS AMEPDREYHF GQAVRFVCNS 180	
GYKIEGDEEM HCSDDGFWSK EKPKCVEISC KSPDVINGSP ISQKIIYKEN ERFQYKCNMG 240	
YEYSERGDAV CTESGWRLPLP SCEEKSCDNP YIPNGDYSP RIKHRTGDEI TYQCRNGFYP 300	
ATRGNTAKCT STGWIPAPRC TLKPCDYPDI KHGGLYHENM RRPYFPVAVG KYYSYCCDEH 360	
FETPSGSYWD HIHCTQDGWS PAVPCLRKY FPYLENGYNQ NYGRKFVQGK SIDVACHPGY 420	
ALPKAQTTVT CMENGWSPTP RCIRVKTCRKY SSIDIENGFI SESQYTYALK EKAKYQCKLG 480	
YVITADGETSG SITCGKDGSW AQPTCIKSD IPVFMNARTK NDFTWFKLND TLDYECHDGY 540	
ESNTGSTTG SITCGKDGSW LPICYCERCE LPKIDVHLVP DRKKDQYKVG EVLKFSCKPG 600	
FTIVGPNVSQ CYHFGLSPDL PICKEQVQSC GPPPEELLNGN VKEKTKEEYG HSEVVVEYYCN 660	
PRFLMKGPNK ITCVGEWTT LPVCIVEEST CGDIPELEHG WAQLSSPPYY YGDSVEFNCS 720	
ESPTMIGHRS ITCIHGWTQ LPQCVIAIDKL KKCKSSNLII LEEHHLKNKKE FDHNSNIRYR 780	
CRGKEGWIHRS VTINGRDWP VNCSMAQIQL CPPPPQIPNS HNMHTTLNYR DGEKVSVLCQ 840	
ENYLIQECEE ITCKDGRWQS ITCVGEWTT LPQCVIAIDKL SQPPQIEHGT INSSRSSQES YAHTKLSYT 900	
CEGGFRISEE NETTCYMGKW SSPPQCEGLP CKSPPEISHG VVAHMSDSYQ YGEEVTYKCF 960	
EFGFIDGPATI AKCLGEKWSH PPSCIKTDCL SLPSFENAIP MGEKKDVYKA GEQVITYTCAT 1020	
YYKMDGASNV TCINSRWTQ PTCRDTSCVN PPTVQNAYIV SRQMSKYPGQ ERVRYQCRSP 1080	
YEMFGDEEVM CLNGNWTEPP OCKDSTGKG PPPPIDNGDI TSFPLSVYAP ASSVEYQCQN 1140	
LYOLEGNKRI TCRNGQWSEP PKCLHPCVIS REIMENYNIA LRWTAKQKLY SRTGESVEFV 1200	
CRGQYRLSSR SHTLRTTCWD GKLEYPTCAK R 1231	
SEQ ID NO: 30	moltype = AA length = 644
FEATURE	Location/Qualifiers
source	1..644
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 30	
GHCQAPDHFL FAKLKTQTNA SDFFIGTSLK YECRPEYYGR PFSITCLDNL VWSSPKDVCK 60	
RKSCKTPDPD VNGMVHVITD IQVGSRINY CTTGHLIGH SSAECILSGN AAHWSTKPP 120	
