Supplemental Data for Hill et al.

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.

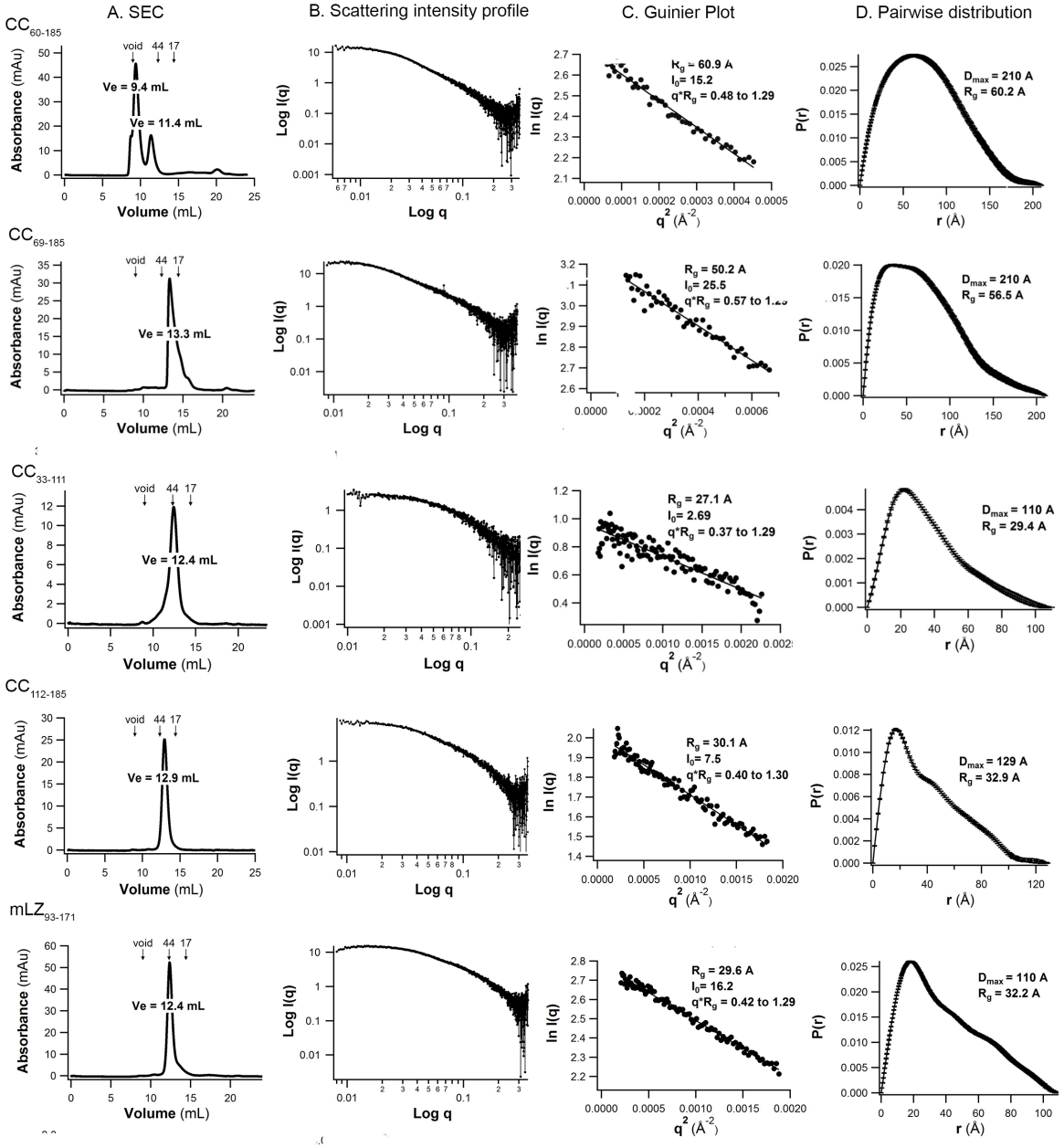


Figure S2. SEC-SAXS analysis of $CC_{60\text{-}185}$, $CC_{69\text{-}185}$, $CC_{33\text{-}111}$, $CC_{112\text{-}185}$, $mLZ_{93\text{-}171}$. Related to Figure 3. (A) Superdex-75 GL traces. (B) Scattering intensity profiles. I(q), scattered intensity; q, scattering vector. (C) Guinier plots with calculated radius of gyration (Rg). (D) Pairwise distribution plots with calculated maximum particle size (Dmax) and Rg.

