Supporting Information

New compstatin peptides containing N-terminal extensions and non-natural amino acids exhibit potent complement inhibition and improved solubility characteristics

Ronald D. Gorham Jr.,[†] David L. Forest,[‡] George A. Khoury,[⊥] James Smadbeck,[⊥] Consuelo N. Beecher,[§] Evangeline D. Healy, [‡] Phanourios Tamamis,[¶,⊥] Georgios Archontis,[¶] Cynthia K. Larive,[§] Christodoulos A. Floudas,[⊥] Monte J. Radeke,[‡] Lincoln V. Johnson,[‡] and Dimitrios Morikis[†,]*

[†]Department of Bioengineering, University of California, Riverside, CA 92521, USA

[‡]Center for the Study of Macular Degeneration, Neuroscience Research Institute, University of California, Santa Barbara, CA 93106, USA

[⊥] Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544, USA

[§]Department of Chemistry, University of California, Riverside, CA 92521, USA

[§]Department of Physics, University of Cyprus, PO20537, CY1678 Nicosia, Cyprus

Corresponding Author: Dimitrios Morikis, Department of Bioengineering, University of California, Riverside, CA 92521; Tel.: +1-951-827-2696; Fax: +1-951-827-6416; E-mail address: dmorikis@ucr.edu.

Table of Contents:

Supporting Tables	S2
Supporting Figure	S13
References	S14

Table S1. List of candidate peptides computationally evaluated in the peptide re-design with non-natural amino acids, and their calculated metrics.[a]

Sequence Solution	Position 4[b]	Position 9[c]	Contact Score	K*	K* rank
7	**meW**	**Nal**	**1.015974**	**2.709E+27**	**1**
12	**Trp**	**Rea**	**0.02375**	**3.082E+12**	**2**
14	**Trp**	**Aal**	**0.021505**	**3.803E+11**	**3**
6	**meW**	**Ala**	**1.017813**	**4.489E+10**	**4**
17	**Trp**	**Ala**	**0.019472**	**1.829E+07**	**5**
2	meW	Sea	1.033698	7.042E+02	6
3	meW	Rea	1.022091	1.315E+02	7
10	**Trp**	**Sea**	**0.035357**	**1.274E+01**	**8**
15	Nmw	Aal	0.021354	5.842E-02	9
19	Nmw	Ala	0.019322	3.412E-02	10
1	meW	Nma	1.862203	3.818E-04	11
11	Nmw	Sea	0.035207	1.069E-06	12
4	meW	Aal	1.019845	1.746E-14	13
16	Trp	Pal	0.019472	1.615E-14	14
18	Nmw	Pal	0.019322	8.196E-15	15
20	**Trp**	**Nal**	**0.017634**	**2.607E-19**	**16**
8	Trp	Nma	0.863863	5.197E-21	17
13	Nmw	Rea	0.0236	3.853E-26	18
5	meW	Pal	1.017813	1.314E-31	19
21	Nmw	Nal	0.017483	1.958E-36	20
9	Nmw	Nma	0.863712	1.656E-53	21

[a] Peptides in bold were experimentally tested. Peptides in bold and red color are controls.

[b] Amino acid abbreviations in position 4: tryptophan (Trp), 5-hydroxytryptophan, 5-methyltryptophan, N-methyltryptophan (Nmw), 1-methyltryptophan (meW), and 7-hydroxytrptophan.

[C] Amino acid abbreviations in position 9: alanine (Ala), 1-naphthylalanine (Nal), N-methylalanine (Nma), Pyrenylalanine (Pal), (R)-α-ethylalanine (Rea), (S)-α-ethylalanine (Sea), and R(+)-α-allyl-alanine (Aal).