CQRIPCGLPP TIANGDFIST NRENPHYGSV VTYRCNPNGSG GRKVFEVLGE PSIYCTSND 180	
QVGIWGPAP QCIIIPNCKTP PVNENGILVS DNRSLFSLNE VVEFRCPQPGF VMKGPRRVKC 240	
QALNKWEPEL PSCSRVCQPP PDVLHAERTQ RDKDNFSPGQ EVFYSCEPGY DLRGAAASMRC 300	
TPOGDWSPA PTCEVKSCDD FMGQLLNGRV LFPVNQLQGA KVDFVCDEGF QLKGSASAC 360	
VLAGMESLWN SSVPVCEQIF CPSPPVIPNG RHTGKPLEVF PFGKAVNYTC DPHPDRGTSF 420	
DLIGESTIWC TSDPQGNGVW SSPAPRCGIL GHQCAPDHFL FAKLKTQTNA SDFFIGTSLK 480	
YECRPEYYGR PFSITCLDNL VWSSPKDVCK RKSCKTPDPD VNGMVHVITD IQVGSRINY 540	
CTTGHLIGH SSAECILSGN TAHWSTKPP CQRIPCGLPP TIANGDFIST NRENPHYGSV 600	
VTYRCNLGSR GRKVFEVLGE PSIYCTSND QVGIWGPAP QCII 644	
SEQ ID NO: 31	moltype = AA length = 644
FEATURE	Location/Qualifiers
source	1..644
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 31	
GHCQAPDHFL FAKLKTQTQA SDFFIGTSLK YECRPEYYGR PFSITCLDNL VWSSPKDVCK 60	
RKSCKTPDPD VNGMVHVITD IQVGSRIQYS CTTGHLIGH SSAECILSGN AAHWSTKPP 120	
CQRIPCGLPP TIANGDFIST NRENPHYGSV VTYRCNPNGSG GRKVFEVLGE PSIYCTSND 180	
QVGIWGPAP QCIIIPNCKTP PVNENGILVS DNRSLFSLNE VVEFRCPQPGF VMKGPRRVKC 240	
QALNKWEPEL PSCSRVCQPP PDVLHAERTQ RDKDNFSPGQ EVFYSCEPGY DLRGAAASMRC 300	
TPOGDWSPA PTCEVKSCDD FMGQLLNGRV LFPVNQLQGA KVDFVCDEGF QLKGSASAC 360	
VLAGMESLWN SSVPVCEQIF CPSPPVIPNG RHTGKPLEVF PFGKAVNYTC DPHPDRGTSF 420	
DLIGESTIRC TSDPQGNGVW SSPAPRCGIL GHQCAPDHFL FAKLKTQTQA SDFFIGTSLK 480	
YECRPEYYGR PFSITCLDNL VWSSPKDVCK RKSCKTPDPD VNGMVHVITD IQVGSRIQYS 540	
CTTGHLIGH SSAECILSGN TAHWSTKPP CQRIPCGLPP TIANGDFIST NRENPHYGSV 600	
VTYRCNLGSR GRKVFEVLGE PSIYCTSND QVGIWGPAP QCII 644	
SEQ ID NO: 32	moltype = AA length = 25