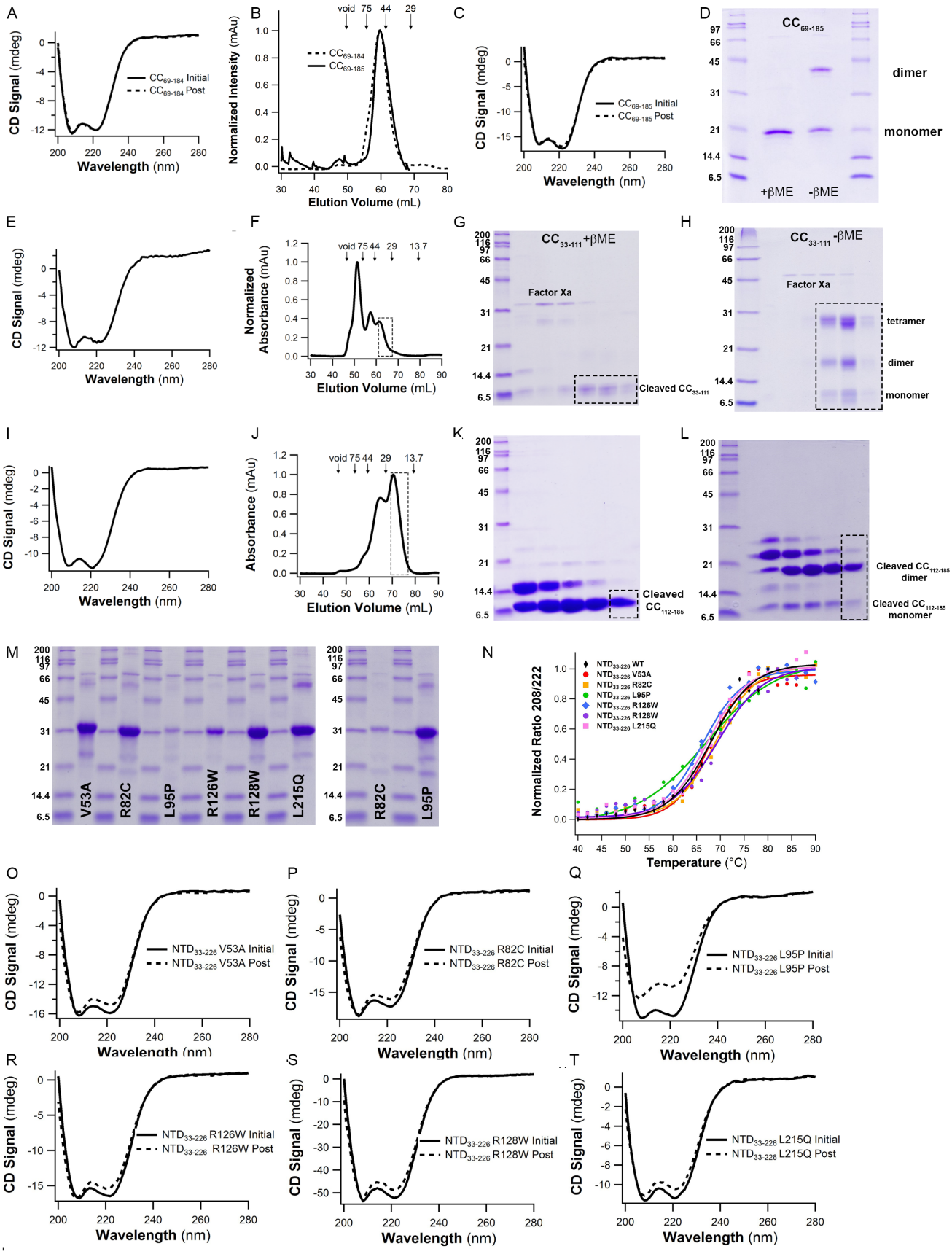

A
CD Signal (mdeg)
Wavelength (nm)
CC69-184 Initial
CC69-184 Post
B
Normalized Intensity (mAu)
void 75 44 29
CC69-184
CC69-185
Elution Volume (mL)
C
CD Signal (mdeg)
Wavelength (nm)
CC69-185 Initial
CC69-185 Post
D
CC69-185
dimer
monomer
+βME -βME
E
CD Signal (mdeg)
Wavelength (nm)
F
Normalized Absorbance (mAu)
void 75 44 29 13.7
Elution Volume (mL)
G
CC33-111 +βME
Factor Xa
Cleaved CC33-111
H
CC33-111 -βME
Factor Xa
tetramer
dimer
monomer
I
CD Signal (mdeg)
Wavelength (nm)
J
Absorbance (mAu)
void 75 44 29 13.7
Elution Volume (mL)
K
Cleaved CC112-185
L
Cleaved CC112-185 dimer
Cleaved CC112-185 monomer
M
V53A
R82C
L95P
R126W
R128W
L215Q
R82C
L95P
N
Normalized Ratio 208/222
Temperature (°C)
NTD33-226 WT
NTD33-226 V53A
NTD33-226 R82C
NTD33-226 L95P
NTD33-226 R126W
NTD33-226 R128W
NTD33-226 L215Q
O
CD Signal (mdeg)
Wavelength (nm)
NTD33-226 V53A Initial
NTD33-226 V53A Post
P
CD Signal (mdeg)
Wavelength (nm)
NTD33-226 R82C Initial
NTD33-226 R82C Post
Q
CD Signal (mdeg)
Wavelength (nm)
NTD33-226 L95P Initial
NTD33-226 L95P Post
R
CD Signal (mdeg)
Wavelength (nm)
NTD33-226 R126W Initial
NTD33-226 R126W Post
S
CD Signal (mdeg)
Wavelength (nm)
NTD33-226 R128W Initial
NTD33-226 R128W Post
T
CD Signal (mdeg)
Wavelength (nm)
NTD33-226 L215Q Initial
NTD33-226 L215Q Post

