Article:
Bruce, AJE, Paci, E and Brockwell, DJ orcid.org/0000-0002-0802-5937 (2023) Collagen-like motifs of SasG: a novel fold for protein mechanical strength. Journal of Molecular Biology, 435 (6). 167980. ISSN 0022-2836

https://doi.org/10.1016/j.jmb.2023.167980

Supporting Information

Collagen-like motifs of SasG: a novel fold for protein mechanical strength.

Alexander J. E. Bruce, Emanuele Paci‡ and David J. Brockwell*

Astbury Centre for Structural and Molecular Biology and School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, LS2 9JT, UK

‡E.P. present address: University of Bologna, Department of Physics, Bologna, Italy

*Corresponding author:
David J. Brockwell
Tel: +44 (0) 113 343 7821
Email: D.J.Brockwell@leeds.ac.uk

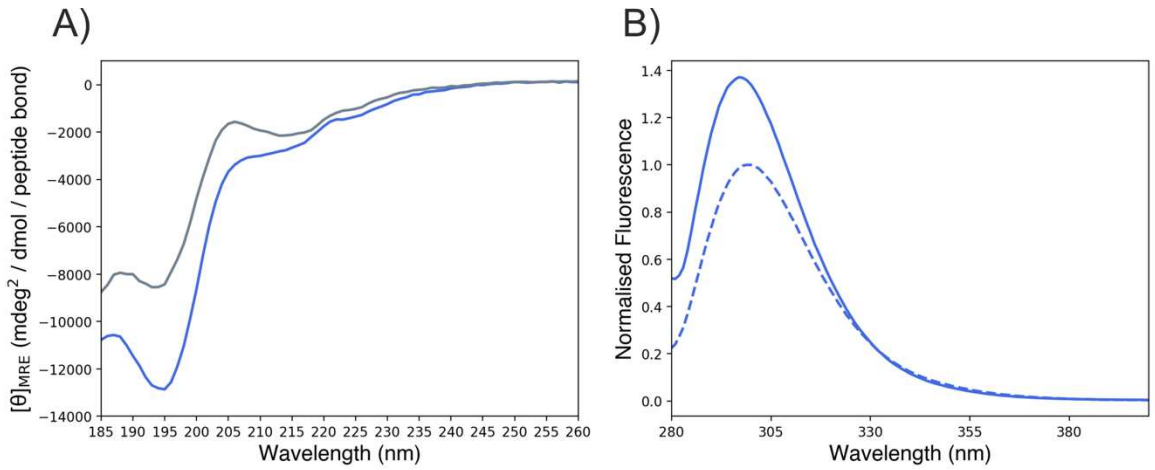

Figure S1. Spectroscopic analysis of pWT $(E\text{-}G5^2)_5$. **A)** Far-UV CD spectra of SasG (grey) and pWT $(E\text{-}G5^2)_5$ (blue). pWT $(E\text{-}G5^2)_5$ shares a similar secondary structure with SasG. **B)** Fluorescence emission spectra of pWT $(E\text{-}G5^2)_5$ in the absence (solid) and presence (dashed) of 8 M urea.