-continued

FEATURE source	Location/Qualifiers 1..25 mol_type = protein organism = Homo sapiens	
SEQUENCE: 32		
ALIVGTLSTGT IFFILLIIFL SWIIL		25
SEQ ID NO: 33	moltype = AA length = 43	
FEATURE source	Location/Qualifiers 1..43 mol_type = protein organism = Homo sapiens	
SEQUENCE: 33		
KHRKGNNAAHE NPKEVAIHLH SQGGSSVHPR TLQTNEENS R VLP		43
SEQ ID NO: 34	moltype = AA length = 8	
FEATURE REGION	Location/Qualifiers 1..8 note = Synthetic construct	
source	1..8 mol_type = protein organism = synthetic construct	
SEQUENCE: 34		
ENLYFQGS		8
SEQ ID NO: 35	moltype = DNA length = 657	
FEATURE source	Location/Qualifiers 1..657 mol_type = other DNA organism = Homo sapiens	
SEQUENCE: 35		
aagcttgcga ccatgagact gctggccaag atcatctgcc ttagtgcgtg ggccatctgc 60 gtggccggac attgtcaggc ccctgaccac ttctctgtcg ccaagctgaa aaccaggacc 120 aacggccagcg acttcctat cgccaccagg ctgaagtacg agtgcagacc cgagtactac 180 ggcagacccct tcagcatcac ctgttggac aacctctgtg ggtctagccc caaggacgtg 240 tgcaggagaaa agagctgcaa gaccctctt gatccctgtga acggcatggt gcacgtgatc 300 accgacatcc aagtgggcag cagaatcaac tacagctgca ccacccggca cagactgatc 360 ggacactcta ggcggcggatg tatttttttttggcggatggc acactgtgc cacaaggct 420 ccaatctgcc agagaatccc ttggggcttg cttccctacaa tcgccaacgg cgattttatc 480 agcaccacaaca gagagaactt ccactacggc tccgtggta cctacagatg caatcctggc 540 agcggccggca gaaagggtttt cgaacttgcg ggcgagccca gcatctactg caccagcaac 600 gtgaccaag tcggcatttg gageggccct gtcctcagt gcatcatcta agatata 657		
SEQ ID NO: 36	moltype = DNA length = 657	
FEATURE source	Location/Qualifiers 1..657 mol_type = other DNA organism = Homo sapiens	
SEQUENCE: 36		
aagcttgcga ccatgagact gctggccaag atcatctgcc ttagtgcgtg ggccatctgc 60 gtggccggac attgtcaggc ccctgaccac ttctctgtcg ccaagctgaa aaccaggaca 120 caggccagcg acttcctat cgccaccagg ctgaagtacg agtgcagacc cgagtactac 180 ggcagacccct tcagcatcac ctgttggac aacctctgtg ggtctagccc caaggacgtg 240 tgcaggagaaa agagctgcaa gaccctctt gatccctgtga acggcatggt gcacgtgatc 300 accgacatcc aagtgggcag cagaatcaac tacagctgca ccacccggca cagactgatc 360 ggcactcta ggcggcggatg tatttttttttggcggatggc acactgtgc cacaaggct 420 ccaatctgcc agagaatccc ttggggcttg cttccctacaa tcgccaacgg cgattttatc 480 agcaccacaaca gagagaactt ccactacggc tccgtggta cctacagatg caatcctggc 540 agcggccggca gaaagggtttt cgaacttgcg ggcgagccca gcatctactg caccagcaac 600 gtgaccaag tcggcatttg gageggccct gtcctcagt gcatcatcta agatata 657		
SEQ ID NO: 37	moltype = DNA length = 659	
FEATURE source	Location/Qualifiers 1..659 mol_type = other DNA organism = Homo sapiens	
SEQUENCE: 37		
aagcttgcga ccatgagact gctggccaag atcatctgcc ttagtgcgtg ggccatctgc 60 gtggccggcc actgtcaggc ccctgatcac ttctctgtcg ccaagctgaa aaccaggacc 120 aacggccagcg acttcctat cgccaccagg ctgaagtacg agtgcagacc cgagtactac 180 ggcagacccct tcagcatcac ctgttggac aacctctgtg ggtctagccc caaggacgtg 240 tgcaggagaaa agagctgcaa gaccctctt gatccctgtga acggcatggt gcacgtgatc 300 accgacatcc aagtgggcag cagaatcaac tacagctgca ccacccggca cagactgatc 360 ggacactcta ggcggcggatg tatttttttttggcggatggc acactgtgc cacaaggct 420 ccaatctgcc agagaatccc ttggggcttg cttccctacaa tcgccaacgg cgattttatc 480 agcaccacaaca gagagaactt ccactacggc tccgtggta cctacagatg caatcctggc 540 agcggccggca gaaagggtttt cgaacttgcg ggcgagccca gcatctactg caccagcaac 600 gtgaccaag tcggcatttg gageggccct gtcctcagt gcatcatccc taagatata 659		

-continued

SEQ ID NO: 38 moltype = DNA length = 659
 FEATURE Location/Qualifiers
 source 1..659
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 38
 aagcttgcga ccatgagact gctggccaag atcatctgcc tgatgtgtg ggccatctgc 60
 gtggccggcc actgtcaggc ccctgatcac ttctgttcg ccaagtgaa aacccagaca 120
 caggccacgc acttccctat cggcaccago ctgaagtctg agtgacgacc cgagttactac 180
 ggcagacccct tcagcatcac ctgtctgtt ggtctatccc caaggacgtg 240
 tgcagaagaaa agagtgcaa gaceccctctt gatccctgtca acggcatgtt gcacgtgtatc 300
 accgacatcc aagtggcag cagaatccag tacagctgca ccacaggcca cagactgtatc 360
 ggcacactta ggcacccatgt tatecttgao ggaacacagc cccactgtgc cacaaggct 420
 ccaatctgcg agagaatccc ttgcggctt cctccatata tcgccaacagg cgatttcatc 480
 agcacaaca a gagaactt ccactacggc tccgtgggtca cctacatgtt caacctgggc 540
 tccagaggcc ggaagggtttt cgaacttgcg ggccggcata gcatctactg caccagcaac 600
 gacgaccaag tcggcatttg gacggacatc gtcctcgtt gcatcatccc taagatata 659