Figure S3. Biochemical characterization of constructs in this study. Related to Figures 3, 4, Table 1 (A-D), CC_{69-184} and CC_{69-185}, (E-L), CC_{33-111} and $CC_{112-185}$ (M-T), NTD_{33-226} variants. (A) Circular dichroism (CD) spectra of CC_{69-184} reveal α-helical signatures and reversible thermal unfolding (B) Superdex-75 preparative SEC traces of CC_{69-184} and CC_{69-185} 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 CC_{69-185} reveal α-helical signatures and reversible thermal unfolding. (D) SDS-PAGE analysis of CC_{69-185} under non-reducing conditions demonstrate predominantly disulfide-dependent dimer species. (E) CD spectrum of CC_{33-111} reveals α-helical signature. (F) Superdex-75 SEC trace of CC_{33-111}. The first peak contains uncleaved and cleaved CC_{33-111}, the second peak contains mostly Factor Xa protease, and the third peak (indicated by a dashed box) contains cleaved CC_{33-111} 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 ~8.8 kDa CC_{33-111} monomer. (G-H) SDS-PAGE analysis of CC_{33-111} fractions from Superdex- 75 trace with (G) and without (H) βME. Dashed box indicates fractions concentrated for SEC-SAXS. (I) CD spectrum of $CC_{112-184}$ reveals an α-helical signature. (J) Superdex-75 SEC trace of $CC_{112-185}$. The first peak contains uncleaved and cleaved $CC_{112-185}$, and the second peak (indicated by dashed box) contains mostly cleaved $CC_{112-185}$ 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 ~9 kDa $CC_{112-185}$ species. (K,L) SDS-PAGE analysis with (K) and without (L) βME of $CC_{112-185}$-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 (~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.