Table S2. List of fold specificity (FS) rankings for peptides chosen for approximate binding affinity validation, in the *de novo* peptide design with natural amino acids. Rankings are out of the consensus 1019 sequences from sequence selection.[a]

Sequence ID	Position -1	Position 0	Position 1	Peptide 9 FS Rank	SS-W4A9 FS Rank	Combined FS Rank
Tri1	Asn	Arg	Asp	-	1	1
Tri2	Asn	Asn	Asn	1	5	3
Tri3	Asn	Arg	Asn	2	4	2
Tri4	Asp	Asn	Asn	3	2	4
Tri5	Gln	Asn	Asn	6	-	-
Tri6	Asp	Arg	Trp	8	13	16
Tri7	Asn	Glu	Asn	-	12	10
Tri8	Asn	Gln	Asp	-	9	13
Tri9	Asn	Thr	Asn	11	19	11
Tri10	Glu	Arg	Trp	22	15	12
Tri11	Asp	Trp	Arg	-	14	15
Tri12	Asn	Asn	Leu	12	20	11
Tri13	Asn	Asn	Tyr	13	23	20
Tri14	Asn	Arg	Leu	15	17	17
Tri15	Asp	Asn	Phe	-	-	22
Di1	Asp	Arg	Ile	9	10	14
Di2	Glu	Arg	Ile	19	18	21
Di3	Asp	Asn	Ile	35	33	19
Di4	Asn	Asn	Ile	31	38	34
Di5	Arg	Gln	Ile	46	-	-
Di6	Asn	Gln	Ile	47	-	-
Di7	Asp	Gln	Ile	50	-	-

[a] "-" indicates the peptide was not ranked in the top 50 for the structure set.

Table S3. List of peptides with approximate binding affinities (K*)[1] calculated to be higher than the reference peptide, W4A9, in the *de novo* peptide design with natural amino acids. Note that the Boltzmann average energy for the protein, q_P, is not provided as no mutations were introduced into the protein so it is a constant in the calculation.[a]

Sequence ID	Position -1	Position 0	Position 1	q_L	q_{PL}	K*
Native	-	-	Ile	6.16E+239	1.83E+16	6.40E+0
Tri8	Asn	Gln	Asp	1.31E+261	6.06E+20	4.13E+17
Tri10	Glu	Arg	Trp	2.75E+250	6.71E+19	7.82E+7
*Tri12	Asn	Asn	Leu	3.06E+254	2.76E+19	2.11E+12
*Tri14	Asn	Arg	Leu	1.39E+245	7.74E+19	3.43E+2
Tri15	Asp	Asn	Phe	2.62E+261	1.46E+20	3.42E+18
*Di2	Glu	Arg	Ile	2.99E+262	2.47E+19	2.30E+20

[a] * indicates a peptide chosen for experimental testing.

Table S4. Summary of IC50 values for compstatin peptides in C3b and C5b-9 ELISAs and hemolytic assays.

Peptide	C3b ELISA (μM)[a]	C5b-9 ELISA (μM)[a]	Hemolysis (μM)[a]
1	1.21 +/- 0.31	0.95 +/- 0.09	0.90 +/- 0.06
2	2.27 +/- 0.39	2.35 +/- 0.25	1.27 +/- 0.13
3	0.77 +/- 0.14	0.83 +/- 0.08	0.69 +/- 0.13
4	0.65 +/- 0.17	0.84 +/- 0.09	0.62 +/- 0.10
5	0.45 +/- 0.14	0.71 +/- 0.07	0.48 +/- 0.08
6	0.84 +/- 0.06	0.50 +/- 0.06	0.78 +/- 0.21
7	0.57 +/- 0.05	0.46 +/- 0.09	0.30 +/- 0.10
8	0.53 +/- 0.04	0.45 +/- 0.04	0.25 +/- 0.09
9	0.58 +/- 0.10	0.59 +/- 0.05	0.25 +/- 0.11
10	2.85 +/- 0.11	1.76 +/- 0.11	2.30 +/- 0.19
11	2.75 +/- 0.11	1.88 +/- 0.09	1.90 +/- 0.18
12	1.90 +/- 0.10	1.20 +/- 0.06	1.13 +/- 0.10
13	2.89 +/- 0.14	1.99 +/- 0.11	1.74 +/- 0.14
14	3.85 +/- 0.17	2.37 +/- 0.21	2.79 +/- 0.17
15	3.71 +/- 0.19	2.40 +/- 0.25	2.17 +/- 0.15
16	2.20 +/- 0.11	1.34 +/- 0.09	1.23 +/- 0.12
17	4.06 +/- 0.22	2.35 +/- 0.16	2.68 +/- 0.18
18	1.26 +/- 0.11	0.88 +/- 0.13	1.49 +/- 0.13
19	1.31 +/- 0.09	0.82 +/- 0.13	1.06 +/- 0.14
20	1.08 +/- 0.07	0.65 +/- 0.10	1.01 +/- 0.11
W4A9	0.77 +/- 0.06	0.60 +/- 0.04	0.66 +/- 0.07
meW4A9	0.43 +/- 0.04	0.49 +/- 0.06	0.59 +/- 0.06
Parent	6.17 +/- 1.03	13.00 +/- 1.13	3.91 +/- 0.91

