

Supplemental Data for Hill et al.

A

```

olmf4 MRPGLSFLALLLFFLQGAAGDLGVDGPPVSPGFSFQVDSSS----SFSS---SSRS
olmf1 MSVPLKIKVVLSTAMITNMW---SQTLSLVSQ
olmf2 -----MAAALFPRLLLLPLV-----LILS-----G
myoc MRFFCAR-CC-----SPFGPEMPAVQLLLACL-----VWDV-----G
olmf3 -----MGSTPLL-----ILFL-----SWSG-----P
          :
          :
olmf4 SSSS-RSLGSGGSV-----SQLFSNFTGSDVDRGTCQCSVSLPDTFFVDR-VE
olmf1 LNTKLSAAGGTLDRSTGVLPTNPEESQVYSSAQDSEGRICITVVAPOQTMCSRDRART
olmf2 RPTADS-----KVF-----GLDQVMTSESGDCRCKIMRPLSKDACSVRVSG
myoc ART-AQL-----RK-ANDQSGRCQYTFVAVS-----
olmf3 LQG-QQH-----HL-VEYME-----RRLALE-----
          :
          :
olmf4 RLE--F-TAHLV-----SQKFEKELSKVREYVQLIS
olmf1 KQ-----L-----RQLLEKVNMSQSI
olmf2 RARVEDFYVETVSSGTDRCSTAPPSSLNPCENEMKELKQAPPELLKQSMVLDLE
myoc -----PNSSCPEQSQAMSVIHLQRD-----SSTQLDLEATKARLSLE
olmf3 -----ERLACQDQSSR-----HAAELRDFKNK-----
          :
          :
olmf4 VYEK--KL-----LNLTVRIDIMEKDTISYTELDFELI
olmf1 VL-----DRRTQRDLQVVEKMNQKGLSEKFKQVEESHKQHLARQKAI
olmf2 GTLYSMDLMKVH-----AYV-----HKVASQNTLEESIKANLSRENEV
myoc SLLHLQLTLDQAARQETQEGQLRELGLRER-----DQLETQRELET-AYSNLRKRSVL
olmf3 -----MLPLLEVA-----
          :
          :
olmf4 KVEKMEKELVQLKESFGSSSEIVDQLEVEIRNMTLLVEKLETLDKNNVLAIRREIVAL
olmf1 KAKMDELRLPLVLE-EYKADAKLVQ-----FKEEVQNL
olmf2 KDSV--RHLSQRL-HYENHSAIMLG-----IKKELSLQLLQKDA--AAPATPATGTGSK
myoc EEK--KRLRQE-----NENLARLES-----SQEVARL
olmf3 EKER--EALRTE-----ADTISGRVDR-----LEREVYL
          :
          :
olmf4 KTKLKECEASKDQNTFV-----VHPPTPGS
olmf1 TSVLNE-----LQEEIGAYDYDEL-----QSRVNLEERLRACM
olmf2 GLQLLQKDAAPATPATGTGSKADQTAGKGDISKYGSQKSPADRQ--LPPKPEKL
myoc -----RRG--CQPTRTDARAIVPPGSR-----VSTWNLDTLAFQELKSELTEVPASRI
olmf3 -----ETQNPALPCVEFDEKV-TGGPGTKGRRNEKYMVDV-----
          :
          :
olmf4 -----QKLA
olmf1 LQVEKLRKESGKGSFLQPTAKFRALAAQAVIRGFTYKAGQEVTEAVADNTLQGTSWL
olmf2 -----KESPSG-----YLRSGEGDTG
myoc -----
olmf3 -----
    
```

B

```

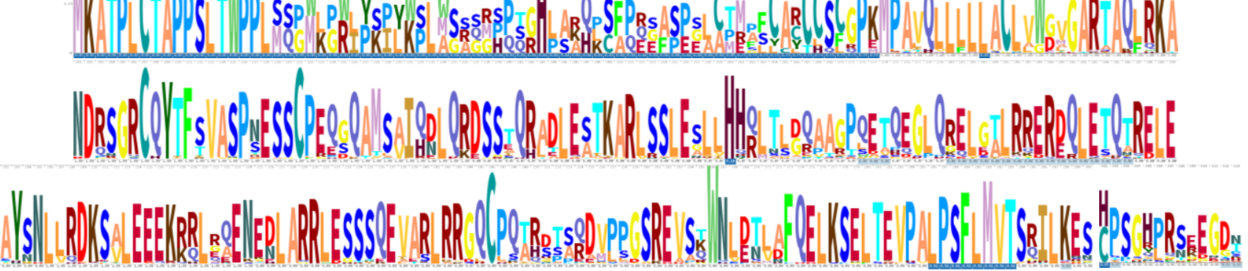
myocilin -----MRFFCARCCSGPEMPAVQL-----LLAC
olmf1 MSVPLKIKVVLSTAMITNMW---SQTLSLVSQ
          :
          :
myocilin LWVDVGARTAQLRKANDQSGRCQYTFVAVSAPNESSCPEQSQAMSVIHLN-----QRDSS-TQ
olmf1 ESQWV-----YSSAQDSEGRICITV--VAPQTMCSRDRARTKQLRQLLEKVNMSQSI
          :
          :
myocilin RLD-----LEATKARLSLESLLHQLTLDQAARQETQEGQLRELGLRERDQLE
olmf1 VLDRRTQRDLQVVEKMNQKGLSEKFKQVEES-----HKQHLARQKAIKAKMDEL
          :
          :
myocilin TQRELETAYSNLRKRSVLEEEKKRLRQENENLARRLESSQVEARLRGQCPQTRDTA
olmf1 PLIPLEE-Y-----KADAKLV-----LQFKEEVQNTLSVNL
          :
          :
myocilin RAVPQSGREVSTWNLDTLAFQELKSELTEVPASRIKESPSGYLRSGEQDTG
olmf1 -----QEIGAYDY-----DELQSRVNLEER--LR-----ACMK-----LA
          :
          :
    
```