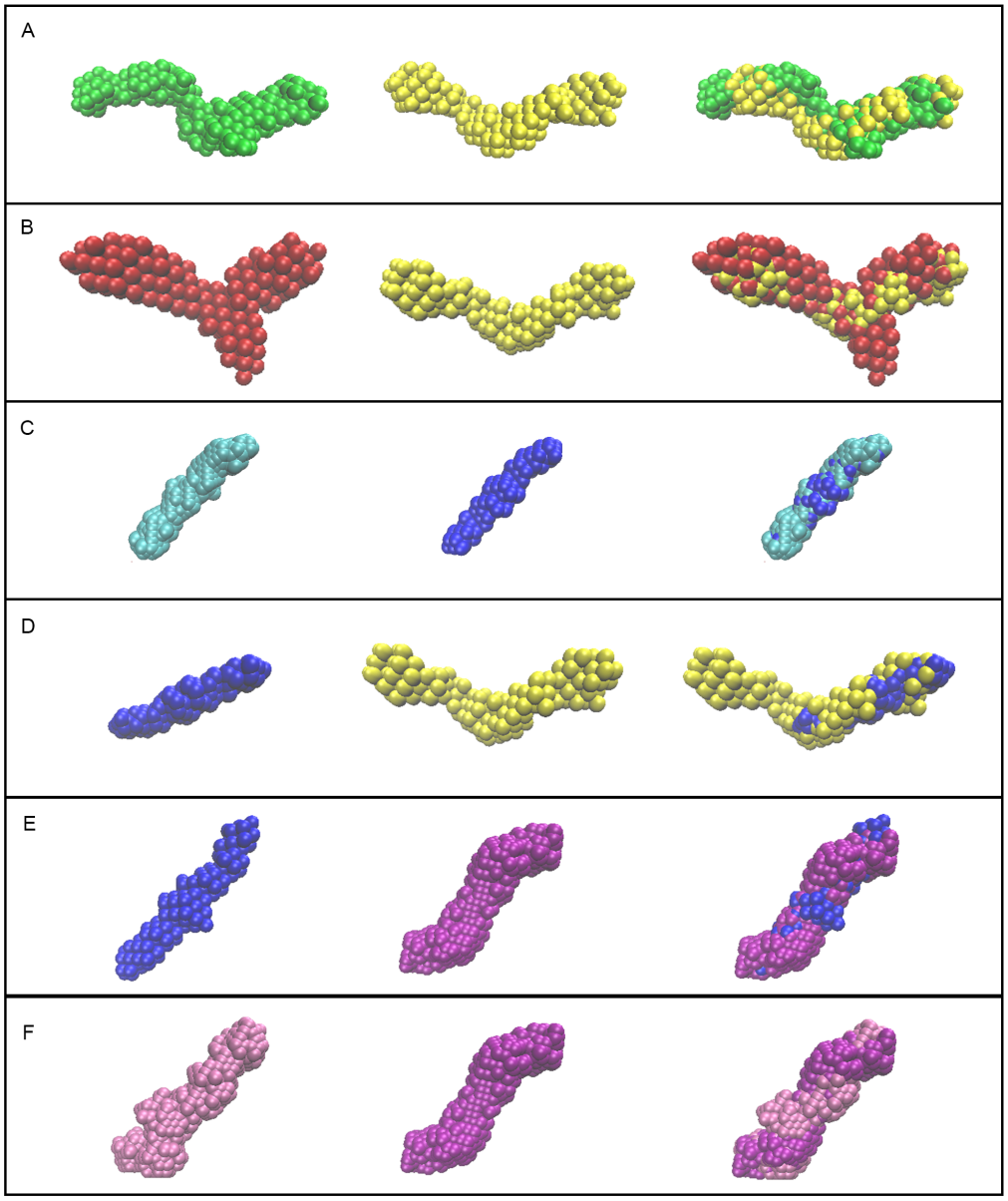


Figure S4. Pairwise comparisons of *ab initio* SEC-SAXS models. Related to Figure 3, 4 and Table 2. (A) Molecular envelopes of $CC_{69\text{-}185}$ with no (P1) symmetry (left), P2 symmetry (middle) and superpositions with SUPCOMB (right). (B) Molecular envelopes of $CC_{60\text{-}185}$ (left), $CC_{69\text{-}185}$ (middle) and superposition with SUPCOMB (right). (C) Molecular envelopes of $CC_{112\text{-}185}$ with no (P1) symmetry (left), P2 symmetry (middle) and superposition with SUPCOMB (right). (D) Molecular envelopes of $CC_{112\text{-}185}$ with P2 symmetry (left), $CC_{69\text{-}185}$ with P2 symmetry (middle) and superposition by manual manipulation in PyMOL (right). (E) Molecular envelopes of $CC_{112\text{-}185}$ with P2 symmetry (left), $mLZ_{93\text{-}171}$ (see F) with P2 symmetry (middle) and superposition with SUPCOMB (right). (F) Molecular envelopes of $mLZ_{93\text{-}171}$ with no (P1) symmetry (left), P2 symmetry (middle), and superposition with SUPCOMB (right).