[a] Values represent mean IC50 value (from three independent experiments) +/- 95% confidence intervals.

Table S5. Statistical analysis (2-tail Mann-Whitney U test) of compstatin peptides compared to the positive control in the RPE cell *in vitro* assay.

Sample	Significance		*p*-value	
	72810	**81309**	**72810**	**81309**
Parent	ns	ns	0.96810	0.94490
Linear	ns	ns	0.30772	0.27134
W4A9	*	*	0.00100	0.00034
Peptide 5	*	*	0.01140	0.00132
Peptide 8	*	*	0.00044	0.00168
Peptide 12	*	*	0.00222	0.00362
Peptide 18	*	*	0.00132	0.00168
Peptide 19	*	*	0.00018	0.01390

* significant difference, p-value $\leq$ 0.05, ns: non-significant

Table S6. Statistical analysis (2-tail Mann-Whitney U test) of compstatin peptides compared to the W4A9 in the RPE cell *in vitro* assay.

Sample	Significance		*p*-value	
	72810	**81309**	**72810**	**81309**
Parent	*	*	0.00058	0.00018
Linear	*	*	0.01140	0.00034
Peptide 5	ns	ns	0.10310	0.08914
Peptide 8	ns	*	0.27134	0.02574
Peptide 12	ns	*	0.24200	0.06432
Peptide 18	ns	ns	0.56868	0.07508
Peptide 19	*	*	0.03156	0.01732

* significant difference, p-value ≤ 0.05, ns: non-significant

Table S7. Apparent solubility of compstatin peptides.

Peptide	Solubility (mg/ml)[a]
1	0.11 ± 0.03
2	0.15 ± 0.01
3	3.60 ± 0.44
4	3.35 ± 0.17
5	3.47 ± 0.17
6	4.29 ± 0.26
7	3.92 ± 0.35
8	1.05 ± 0.07
9	3.69 ± 0.16
10	0.38 ± 0.04
11	0.27 ± 0.12
12	0.34 ± 0.14
13	0.39 ± 0.04
14	0.16 ± 0.03
15	0.27 ± 0.05
16	0.30 ± 0.07
17	0.20 ± 0.05
18	2.84 ± 2.30
19	2.11 ± 0.21
20	2.17 ± 0.22
W4A9	3.22 ± 0.11
meW4A9	1.92 ± 0.13
Parent	4.46 ± 0.25

[a] Values represent mean IC50 value (from three independent measurements of same sample) +/- standard deviation.

Table S8. RP-HPLC retention times and factors for compstatin peptides, separated by substitution patterns as in Table 1.