C

```

myocilin -----MRFFCAR-----CCSGPEMPAVQLLLACL-----WD-----
olmf4 MRPGLSFLALLLFFLQGAAGDLGVDGPPVSPGFSFQVDSSSSFSSSSRSGSSSRL
          :
          :
myocilin -VGAR-----TAQLRKANDQSGRCQYTFVAVSAPNESSCPEQSQAMSVI-HN-----LQRDS-
olmf4 GSGGSVSQLFSNFTGSDVDRGTCQCSVSLPDTFFP--VDRVERLEFTAHLVSKQFEKELS
          :
          :
myocilin -----STQRDLDEATKARLSLESLLHQLTLDQA
olmf4 KRVREVQLISVYEKLLNLTVRIDIMEKDTISYTELDFELIKVEKEMKLVIQ-----
          :
          :
myocilin ARPQETQEGQLRELGLRERDQLETQRELETAYSNLRKRSVLEEEKKRLRQENENLA
olmf4 -----LKESFGSSEIVDQLEVEIRN-----LLVEKLETLDRN-NVLA
          :
          :
myocilin RRLESSQVEARLRGQCPQTRDTA--RAVPQSGREVSTWNLDTLAFQELKSELTEVPAS
olmf4 IREIV-----ALKTKLKECEASKDQNTFVHPPTPGS-----
          :
          :
myocilin RILKESPSGYLRSGEQDTG
olmf4 -----
    
```

G



D

```

myocilin MRFFCARCCSGPEMPAVQLLLACLWVDVGARTA-----QLRKA-NDQSGRCQY
olmf2 -----MAAALFPRLLLLPLVLLLSGRFTRADSKVFDLDDQVMTSESGDCRCK
          :
          :
myocilin TFS-----VASPNE-----SSCPEQSQAMSVI-----NLQR
olmf2 IMRPLSKDACSVRSGRARVEDFYVETVSSGDCRCKTAPPSSLNPCENEMKELK
          :
          :
myocilin DSSTQRDLDEAT-----KAR-----LSSLESLLHQLTLDQAARQETQEGQLRELGLRRE
olmf2 Q-APPELLKQSMVLDLEGLTYSMDLMKVHAYVHRKVASQMTLEESIKANLSRENEV
          :
          :
myocilin RDQLETQRELETAYSNLRKRSVLEEEKKRLRQENENLARRLESSQVEARLRGQCPQ
olmf2 VRHLEQLRHYNHSAIMLG-----IKKELSLQLLQKDA--AAPATPATGTGSK
          :
          :
myocilin TRDTARAVPQSGREVSTWNLDTLAFQELKSELTEVPASRI
olmf2 AQDTARGK--GKDISKYGVQKSFAD--RLEKFKPKKELLQVEKLRKESGKGSFLQPTA
          :
          :
myocilin KPRALAAQAVIRGFTYKAGQEVTEAVADNTLQGTSWLEQLPPEKVEGSRNSAEPNSAE
olmf2 -----
          :
          :
myocilin -----KESPSGYLRSGEQDTG-
olmf2 QDEAEFRSSERVDLASGTPPTSIPATTTTATTTPTTLLPTEPFGVESSQGREAS
          :
          :
myocilin MRFFCARCCSGPEMPAVQLLLACLWVDVGARTAQLRKANDQSGRCQYTFVAVSAPNESS
olmf3 -----MGSTPLL-----ILFLSWSGFLQGGQKHLVEYME-----RRLALEERLQA
          :
          :
myocilin CPEQSQAMSVIHLNQRDSTQRDLDEATKARLSLESLLHQLTLDQAARQETQEGQLRE
olmf3 CQDQSSR-----HAAELRDFKNK-----
          :
          :
myocilin LGTLRERDQLETQRELETAYSNLRKRSVLEEEKKRLRQENENLARRLESSQVEARL
olmf3 -----MLPLLEVAEKERALEATDTSIGRVDLEREDVYL
          :
          :
myocilin RRG--CQPTRTDARAIVPPGSR-----VSTWNLDTLAFQELKSELTEVPASRIKESP
olmf3 ETQNPALPCVEFDEKV-TGGPGTKGRRNEKYMVDV-----
          :
          :
myocilin SGYLRSGEGDTG
olmf3 -----
    
```

E

```

myocilin MRFFCARCCSGPEMPAVQLLLACLWVDVGARTAQLRKANDQSGRCQYTFVAVSAPNESS
olmf3 -----MGSTPLL-----ILFLSWSGFLQGGQKHLVEYME-----RRLALEERLQA
          :
          :
myocilin CPEQSQAMSVIHLNQRDSTQRDLDEATKARLSLESLLHQLTLDQAARQETQEGQLRE
olmf3 CQDQSSR-----HAAELRDFKNK-----
          :
          :
myocilin LGTLRERDQLETQRELETAYSNLRKRSVLEEEKKRLRQENENLARRLESSQVEARL
olmf3 -----MLPLLEVAEKERALEATDTSIGRVDLEREDVYL
          :
          :
myocilin RRG--CQPTRTDARAIVPPGSR-----VSTWNLDTLAFQELKSELTEVPASRIKESP
olmf3 ETQNPALPCVEFDEKV-TGGPGTKGRRNEKYMVDV-----
          :
          :
myocilin SGYLRSGEGDTG
olmf3 -----
    
```