SEQ ID NO: 39 moltype = DNA length = 714
 FEATURE Location/Qualifiers
 source 1..714
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 39
 aagcttgcga ccatgagact gctggccaag atcatctgcc tgatgtgtg ggccatctgc 60
 gtggccacc accatcacca tcacggcago a cggcagaacc tgtaacttca aggatcttct 120
 ggcggccact gtcaggcccccc tcgtacttc ctgtcgccca agctgaaaac ccagaccaac 180
 gccagcgact tccctatcg caccagccgt aagtacgatc gcaagacccga gtactacggc 240
 agacccttca gcatcacccatc tctggacatc ctgtgtgtt ctggcccaaa ggacgtgtgc 300
 aagagaaaga gtcgaagac ccctctgtat cctgtgaacgc gcatgtgtca cgtgtatcacc 360
 gacatccaa gacccatccc aatccactac agtgcacca ccggccacag actgtatcg 420
 cactctagcg ccgagtttat cctgagccgc aatggccac acgtgttccac caagccttca 480
 atctgcccaga gatcccttg cggccgtt cctacaatcc ccaacggcga tttcatcagc 540
 accaacacag a gaaacttcca ctaacgttcc gttgttccatc acatgtccaa tccctggcagc 600
 ggccggcagaa aggtgttca acttgtggc gagcccgacatc tctactgcac cagcaacgat 660
 gaccaagtcg gcatattggag cggccctgtt cctcgttca tcatctaaga tata 714

SEQ ID NO: 40 moltype = AA length = 231
 FEATURE Location/Qualifiers
 source 1..231
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 40
 MRLLAKIICL MLWAICVAHH HHHHGSSENLYFQGSSGGHQAPDHFLFAKLKTQTNASDF 60
 PIGTSLKYEC RPEYYGRPFS ITCLDNLVWS SPKDVKCRKS CKTPDPVNG MVHVITDIQV 120
 GSRINYSCTT GHRLIGHSSA ECILSGNAAH WSTKPPICQR IPCGLPTIA NGDFISTNRE 180
 NFHYGSVVTVY RCNPNGSGGRK VFELVGEPSI YCTSNDQVG IWSPGPAPQCI I 231

SEQ ID NO: 41 moltype = DNA length = 714
 FEATURE Location/Qualifiers
 source 1..714
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 41
 aagcttgcga ccatgagact gctggccaag atcatctgcc tgatgtgtg ggccatctgc 60
 gtggccacc accatcacca tcacggcago a cggcagaacc tgtaacttca aggatcttct 120
 ggcggccact gtcaggcccccc tcgtacttc ctgtcgccca agctgaaaac ccagacacag 180
 gccagcgact tccctatcg caccagccgt aagtacgatc gcaagacccga gtactacggc 240
 agacccttca gcatcacccatc tctggacatc ctgtgtgtt ctggcccaaa ggacgtgtgc 300
 aagagaaaga gtcgaagac ccctctgtat cctgtgaacgc gcatgtgtca cgtgtatcacc 360
 gacatccaa gacccatccc aatccactac agtgcacca ccggccacag actgtatcg 420
 cactctagcg ccgagtttat cctgagccgc aatggccac acgtgttccac caagccttca 480
 atctgcccaga gatcccttg cggccgtt cctacaatcc ccaacggcga tttcatcagc 540
 accaacacag a gaaacttcca ctaacgttcc gttgttccatc acatgtccaa tccctggcagc 600
 ggccggcagaa aggtgttca acttgtggc gagcccgacatc tctactgcac cagcaacgat 660
 gaccaagtcg gcatattggag cggccctgtt cctcgttca tcatctaaga tata 714