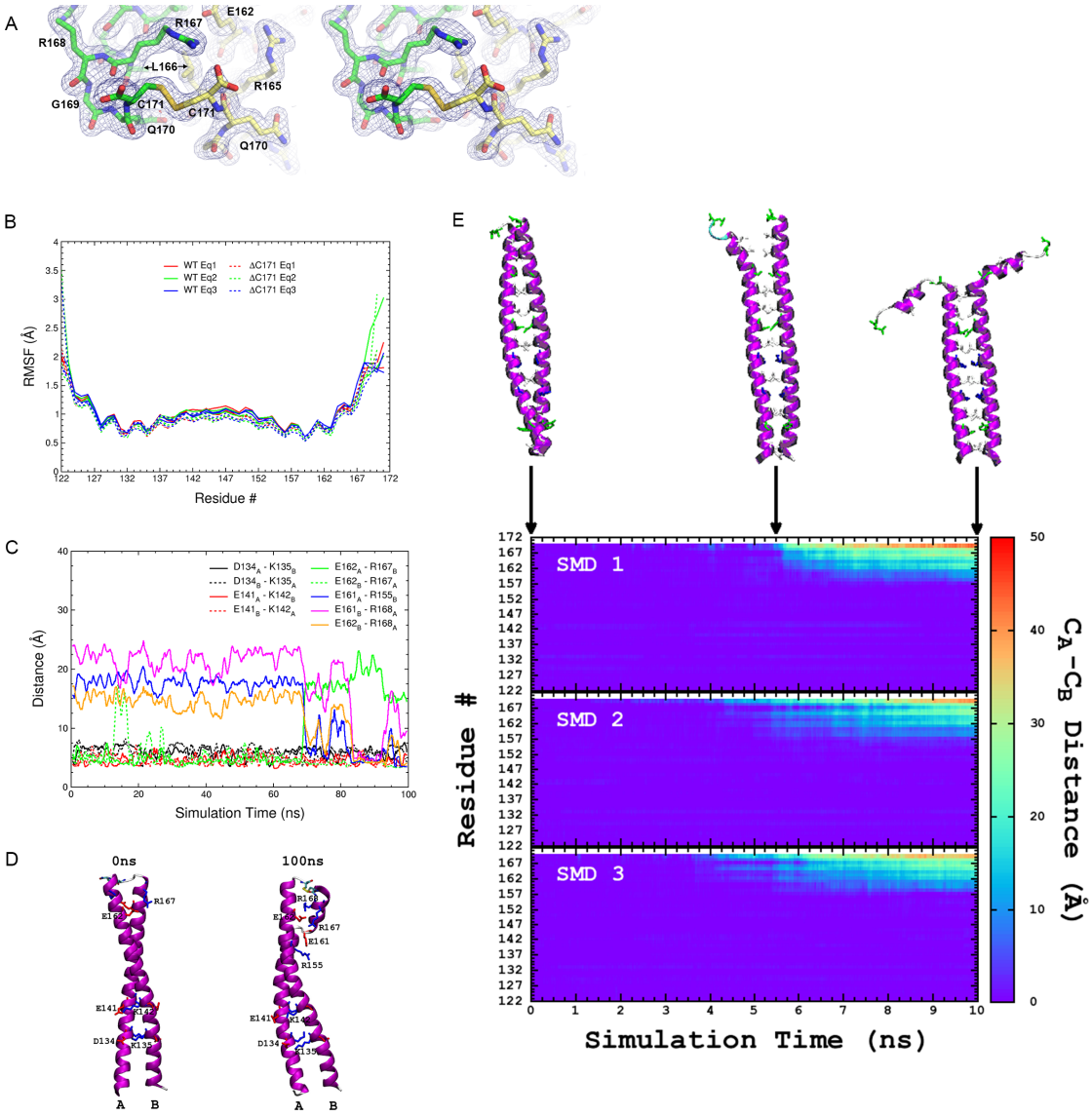
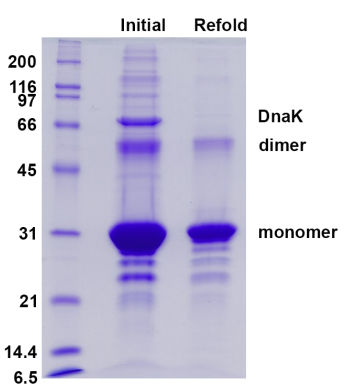