F

Accession	Species	Length	Score
gi_3065674_h_sapiens	1	MRFFCARCCSGPEMPAVQLLLACLWVDVGARTAQLRKANDQSGRCQYTFVAVSAPNESS	60
gi_15077142_m_musculus	1	-----MPALHLFLACLWVGARTAQFRKANDRSGRQYTFVAVSAPNESS	46
gi_74356501_b_taurus	1	-----MPAVQLLLACLWVGARTAQFRKANDRSGRQYTFVAVSAPNESS	46
gi_3845607_r_norvegicus	1	MP-SCARCCSQPMPALQLLFLACLWVGARTAQFRKANDRSGRQYTFVAVSAPNESS	59
gi_62632725_d_gerio	1	MN-FL-----AVLMSISLMSGQSSANLRANRANGRCQYTFVAVSAPNESS	47

Figure S1. Sequence alignments of CC-containing regions of olfactomedin-containing proteins and myocilins across organisms. Related to Figure 1. (A) Overall sequence alignment of CC-containing regions of human proteins found within different olfactomedin subfamilies. (B) Pairwise sequence alignment of myocilin and olfactomedin-1 (22% identity). (C) Pairwise sequence alignment of myocilin and olfactomedin-4 (22% identity) (D). Pairwise sequence alignment of myocilin and olfactomedin-like-2 (22% identity). (E) Pairwise sequence alignment of myocilin and olfactomedin-like-3 (21% identity). (F) Multiple sequence alignment of N-terminal region of myocilin from human, mouse, bovine, rat, and zebrafish (OLF domain excluded). X above alignment indicates position of mutation investigated in this study. In (A-F), identical residues *; similar residues : and . (G) HMM weblogo representation of conservation across the same region as in (F) across myocilin from 75 species.

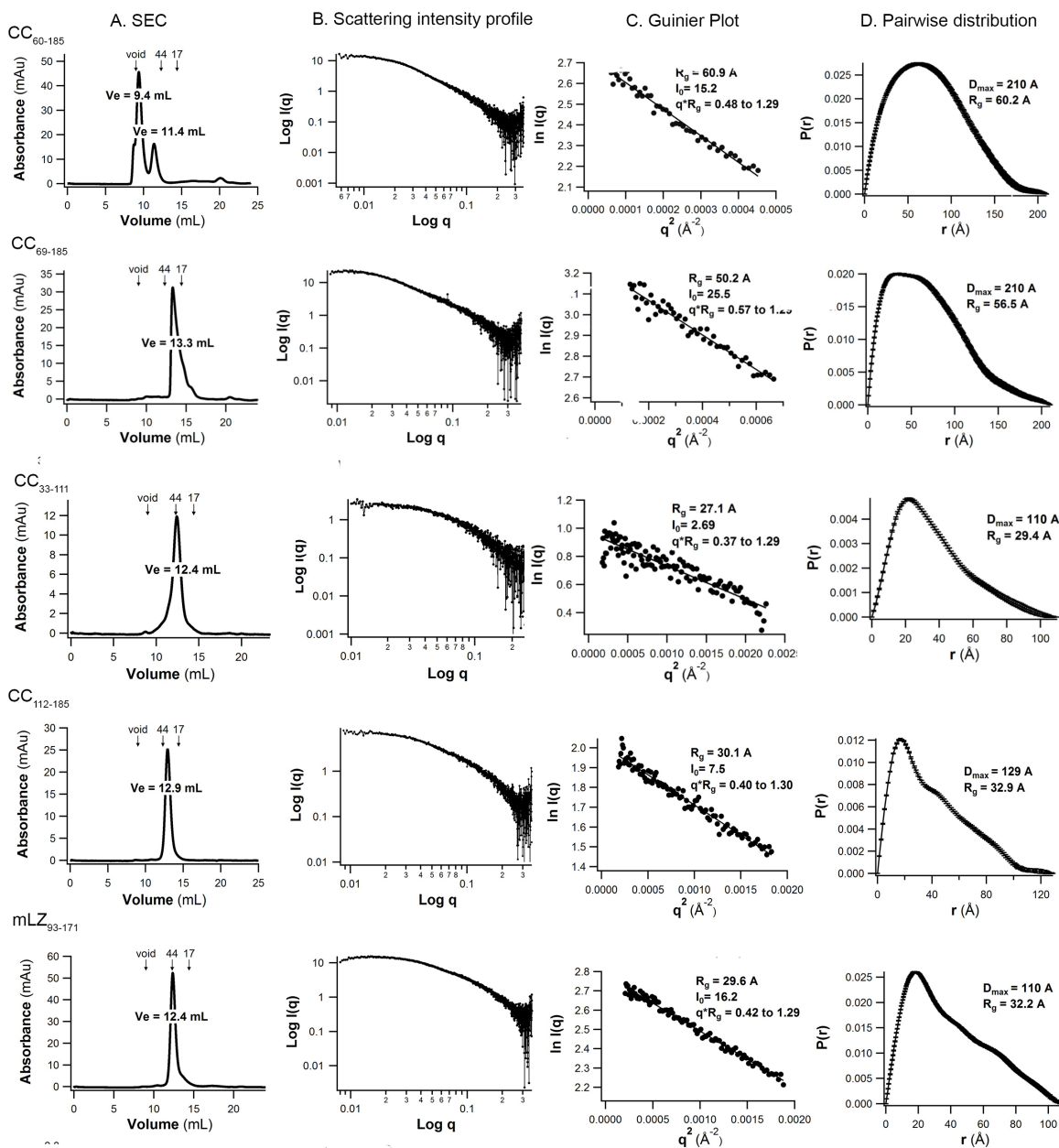


Figure S2. SEC-SAXS analysis of CC₆₀₋₁₈₅, CC₆₉₋₁₈₅, CC₃₃₋₁₁₁, CC₁₁₂₋₁₈₅, mLZ₉₃₋₁₇₁. Related to Figure 3. (A) Superdex-75 GL traces. (B) Scattering intensity profiles. I(q), scattered intensity; q, scattering vector. (C) Guinier plots with calculated radius of gyration (R_g). (D) Pairwise distribution plots with calculated maximum particle size (D_{max}) and R_g.