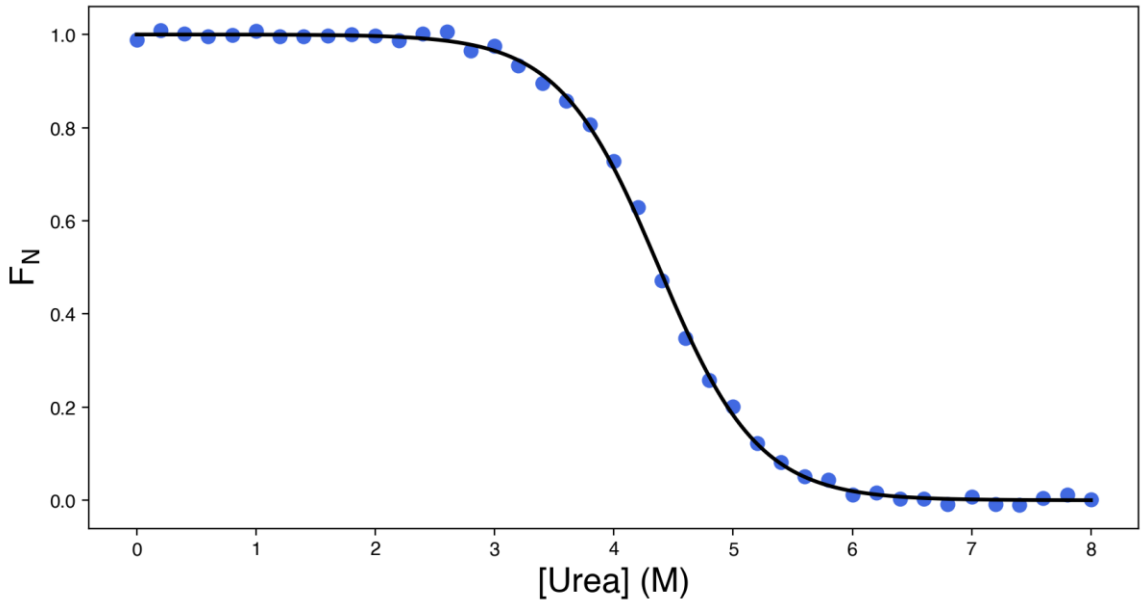

Figure S2. Urea equilibrium denaturation of pWT $(E\text{-}G5^2)_5$ measured using fluorescence spectroscopy in phosphate buffered saline, pH 7.4 at 25 °C. Presented as a fraction of folded/unfolded protein (F_N). A two-state model was fit to the data points (black). A single transition indicates cooperativity of the E and $G5^2$ domains during unfolding.