Figure S5. Molecular dynamics simulations of mLZ$_{122-171}$. Related to Figure 4. (A) Stereoview of Figure 4B. (B) Root mean squared fluctuation (RMSF) values for individual mLZ$_{122-171}$ residues for simulations conducted at 310 K. (C)/(D) Salt bridges formed between individual coils A and B during 100-ns simulation of mLZ$_{122-171}$ 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 mLZ$_{122-171}$ 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.

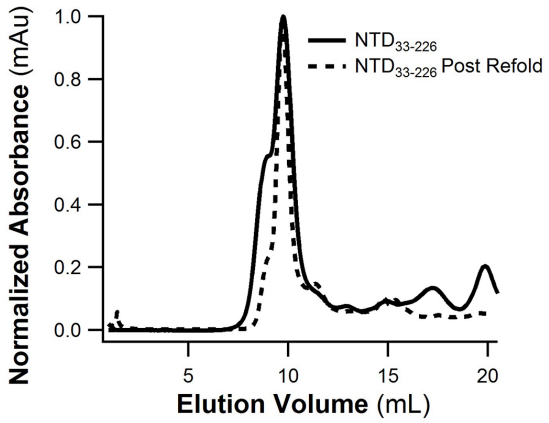
A

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
[G].RVTGTQEAQEGLQGQLGALR.[R]	11
[G].RVTGTQEAQEGLQGQLGALR.[R]	10
[R].VTGTQEAQEGLQGQLGALR.[R]	159
[R].VTGTQEAQEGLQGQLGALRR.[E]	25
[R].VTGTQEAQEGLQGQLGALRRER.[D]	8
[R].VTGTQEAQEGLQGQLGALRRE.[R]	1
[R].VTGTQEAQEGLQGQ.[L]	1
[R].VTGTQEAQEGLQGQLGALR.[R]	152
[R].VTGTQEAQEGLQGQLGALRR.[E]	25
[R].VTGTQEAQEGLQGQLGALRRER.[D]	6
[R].VTGTQEAQEGLQGQLGALRRE.[R]	1
[V].TGTQEAQEGLQGQLGALR.[R]	7
[V].TGTQEAQEGLQGQLGALR.[R]	7
[T].GTQEAQEGLQGQLGALR.[R]	12
[T].GTQEAQEGLQGQLGALR.[R]	13
[G].TQEAQEGLQGQLGALR.[R]	13
[G].TQEAQEGLQGQLGALR.[R]	14
[T].QEAQEGLQGQLGALR.[R]	2
[T].QEAQEGLQGQLGALR.[R]	2
[Q].EAQEGLQGQLGALR.[R]	13
[Q].EAQEGLQGQLGALR.[R]	14
[A].QEGLQGQLGALR.[R]	4
[A].QEGLQGQLGALR.[R]	4
[Q].EGLQGQLGALR.[R]	10
[Q].EGLQGQLGALR.[R]	13
[E].GLQGQLGALR.[R]	3
[E].GLQGQLGALR.[R]	3
[L].QGQLGALR.[R]	1
[R].ERDQLETQTR.[D]	2
[R].ERDQLETQTRDLEAAYNNLLR.[D]	1
[R].ERDQLETQTRDLEAAYNNLLR.[D]	4
[R].ERDQLETQTR.[D]	1