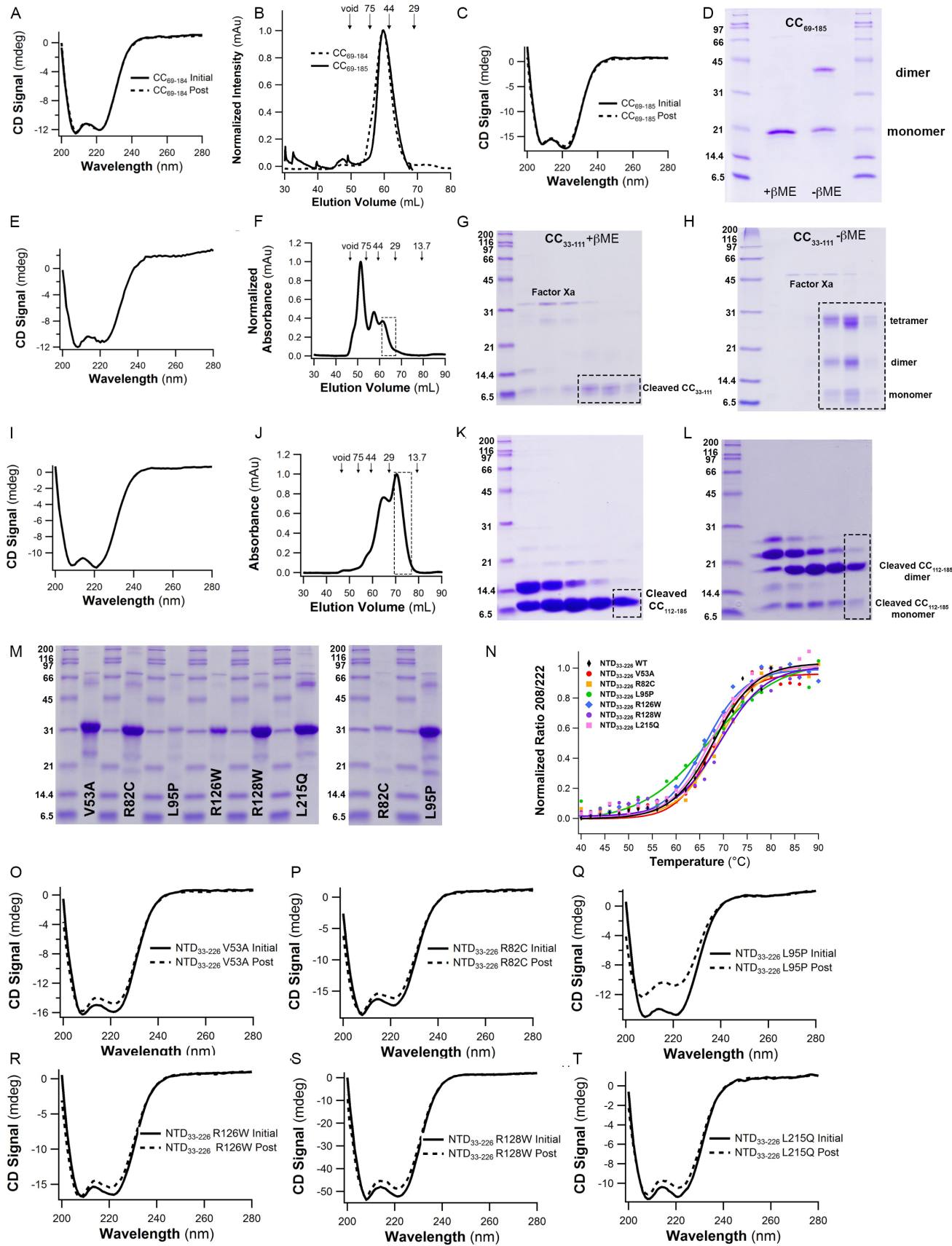


Figure S3. Biochemical characterization of constructs in this study. Related to Figures 3, 4, Table 1 (A-D), CC₆₉₋₁₈₄ and CC₆₉₋₁₈₅, (E-L), CC₃₃₋₁₁₁ and CC₁₁₂₋₁₈₅ (M-T), NTD₃₃₋₂₂₆ variants. (A) Circular dichroism (CD) spectra of CC₆₉₋₁₈₄ reveal α -helical signatures and reversible thermal unfolding (B) Superdex-75 preparative SEC traces of CC₆₉₋₁₈₄ and CC₆₉₋₁₈₅ reveal a total molecular mass of 53 kDa based on a standard calibration curve. Given the calculated mass of 13.7 kDa/monomer the species is a tetramer. (C) CD spectra of CC₆₉₋₁₈₅ reveal α -helical signatures and reversible thermal unfolding. (D) SDS-PAGE analysis of CC₆₉₋₁₈₅ under non-reducing conditions demonstrate predominantly disulfide-dependent dimer species. (E) CD spectrum of CC₃₃₋₁₁₁ reveals α -helical signature. (F) Superdex-75 SEC trace of CC₃₃₋₁₁₁. The first peak contains uncleaved and cleaved CC₃₃₋₁₁₁, the second peak contains mostly Factor Xa protease, and the third peak (indicated by a dashed box) contains cleaved CC₃₃₋₁₁₁ used in subsequent experiments. The molecular weight of the peak within the dashed box is 40 kDa based on a standard calibration curve, consistent with a 4 or 5-mer of an \sim 8.8 kDa CC₃₃₋₁₁₁ monomer. (G-H) SDS-PAGE analysis of CC₃₃₋₁₁₁ fractions from Superdex-75 trace with (G) and without (H) β ME. Dashed box indicates fractions concentrated for SEC-SAXS. (I) CD spectrum of CC₁₁₂₋₁₈₄ reveals an α -helical signature. (J) Superdex-75 SEC trace of CC₁₁₂₋₁₈₅. The first peak contains uncleaved and cleaved CC₁₁₂₋₁₈₅, and the second peak (indicated by dashed box) contains mostly cleaved CC₁₁₂₋₁₈₅ used in subsequent experiments. The molecular mass of the species is 23 kDa based on a standard calibration curve and thus consistent with a 2- or 3-mer of an \sim 9 kDa CC₁₁₂₋₁₈₅ species. (K,L) SDS-PAGE analysis with (K) and without (L) β ME of CC₁₁₂₋₁₈₅-containing fractions from Superdex-75 trace shown in (J). Dashed box indicates fraction selected for SEC-SAXS. (M) (M) SDS-PAGE analysis (left) of all six purified variants (\sim 30 kDa monomer) corresponding to the fraction (left asterisk) on the chromatograms in Figure 4F. SDS-PAGE analysis (right) of R82C and L95P variants with disrupted tetramer arrangement corresponding to the fraction (right asterisk) on the chromatogram in Figure 4F. (N) CD thermal melts show six variants are indistinguishable from wild-type stability. (O-T) Comparison of CD spectra before (initial) and after (post) thermal melt shows a high level of reversibility for six variants studied.