SEQ ID NO: 42 moltype = AA length = 231
 FEATURE Location/Qualifiers
 source 1..231
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 42
 MRLLAKIICL MLWAICVAHH HHHHGSSENLYFQGSSGGHQAPDHFLFAKLKTQTNASDF 60
 PIGTSLKYEC RPEYYGRPFS ITCLDNLVWS SPKDVKCRKS CKTPDPVNG MVHVITDIQV 120
 GSRIQYSCCTT GHRLIGHSSA ECILSGNAAH WSTKPPICQR IPCGLPTIA NGDFISTNRE 180
 NFHYGSVVTVY RCNPNGSGGRK VFELVGEPSI YCTSNDQVG IWSPGPAPQCI I 231

SEQ ID NO: 43 moltype = DNA length = 716

-continued

FEATURE source	Location/Qualifiers 1..716 mol_type = other DNA organism = Homo sapiens
SEQUENCE: 43	
	aagcttgcga ccatgagact gctggccaag atcatctgcc tgatgtgtg ggccatctgc 60 gtggccacc accatcacca tcacggcago agcgagaacc tgcacttca aggatctct 120 ggcgccact gtcaggcccc tgatcacttc ctgttgcaca agctgaaaac ccagaccaa 180 gcacgact tccctatcgg caccagcttg aagtacgact gcagacccga gtactacggc 240 agacccctca gcatcaccc tctggacaad ctgcgtgtgg ctgcgttccaa ggacgtgtgc 300 aagagaaaaga gtcgaagac ccctctgtat cttgtgaacg gcatgtgtca cgtgttccaa 360 gacatccaag tgggcagcag aatcaactac agtgcacca cccggccacag actgtatcgga 420 cactctagcg ccgactgttat cctgcggcact aatgcgcac acgtgttccac caagcttcca 480 atctgcaga gaatcccttg cggcgtgcct cttacaatcc ccaacggcga ttcatcagc 540 accaacagag agaacttcca ctacggctcc gtgttccactt acagatgcaaa tcctggcagc 600 ggcgccagaa aggtgttca gacttgtggc gagcccacca tctactgcac cagcaacgat 660 gaccaagtgcg catatggag cggccctgtt cttcagtgc tcataccctaa gatatc 716
SEQ ID NO: 44	
FEATURE source	moltype = AA length = 231 Location/Qualifiers 1..231 mol_type = protein organism = Homo sapiens
SEQUENCE: 44	
MRLLAIIICL MLWAICVAHH HHHHGSSENLYFQGSSGGHC QAPDHFLFAK LKTQTNASDF 60 PIGTSKYEC RPEYYGRPFS ITCLDNLVWS SPKDVKRKS CKTPDPVNG MVHVITDIQV 120 GSRINYSCCTT GHRLIGHSSA ECILSGNTAH WSTKPPICQR IPCGLPPTIA NGDFISTNRE 180 NFHYGSVVY RCNLGSRGRK VFELVGEPSI YCTSNDQVG IWSGPAPQCI I 231	
SEQ ID NO: 45	
FEATURE source	moltype = DNA length = 716 Location/Qualifiers 1..716 mol_type = other DNA organism = Homo sapiens
SEQUENCE: 45	
	aagcttgcga ccatgagact gctggccaag atcatctgcc tgatgtgtg ggccatctgc 60 gtggccacc accatcacca tcacggcago agcgagaacc tgcacttca aggatctct 120 ggcgccact gtcaggcccc tgatcacttc ctgttgcaca agctgaaaac ccagacacag 180 gcacgact tccctatcgg caccagcttg aagtacgact gcagacccga gtactacggc 240 agacccctca gcatcaccc tctggacaad ctgcgtgtgg ctgcgttccaa ggacgtgtgc 300 aagagaaaaga gtcgaagac ccctctgtat cttgtgaacg gcatgtgtca cgtgttccaa 360 gacatccaag tgggcagcag aatccctgtat cgcgttccac cggccacac agtgcacca 420 cactctagcg ccgactgttat cctgcggcact aacacggcc acgtgttccac caagcttcca 480 atctgcaga gaatcccttg cggcgtgcct cttacaatcc ccaacggcga ttcatcagc 540 accaacagag agaacttcca ctacggctcc gtgttccactt acagatgcaaa cctggctcc 600 agaggccgaa aggtgttca gacttgtggc gagccataca tctactgcac cagcaacgac 660 gaccaagtgcg catatggag cggccctgtt cttcagtgc tcataccctaa gatatc 716
SEQ ID NO: 46	
FEATURE source	moltype = AA length = 231 Location/Qualifiers 1..231 mol_type = protein organism = Homo sapiens
SEQUENCE: 46	
MRLLAIIICL MLWAICVAHH HHHHGSSENLYFQGSSGGHC QAPDHFLFAK LKTQTNASDF 60 PIGTSKYEC RPEYYGRPFS ITCLDNLVWS SPKDVKRKS CKTPDPVNG MVHVITDIQV 120 GSRIQYSCCTT GHRLIGHSSA ECILSGNAAH WSTKPPICQR IPCGLPPTIA NGDFISTNRE 180 NFHYGSVVY RCNPNGSGGRK VFELVGEPSI YCTSNDQVG IWSGPAPQCI I 231	
SEQ ID NO: 47	
FEATURE REGION source	moltype = AA length = 212 Location/Qualifiers 1..212 note = Synthetic construct 1..212 mol_type = protein organism = synthetic construct
SEQUENCE: 47	
MRLLAIIICL MLWAICVAGH CQAPDHFLFA KLKTQTNASD FPIGTSKYEC CRPEYYGRPF 60 STCLDNLVW SSPKDVKRKS SCKTPDPVN GMVHVITDIQ VGSRINYSCCTT GHRLIGHSS 120 ACILSGNAAH WSTKPPICQ RIPCGLPPTI ANGDFISTNRE ENFHGSVVY YRCNPNGSGGR 180 KVFEVLGEPSI YCTSNDQVG GIWSGPAPQCI II 212	
SEQ ID NO: 48	
FEATURE REGION source	moltype = AA length = 212 Location/Qualifiers 1..212 note = Synthetic construct 1..212 mol_type = protein organism = synthetic construct