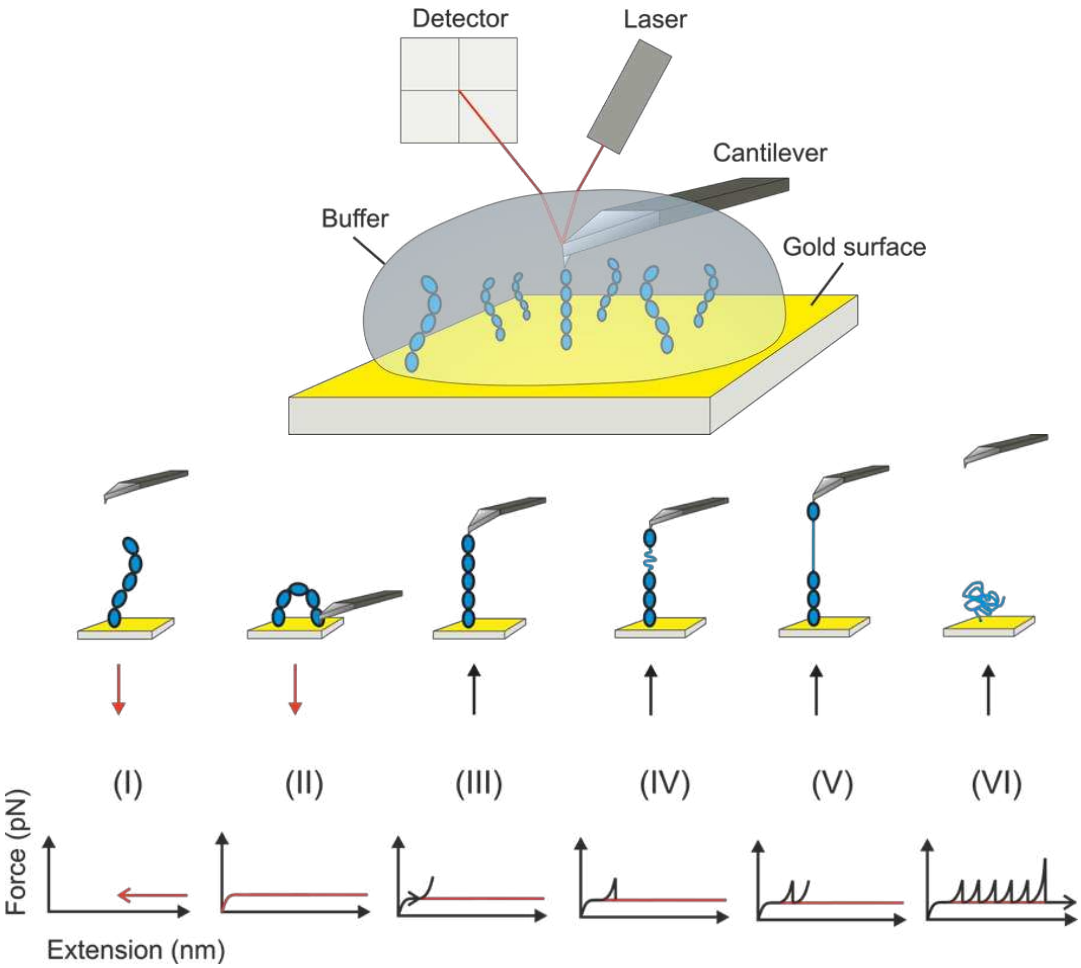

Figure S3. Top: schematic of a typical AFM setup for polyprotein unfolding assays.The gold surface is decorated in polyproteins (blue beads) via gold-thiol bonding. A laser is reflected from the tip of the cantilever probe, which detects any deviations (bending) of the cantilever in response to entropic restoration forces of protein domains under mechanical tension. Bottom: schematic showing the steps in obtaining a force-extension sawtooth profile for a five domain polyprotein unfolding at a constant retraction velocity. The features of the polyprotein cartoon correspond to the force-extension profile underneath. In force-extension mode, the cantilever is moved towards the surface under piezoelectric control (I) and presses into the surface causing a change in deflection (II). Once a threshold is reached the cantilever is retracted at a constant speed and if a protein has adsorbed onto the cantilever tip, an entropic restoring force is observable as the domain resists the forced decrease in entropy (due to preference of a polypeptide chain to form a random coil[1]) as the distance between the tip and the surface increases (III). The chain is extended to a point where the force exerted on the folded domain reaches a point capable of permitting thermally activated transitions and unfolding of the domain occurs causing an abrupt change in force. This force (apex of the 'tooth') at rupture (y axis) is considered the rupture force. The vertical linear decrease in force is because the domain unfolding occurs at a rate faster than the extension rate. (IV). This unfolded domain will continue to unravel adding to the effective length of the chain until fully stretched out, where an entropic restoring force will be observed again as force is applied across another domain (V). After all five domains have unfolded, the protein will detach from the tip resulting in a usually large detachment peak.

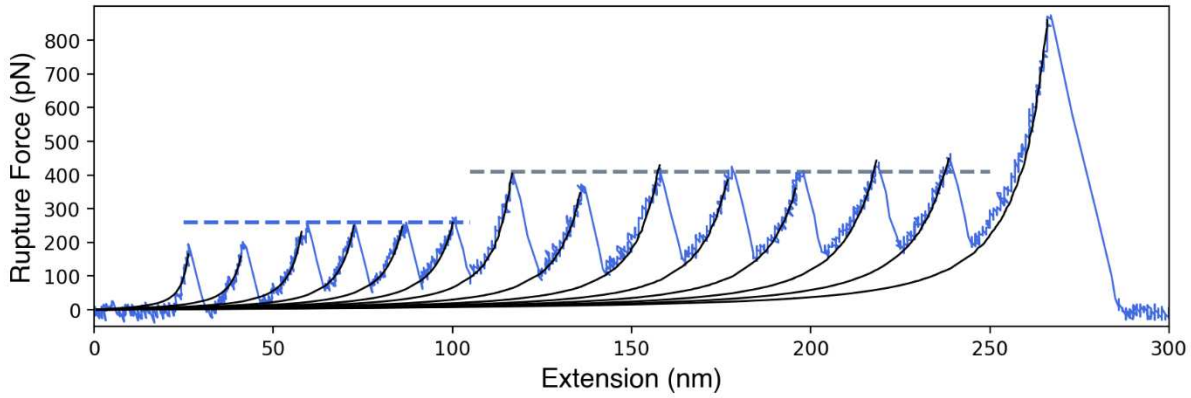

Figure S4. Example native SasG 'sawtooth' force-extension profiles displaying five events of both E and G5 sub domains unfolding at a retraction velocity of 1500 nms^{-1} in phosphate buffered saline, pH 7.4 at room temperature. WLC model fitted to E and G5 sub-domains are shown in black. The average unfolding forces for E and G5 at a retraction velocity of 1500 nms^{-1} are shown as blue and grey broken lines, respectively.