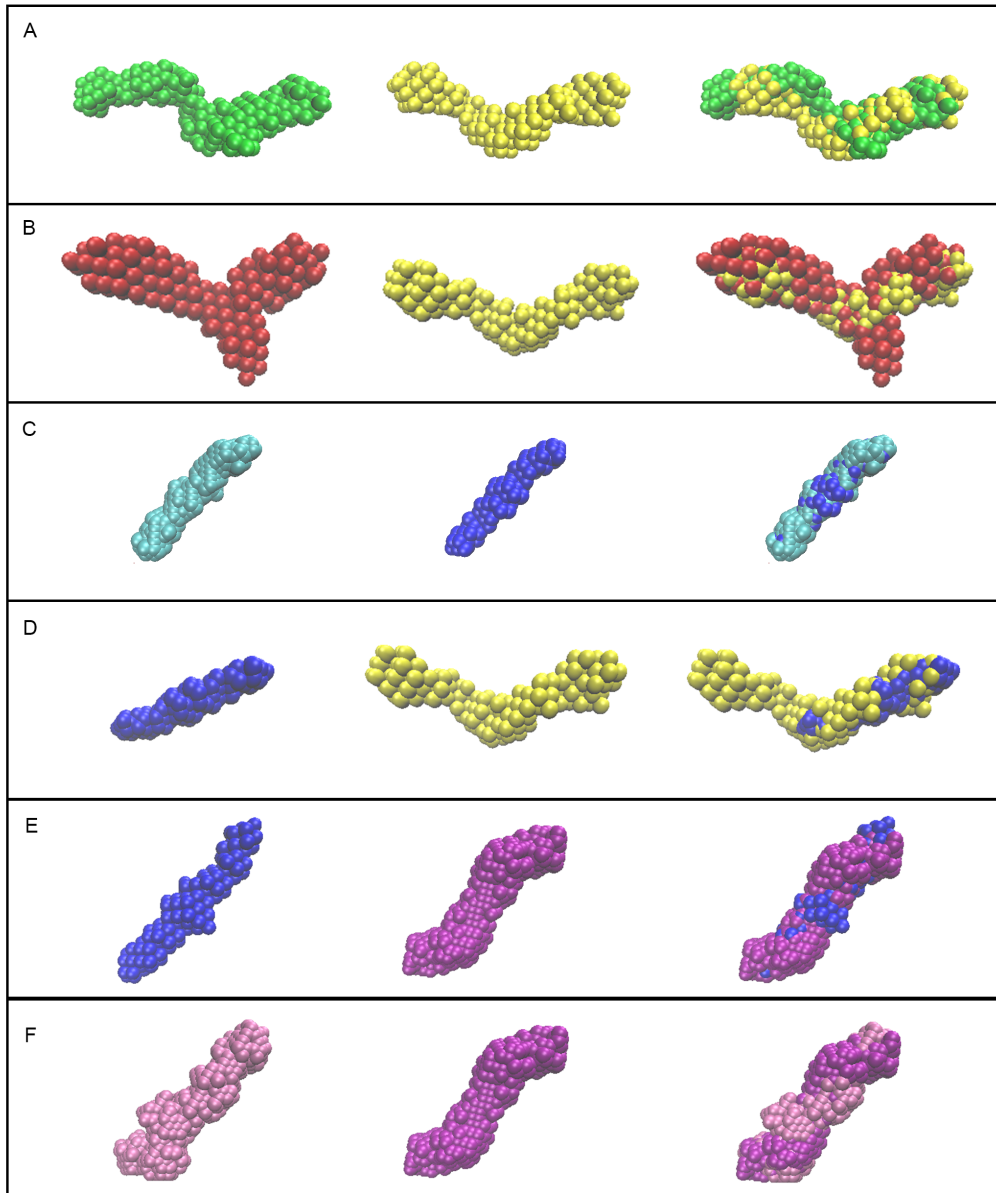


Figure S4. Pairwise comparisons of *ab initio* SEC-SAXS models. Related to Figure 3, 4 and Table 2. (A) Molecular envelopes of CC₆₉₋₁₈₅ with no (P1) symmetry (left), P2 symmetry (middle) and superpositions with SUPCOMB (right). (B) Molecular envelopes of CC₆₀₋₁₈₅ (left), CC₆₉₋₁₈₅ (middle) and superposition with SUPCOMB (right). (C) Molecular envelopes of CC₁₁₂₋₁₈₅ with no (P1) symmetry (left), P2 symmetry (middle) and superposition with SUPCOMB (right). (D) Molecular