-continued

SEQUENCE: 48

MRLLAIIICL	MLWAICVAGH	CQAPDHFLFA	KLKTQTQASD	FPIGTSLSKYE	CRPEYYGRPF	60
SITCLDNLVW	SSPKDVCKRK	SCKTPDPVN	GMVHVITDIQ	VGSRIQYSCT	TGHRLIGHSS	120
AECILSGNAA	HWSRKPPICQ	RIPCGLPPTI	ANGDFISTNR	ENFHYGSVVT	YRCNPMSGGR	180
KVFELVGEPS	IYCTSNDQV	GIWSPAPQC	II			212

SEQ ID NO: 49 moltype = AA length = 212
 FEATURE Location/Qualifiers
 REGION 1..212
 note = Synthetic construct
 source 1..212
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 49

MRLLAIIICL	MLWAICVAGH	CQAPDHFLFA	KLKTQTQASD	FPIGTSLSKYE	CRPEYYGRPF	60
SITCLDNLVW	SSPKDVCKRK	SCKTPDPVN	GMVHVITDIQ	VGSRIQYSCT	TGHRLIGHSS	120
AECILSGNTA	HWSRKPPICQ	RIPCGLPPTI	ANGDFISTNR	ENFHYGSVVT	YRCNLGSRGR	180
KVFELVGEPS	IYCTSNDQV	GIWSPAPQC	II			212

SEQ ID NO: 50 moltype = AA length = 212
 FEATURE Location/Qualifiers
 REGION 1..212
 note = Synthetic construct
 source 1..212
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 50

MRLLAIIICL	MLWAICVAGH	CQAPDHFLFA	KLKTQTQASD	FPIGTSLSKYE	CRPEYYGRPF	60
SITCLDNLVW	SSPKDVCKRK	SCKTPDPVN	GMVHVITDIQ	VGSRIQYSCT	TGHRLIGHSS	120
AECILSGNTA	HWSRKPPICQ	RIPCGLPPTI	ANGDFISTNR	ENFHYGSVVT	YRCNLGSRGR	180
KVFELVGEPS	IYCTSNDQV	GIWSPAPQC	II			212