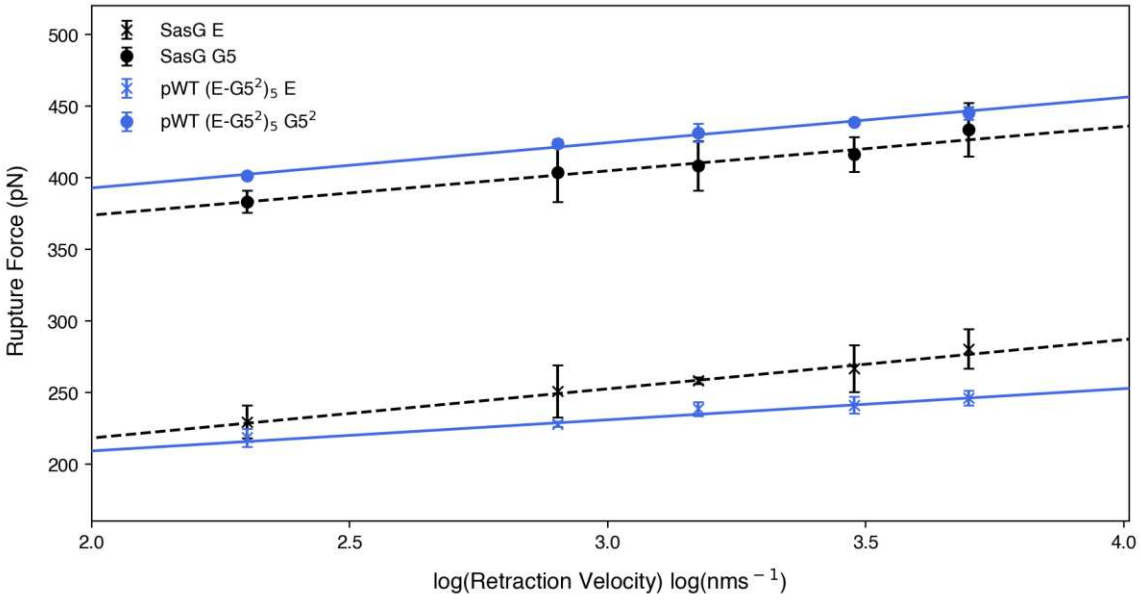

Figure S5. The dependence of rupture force on the logarithm of retraction velocity of E (crosses) and G5 (filled circles) sub domains of pWT $(E\text{-}G5^2)_5$ and SasG. Points and errors are the mean and standard deviation of triplicate datasets. Linear fit is weighted with the inverse of the standard deviation error of the triplicate datasets.