envelopes of CC₁₁₂₋₁₈₅ with P2 symmetry (left), CC₆₉₋₁₈₅ with P2 symmetry (middle) and superposition by manual manipulation in PyMOL (right). (E) Molecular envelopes of CC₁₁₂₋₁₈₅ with P2 symmetry (left), mLZ₉₃₋₁₇₁ (see F) with P2 symmetry (middle) and superposition with SUPCOMB (right). (F) Molecular envelopes of mLZ₉₃₋₁₇₁ with no (P1) symmetry (left), P2 symmetry (middle), and superposition with SUPCOMB (right).

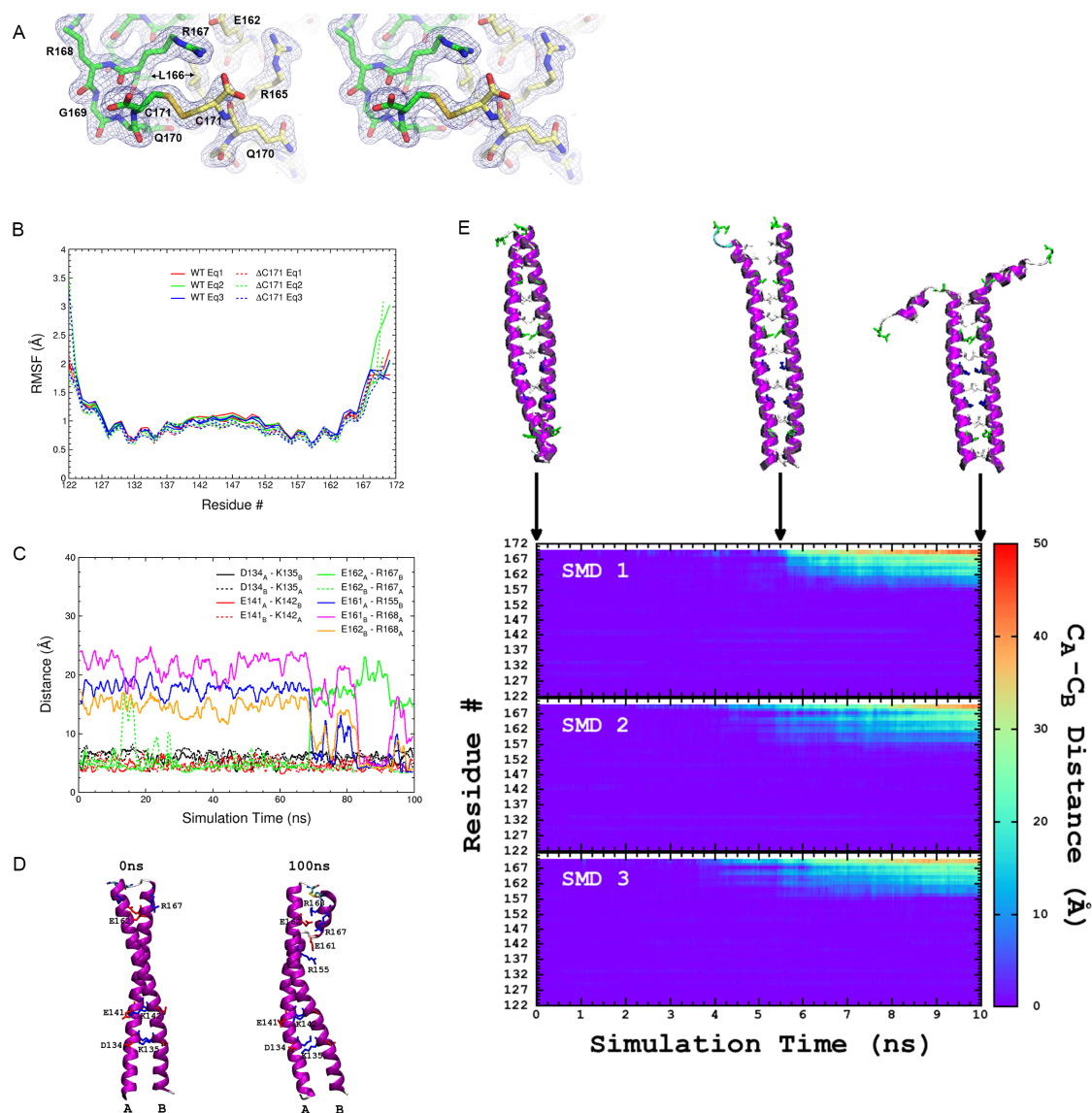
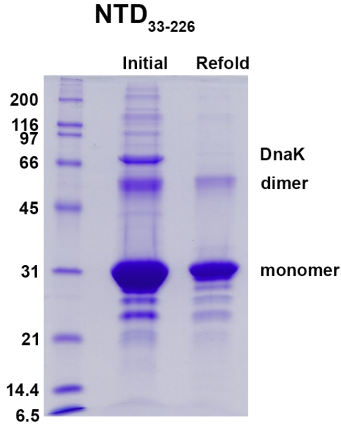