SEQ ID NO: 51 moltype = AA length = 406
 FEATURE Location/Qualifiers
 REGION 1..406
 note = Synthetic construct
 source 1..406
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 51

MRLLAIIICL	MLWAICVAGH	CQAPDHFLFA	KLKTQTQASD	FPIGTSLSKYE	CRPEYYGRPF	60	
SITCLDNLVW	SSPKDVCKRK	SCKTPDPVN	GMVHVITDIQ	VGSRIQYSCT	TGHRLIGHSS	120	
AECILSGNAA	HWSRKPPICQ	RIPCGLPPTI	ANGDFISTNR	ENFHYGSVVT	YRCNPMSGGR	180	
KVFELVGEPS	IYCTSNDQV	GIWSPAPQC	II	IIGHCQADPH	FLFAKLKTQT	NASDFPIGTS	240
LKYECRPEYY	GRFSITCLD	NLVWSSPKDV	CKRKSCCKTPP	DPVNGMVHVI	TDIQVGSRIN	300	
YSCTTGHLI	GHSSAECILS	GNTAHWSTKP	PICQRIPCG	PPTIANGDFI	STNRENFH	360	
SVVTYRCNLG	SRGRKVFELV	GEPSIYCTSN	DDQVGIWSPG	APQCII		406	

SEQ ID NO: 52 moltype = AA length = 406
 FEATURE Location/Qualifiers
 REGION 1..406
 note = Synthetic construct
 source 1..406
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 52

MRLLAIIICL	MLWAICVAGH	CQAPDHFLFA	KLKTQTQASD	FPIGTSLSKYE	CRPEYYGRPF	60	
SITCLDNLVW	SSPKDVCKRK	SCKTPDPVN	GMVHVITDIQ	VGSRIQYSCT	TGHRLIGHSS	120	
AECILSGNAA	HWSRKPPICQ	RIPCGLPPTI	ANGDFISTNR	ENFHYGSVVT	YRCNPMSGGR	180	
KVFELVGEPS	IYCTSNDQV	GIWSPAPQC	II	IIGHCQADPH	FLFAKLKTQT	QASDFPIGTS	240
LKYECRPEYY	GRFSITCLD	NLVWSSPKDV	CKRKSCCKTPP	DPVNGMVHVI	TDIQVGSRIQ	300	
YSCTTGHLI	GHSSAECILS	GNTAHWSTKP	PICQRIPCG	PPTIANGDFI	STNRENFH	360	
SVVTYRCNLG	SRGRKVFELV	GEPSIYCTSN	DDQVGIWSPG	APQCII		406	

SEQ ID NO: 53 moltype = AA length = 406
 FEATURE Location/Qualifiers
 REGION 1..406
 note = Synthetic construct
 source 1..406
 mol_type = protein
 organism = synthetic construct

SEQUENCE: 53

MRLLAIIICL	MLWAICVAGH	CQAPDHFLFA	KLKTQTQASD	FPIGTSLSKYE	CRPEYYGRPF	60	
SITCLDNLVW	SSPKDVCKRK	SCKTPDPVN	GMVHVITDIQ	VGSRIQYSCT	TGHRLIGHSS	120	
AECILSGNAA	HWSRKPPICQ	RIPCGLPPTI	ANGDFISTNR	ENFHYGSVVT	YRCNPMSGGR	180	
KVFELVGEPS	IYCTSNDQV	GIWSPAPQC	II	IIGHCQADPH	FLFAKLKTQT	QASDFPIGTS	240
LKYECRPEYY	GRFSITCLD	NLVWSSPKDV	CKRKSCCKTPP	DPVNGMVHVI	TDIQVGSRIQ	300	
YSCTTGHLI	GHSSAECILS	GNTAHWSTKP	PICQRIPCG	PPTIANGDFI	STNRENFH	360	
SVVTYRCNLG	SRGRKVFELV	GEPSIYCTSN	DDQVGIWSPG	APQCII		406	

```

SEQ ID NO: 54          moltype = AA  length = 406
FEATURE
REGION
1..406
note = Synthetic construct
1..406
source
mol_type = protein
organism = synthetic construct

SEQUENCE: 54
MRLLAIIICL MLWAICVAGH CQAPDHFLFA KLKTQTQASD FPIGTSALKYE CRPEYYGRPF 60
SITCLDNLVW SSPKDVKCRK SCKTPDPVN GMVHVITDIQ VGSRIQYSCT TGHLIGHSS 120
AECILSGNAA HWSTKPPICQ RIPCGLPPTI ANGDFISTNR ENFHYGSVVT YRCNPMSGGR 180
KVPELVGEPS IYCTSNDQV GIWSGPAPQC IIGHCQAPDH FLFAKLKTQT NASDFPIGTS 240
LYKECRPEYY GRPFSITCLD NLVWSPKD VCKRKSCKTPP DPVNGMVHVI TDIQVGSRIN 300
YSCTTGHLI GHSSAECILS GNTAHWSTKP PICQRIPCGL PPTIANGDFI STNRENFHYG 360
SVTYRCNLG SRGRKVFELV GEPSIYCTSN DDQVGIWSPG APQCI 406