Annotated Sequence	#PSMs
[R].DQLETQTRDLEAAYNNLLR.[D]	4
[R].DQLETQTRDLEAAYNNLLR.[D]	3
[R].DLEAAYNNLLR.[D]	164
[R].DLEAAYNNLLRDK.[S]	7
[R].DLEAAYNNLLRDKSALEEEK.[R]	3
[R].DLEAAYNNLLRDKSALEEEKR.[Q]	2
[R].DLEAAYNNLLR.[D]	174
[R].DLEAAYNNLLRDK.[S]	9
[R].DLEAAYNNLLRDKSALEEEK.[R]	3
[R].DLEAAYNNLLRDKSALEEEKR.[Q]	2
[D].LEAAYNNLLR.[D]	5
[D].LEAAYNNLLR.[D]	5
[L].EAAYNNLLR.[D]	9
[L].EAAYNNLLR.[D]	9
[R].DKSALEEEKRQLEQENEDLAR.[R]	12
[R].DKSALEEEKRQLEQENEDLAR.[R]	15
[K].SALEEEKRQLEQENEDLAR.[R]	6
[K].SALEEEKRQLEQENEDLAR.[R]	8
[K].SALEEEKR.[Q]	1
[K].RQLEQENEDLAR.[R]	15
[K].RQLEQENEDLAR.[R]	15
[R].QLEQENEDLAR.[R]	93
[R].QLEQENEDLAR.[R]	91
[Q].LEQENEDLAR.[R]	1
[Q].LEQENEDLAR.[R]	1
[L].EQENEDLAR.[R]	6
[L].EQENEDLAR.[R]	5
[E].QENEDLAR.[R]	4
[E].QENEDLAR.[R]	3
[R].RLESSSEEVTR.[L]	11
[R].RLESSSEEVTR.[L]	9
[R].LESSSEEVTR.[L]	1

E

Primers used in this study.

Construct	Primers for pet 30 Xa/LIC subcloning
NTD_{33-226}	5'GGTATTGAGGGTCGCCGCACCGCTCAGCTGCGCAAGGCGAATG
CC_{60-185}	5' GGTATTGAGGGTCGCTCTTGCCCGGAGCAG
CC_{69-184}	5'GGTATTGAGGGTCGCTCGGTGATTCACAATCTGCAGCGTG
$CC_{112-184}$	5'GGTATTGAGGGTCGCGAAACTCAGGAAGGCTTGCAACGTG
Construct	**Primers for Site Directed Mutagenesis**
CC_{69-185} [a]	5'CTTGCGTCGTGGTCAGTGTTAAGGCTCTAACTCTC
$CC_{112-185}$ [a]	5'CTTGCGTCGTGGTCAGTGTTAAGGCTCTAACTCTC
CC_{33-111} [b]	5'CAAGCTGCGCGTCCGCAGTAAACTCAGGAAGGCTTGC
NTD_{33-226} $C_{47}S$ [c]	5'GACCAGTCGGGCCGTAGTCAGTATACCTTTTC
NTD_{33-226} $C_{61}S$ [c]	5'GCCCTAATGAGTCCTCTAGCCCGGAGC
NTD_{33-226} $C_{185}S$ [c]	5'GCGTCGTGGTCAGAGTCCGCAAACGCG
NTD_{33-226} V53A	5'TCAGTATACCTTTTCCGCGGCAAGCCCTAATGAGT
NTD_{33-226} R82C	5'GATTCCAGCACGCAGTGTCTGGACCTGGAAG
NTD_{33-226} L95P	5'GCCTGAGCAGCCCGGAGAGCCTGCT
NTD_{33-226} R126W	5'GTGAATTGGGTACGCTGCGTTGGGAACGTGACCAGCTGGAAAC
NTD_{33-226} R128W	5'CGCTGCGTCGTGAATGGGACCAGCTGGAAAC
NTD_{33-226} L215Q	5'CTGGCGTTTCAAGAGCAGAAAAGCGAGCTGACC

[a] CC_{69-185} and $CC_{112-185}$ produced by insertion of Cys_{185} to CC_{69-184} and $CC_{112-184}$, respectively.
[b] CC_{33-111} produced by mutation to stop codon after Gln111 in NTD_{33-226} construct.
[c] NTD_{33-226} 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 NTD_{33-226} before and after unfolding/refolding procedure. (C) Comparison of Superdex 200 GL elution profile of two samples in (B). (D) Identification of mLZ_{93-171} as predominant product after cleavage of mLZ_{55-171} by Factor Xa. (E) Primers used in this study.