Figure S5. Molecular dynamics simulations of m LZ₁₂₂₋₁₇₁. Related to Figure 4. (A) Stereoview of Figure 4B. (B) Root mean squared fluctuation (RMSF) values for individual m LZ₁₂₂₋₁₇₁ residues for simulations conducted at 310 K. (C)/(D) Salt bridges formed between individual coils A and B during 100-ns simulation of m LZ₁₂₂₋₁₇₁ at 400 K. For (D), backbone structures are shown for 0 ns (left) and 100 ns (right) of simulation time with salt-bridge-forming residues shown explicitly. Salt bridges were identified by ≤ 3.2 Å distance between side chain oxygen and nitrogen atoms. (E) Distances between backbone carbonyl carbon atoms of chains A and B for each residue of m LZ₁₂₂₋₁₇₁ during steered molecular dynamics (SMD) simulations at 310 K. Backbone structures from the SMD1 trajectory are shown above the graph with side chains of heptad positions 'a' and 'd' shown explicitly. For all protein structures shown: Individual protein chains are labeled A and B. Protein backbone is shown in cartoon representation colored by secondary structure: (magenta) α -helix and (white) random coil. Explicit side chains are shown in licorice representation colored by residue type: (blue) positively charged, (red) negatively charged, (green) polar, and (white) hydrophobic.

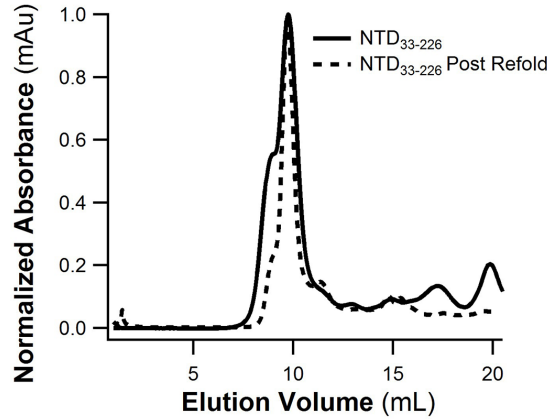
A

Accession	Description	Score	Coverage	# Proteins	# Unique Peptides	# Peptides	# PSMs	# AAs	MW [kDa]	calc. pI
157159481	molecular chaperone DnaK [Escherichia coli HS]	510.06	54.23	566	11	28	100	638	69.1	4.97
16763402	molecular chaperone DnaK [Salmonella enterica subsp. enterica serovar Typhimurium str. LT2]	304.82	36.52	552	1	18	62	638	69.2	4.97
16130190	fused UDP-L-Ara4N formyltransferase/UDP-GlcA C-4'-decarboxylase [Escherichia coli str. K-12 substr. MG1655]	76.90	19.70	22	9	9	16	660	74.2	6.87
160961491	keratin, type II cytoskeletal 1 [Pan troglodytes]	76.64	20.88	46	6	8	18	637	65.4	7.81
157376541	molecular chaperone DnaK [Shewanella sediminis HAW-EB3]	54.60	6.09	519	1	3	11	640	69.0	4.73
114667511	PREDICTED: keratin, type I cytoskeletal 10 isoform 3 [Pan troglodytes]	51.02	17.85	15	6	6	11	577	58.2	5.16
47132620	keratin, type II cytoskeletal 2 epidermal [Homo sapiens]	50.72	14.24	30	2	4	11	639	65.4	8.00