```

The invention claimed is:

1. A method for inhibiting complement activation in a subject, the method comprising administering or expressing a polypeptide which is capable of binding to C3b, the polypeptide comprising an amino acid sequence wherein said amino acid sequence consists of a sequence having at least 90% identity to SEQ ID NO:2, wherein the amino acid at position 111 of SEQ ID NO:2 is A, the amino acid at position 157 of SEQ ID NO:2 is P, and the amino acid at position 160 of SEQ ID NO:2 is G, and wherein the polypeptide has a total length of 450 amino acids or fewer.
2. The method according to claim 1, wherein the subject has a disease associated with complement dysregulation or overactivation, wherein the disease is not an ocular disease.
3. The method according to claim 1, wherein the subject has a kidney disease or condition.
4. The method according to claim 1, wherein the subject has atypical Haemolytic Uremic Syndrome (aHUS), Membranoproliferative Glomerulonephritis Type II (MPGN II), sepsis, or Paroxysmal nocturnal hemoglobinuria (PNH).
5. The method according to claim 1, wherein the polypeptide has a total length of 250 amino acids or fewer.
6. The method according to claim 1, wherein the method comprises administering a nucleic acid encoding the polypeptide.
7. The method according to claim 6, wherein the nucleic acid comprises a sequence having at least 85% sequence identity to SEQ ID NO:35.
8. The method according to claim 1, wherein the method comprises modifying at least one cell of a subject to express or comprise a nucleic acid encoding the polypeptide.
9. The method according to claim 8, wherein the nucleic acid comprises a sequence having at least 85% sequence identity to SEQ ID NO:35.
10. The method according to claim 1, wherein the method comprises administering a vector comprising a nucleic acid encoding the polypeptide to at least one cell of a subject.

11. The method according to claim 10, wherein the nucleic acid comprises a sequence having at least 85% sequence identity to SEQ ID NO:35.

12. The method according to claim 10, wherein the vector is a viral vector.

13. A method according to claim 10, wherein the vector is a lentiviral vector, a retroviral vector, a gammaretroviral vector, an adenoviral vector, an adeno-associated viral (AAV) vector, a vaccinia virus vector, or a herpesvirus vector.

14. A method according to claim 10, wherein the vector is an adeno-associated viral (AAV) vector comprising a serotype selected from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, or AAV11, or hybrids and/or mutants thereof.

15. The method according to claim 1, wherein the polypeptide or amino acid sequence comprises or consists of SEQ ID NO:2 or SEQ ID NO:3.

16. The method according to claim 1, wherein the polypeptide or amino acid sequence comprises a secretory pathway sequence comprising, or consisting of, SEQ ID NO: 7.

17. The method according to claim 16, wherein the polypeptide or amino acid sequence comprises, or consists of, a sequence having at least 95% sequence identity to SEQ ID NO:47 or 49.

18. The method according to claim 16, wherein the polypeptide or amino acid sequence comprises a cleavage site for removing the secretory pathway sequence.

19. The method according to claim 1, wherein the method comprises simultaneous or sequential administration or expression of Complement Factor I.

20. The method according to claim 1, wherein the method comprises simultaneous or sequential administration or expression of a nucleic acid encoding Complement Factor I.

* * * * *