Peptide	Retention Time (min)	*k*	log(*k*)	Retention Time (min)	*k*	log(*k*)
	Method 1 (32% B)			Method 2 (28% B)		
1	15.786 ± 0.019	10.132 ± 0.005	1.006			
2	8.511 ± 0.013	5.002 ± 0.006	0.699			
3	3.275 ± 0.006	1.310 ± 0.008	0.117	6.976 ± 0.015	3.895 ± 0.008	0.591
4	3.355 ± 0.002	1.366 ± 0.006	0.136	7.276 ± 0.009	4.106 ± 0.007	0.613
5	3.131 ± 0.006	1.208 ± 0.010	0.082	6.459 ± 0.004	3.533 ± 0.007	0.548
6	1.868 ± 0.002	0.317 ± 0.020	-0.499	2.309 ± 0.001	0.620 ± 0.015	-0.208
7	1.979 ± 0.001	0.396 ± 0.015	-0.402	2.917 ± 0.001	1.047 ± 0.010	0.020
8	2.185 ± 0.001	0.541 ± 0.011	-0.267	3.519 ± 0.002	1.469 ± 0.009	0.167
9	2.118 ± 0.001	0.494 ± 0.013	-0.306	3.346 ± 0.002	1.348 ± 0.009	0.130
10	3.297 ± 0.007	1.325 ± 0.010	0.122			
11	4.381 ± 0.005	2.089 ± 0.006	0.320			
12	5.231 ± 0.010	2.689 ± 0.007	0.430			
13	5.122 ± 0.003	2.612 ± 0.005	0.417			
14	6.053 ± 0.002	3.269 ± 0.004	0.514			
15	9.782 ± 0.005	5.899 ± 0.004	0.771			
16	11.768 ± 0.017	7.299 ±0.005	0.863			
17	11.452 ± 0.016	7.076 ± 0.005	0.850			
18	1.858 ± 0.005	0.310 ± 0.030	-0.509	2.104 ± 0.003	0.477 ± 0.020	-0.322
19	3.077 ± 0.001	1.170 ± 0.006	0.068	7.134 ± 0.003	4.007 ± 0.006	0.603
20	3.168 ± 0.001	1.234 ± 0.006	0.091	7.483 ± 0.003	4.251 ± 0.006	0.628
W4A9	2.778 ± 0.002	0.959 ± 0.008	-0.018	5.226 ± 0.002	2.668 ± 0.007	0.426
meW4A9	4.016 ± 0.007	1.832 ± 0.007	0.263	9.834 ± 0.005	5.901 ± 0.006	0.771
Parent	1.932 ± 0.003	0.363 ± 0.020	-0.440	2.582 ± 0.002	0.812 ± 0.013	-0.090

[a] Values represent mean IC50 value (from three independent measurements of same sample) +/- standard deviation. Methods 1 and 2 are described in the Experimental Section.

Table S9. List of compstatin peptides in order of increasing lipophilicity based on their log(*k*) using Method 2.

Peptide	**log(*k*)**
18	-0.509
6	-0.499
Parent	-0.440
7	-0.402
9	-0.306
8	-0.267
W4A9	-0.018
19	0.068
5	0.082
20	0.091
3	0.117
10	0.122
4	0.136
meW4A9	0.263
11	0.320
13	0.417
12	0.430
14	0.514
2	0.699
15	0.771
17	0.850
16	0.863
1	1.006

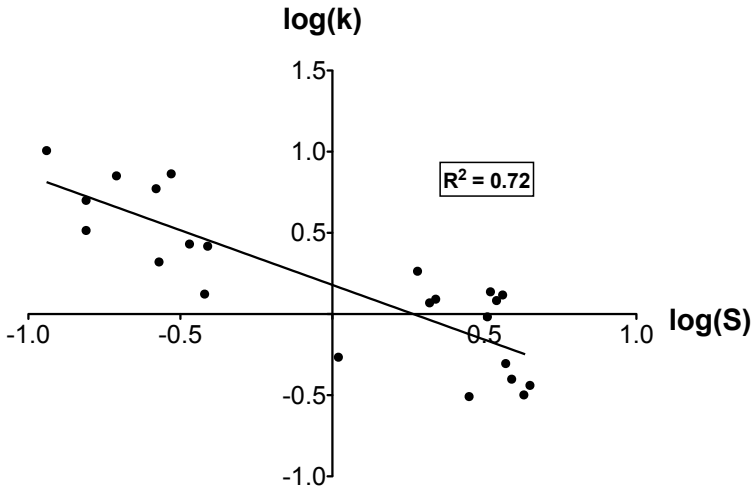

Figure S1. Log-log plot of retention factor (lipophilicity) versus solubility. The trendline shows negative correlation between log(*k*) and log(*S*) values.

References

(1) Smadbeck, J.; Peterson, M. B.; Khoury, G. A.; Taylor, M. S.; Floudas, C. A. Protein WISDOM: a Workbench for in Silico De Novo Design of Biomolecules. *J. Vis. Exp.* **2013**, e50476–e50476.