B



C



D

Annotated Sequence	# PSMs	Annotated Sequence	#PSMs
[G].RVGTQEAQEGQLGQALR.[R]	11	[R].DQLETQTRDLEAAYNNLLR.[D]	4
[G].RVGTQEAQEGQLGQALR.[R]	10	[R].DQLETQTRDLEAAYNNLLR.[D]	3
[R].VTGTQEAQEGQLGQALR.[R]	159	[R].DLEAAYNNLLR.[D]	164
[R].VTGTQEAQEGQLGQALRR.[E]	25	[R].DLEAAYNNLLRDK.[S]	7
[R].VTGTQEAQEGQLGQALRR.[D]	8	[R].DLEAAYNNLLRDKSALEEEK.[R]	3
[R].VTGTQEAQEGQLGQALRR.[R]	1	[R].DLEAAYNNLLRDKSALEEEK.[Q]	2
[R].VTGTQEAQEGQLGQ.[L]	1	[R].DLEAAYNNLLR.[D]	174
[R].VTGTQEAQEGQLGQALR.[R]	152	[R].DLEAAYNNLLRDK.[S]	9
[R].VTGTQEAQEGQLGQALRR.[E]	25	[R].DLEAAYNNLLRDKSALEEEK.[R]	3
[R].VTGTQEAQEGQLGQALRR.[D]	6	[R].DLEAAYNNLLRDKSALEEEK.[Q]	2
[R].VTGTQEAQEGQLGQALRR.[R]	1	[D].LEAAYNNLLR.[D]	5
[V].TGTQEAQEGQLGQALR.[R]	7	[D].LEAAYNNLLR.[D]	5
[V].TGTQEAQEGQLGQALR.[R]	7	[L].EAAYNNLLR.[D]	9
[T].GTQEAQEGQLGQALR.[R]	12	[L].EAAYNNLLR.[D]	9
[T].GTQEAQEGQLGQALR.[R]	13	[R].DKSALEEEKRQLEQENEDLAR.[R]	12
[G].TQEAQEGQLGQALR.[R]	13	[R].DKSALEEEKRQLEQENEDLAR.[R]	15
[G].TQEAQEGQLGQALR.[R]	14	[K].SALEEEKRQLEQENEDLAR.[R]	6
[T].QEAQEGQLGQALR.[R]	2	[K].SALEEEKRQLEQENEDLAR.[R]	8
[T].QEAQEGQLGQALR.[R]	2	[K].SALEEEKR.[Q]	1
[Q].EAQEGQLGQALR.[R]	13	[K].RQLEQENEDLAR.[R]	15
[Q].EAQEGQLGQALR.[R]	14	[K].RQLEQENEDLAR.[R]	15
[A].QEGQLGQALR.[R]	4	[R].QLEQENEDLAR.[R]	93
[A].QEGQLGQALR.[R]	4	[R].QLEQENEDLAR.[R]	91
[Q].EGLGQALR.[R]	10	[Q].LEQENEDLAR.[R]	1
[Q].EGLGQALR.[R]	13	[Q].LEQENEDLAR.[R]	1
[E].GLGQALR.[R]	3	[L].EQENEDLAR.[R]	6
[E].GLGQALR.[R]	3	[L].EQENEDLAR.[R]	5
[L].QGQALR.[R]	1	[E].QENEDLAR.[R]	4
[R].ERDQLETQTR.[D]	2	[E].QENEDLAR.[R]	3
[R].ERDQLETQTRDLEAAYNNLLR.[D]	1	[R].RLESSEEVTR.[L]	11
[R].ERDQLETQTRDLEAAYNNLLR.[D]	4	[R].RLESSEEVTR.[L]	9
[R].ERDQLETQTR.[D]	1	[R].RLESSEEVTR.[L]	1

E

Primers used in this study.

Construct	Primers for pet 30 Xa/LIC subcloning
NTD ₃₃₋₂₂₆	5'GGTATTGAGGGTCGCCACCCGCTCAGCTCGCAAGGCGAATG
CC ₆₀₋₁₈₅	5'GGTATTGAGGGTCGCTCTGCCGGAGCAG
CC ₆₉₋₁₈₄	5'GGTATTGAGGGTCGCTCGGTGATTACAACTCGACCGTG
CC ₁₁₂₋₁₈₄	5'GGTATTGAGGGTCGCCAAACTCAGGAAGGCTTGCAACGTG
Construct	Primers for Site Directed Mutagenesis
CC ₆₉₋₁₈₅ ^a	5'CTTGCCTCGTGGTCAGTGTAAAGGCTCTAACTCTC
CC ₁₁₂₋₁₈₅ ^a	5'CTTGCCTCGTGGTCAGTGTAAAGGCTCTAACTCTC
CC ₃₃₋₁₁₁ ^b	5'CAAGCTCGCGCTCCGAGTAACTCAGGAAGGCTTGC
NTD ₃₃₋₂₂₆ C ₄₇ S ^c	5'GACCAGTCGGGCCGTAGTCAGTATACCTTTTC
NTD ₃₃₋₂₂₆ C ₉₁ S ^c	5'GCCCTAATGAGTCTCTAGCCCGGAGC
NTD ₃₃₋₂₂₆ C ₁₈₅ S ^c	5'GCGTCGTGGTCAGAGTCGCAAAACGCG
NTD ₃₃₋₂₂₆ V53A	5'TCAGTATACCTTTTCGCGGCAAGCCCTAATGAGT
NTD ₃₃₋₂₂₆ R82C	5'GATTCCAGCACGACGAGTGTCTGGACCTGGAAG
NTD ₃₃₋₂₂₆ L95P	5'GCCTGAGCAGCCCGGAGAGCCTGCT
NTD ₃₃₋₂₂₆ R126W	5'GTGAATTTGGGTACGCTGCGTGGGAACGTGACCAGCTGGAAC
NTD ₃₃₋₂₂₆ R128W	5'CGCTGCGTCGTGAATGGGACCAGCTGGAAC
NTD ₃₃₋₂₂₆ L215Q	5'CTGGCGTTTCAAGAGCAGAAAAGCGAGCTGACC

^aCC₆₉₋₁₈₅ and CC₁₁₂₋₁₈₅ produced by insertion of Cys₁₈₅ to CC₆₉₋₁₈₄ and CC₁₁₂₋₁₈₄, respectively.

^bCC₃₃₋₁₁₁ produced by mutation to stop codon after Gln111 in NTD₃₃₋₂₂₆ construct.

^cNTD₃₃₋₂₂₆ double and triple cysteine to serine variants produced by multiple rounds of mutagenesis.

Figure S6. Additional supplemental material accompanying STAR methods. (A) Identification of DnaK by mass spectrometry as main contaminant (~60 kDa) visible after purification on SDS-PAGE. (B) SDS-PAGE analysis of NTD₃₃₋₂₂₆ before and after unfolding/refolding procedure. (C) Comparison of Superdex 200 GL elution profile of two samples in (B). (D) Identification of mLZ₉₃₋₁₇₁ as predominant product after cleavage of mLZ₅₅₋₁₇₁ by Factor Xa. (E) Primers used in this study.