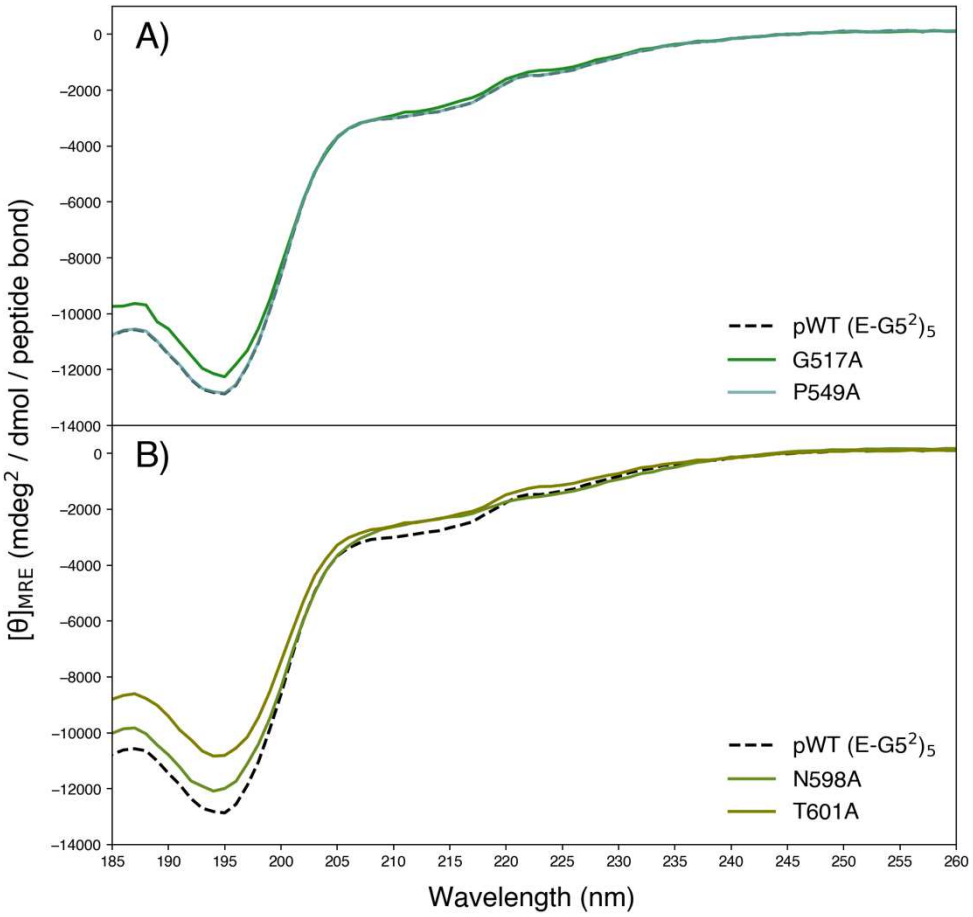

Figure S6. Far-UV CD spectra of the interface variants. **A)** G517A and P549A. **B)** N598A and T601A. These spectra show G517A and P549A have similar secondary structure to pWT $(E\text{-}G5^2)_5$. pWT $(E\text{-}G5^2)_5$ displayed in every spectrum for visualisation of secondary structure deviation.

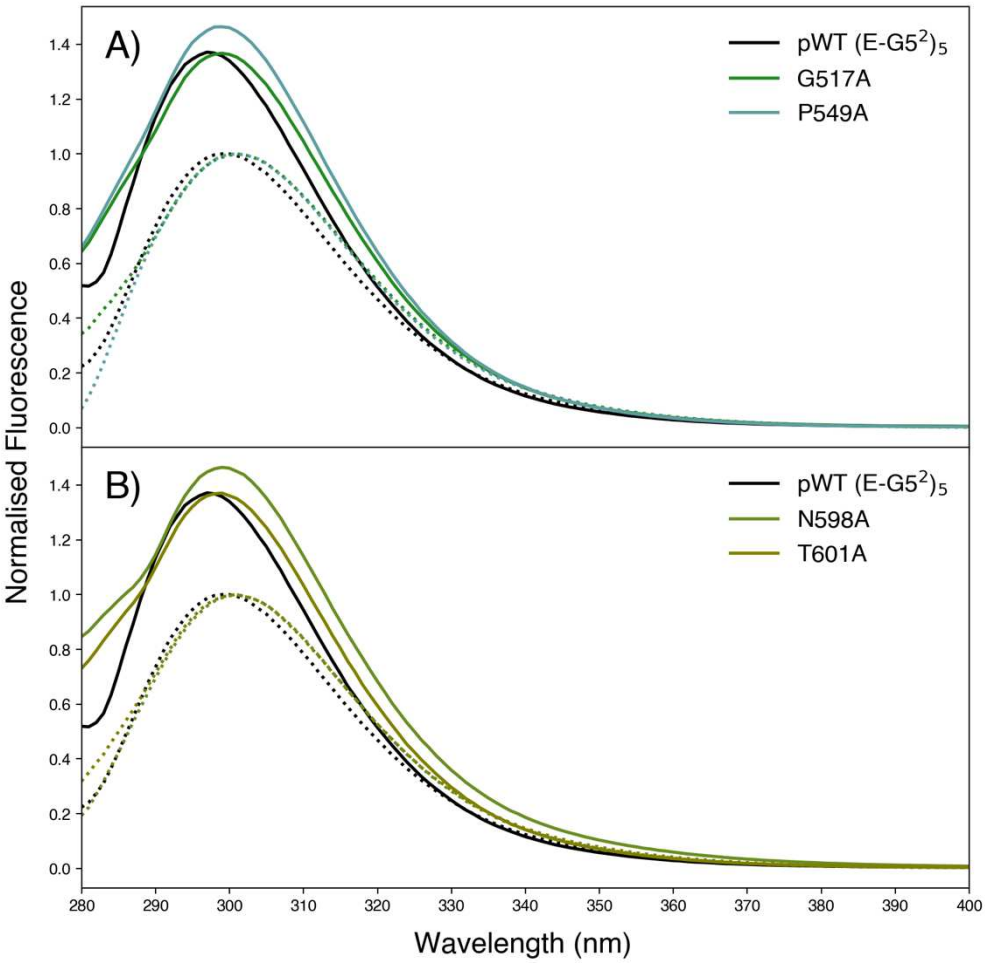

Figure S7. Intrinsic tyrosine fluorescence emission spectra of the interface variants. **A)** G517A and P549A and **B)** N598A and T601A. Absence of urea (solid line) and presence of 8 M urea (dashed lines). pWT $(E\text{-}G5^2)_5$ is included for reference. All variants display tertiary structure, which is lost in the presence of chemical denaturant.

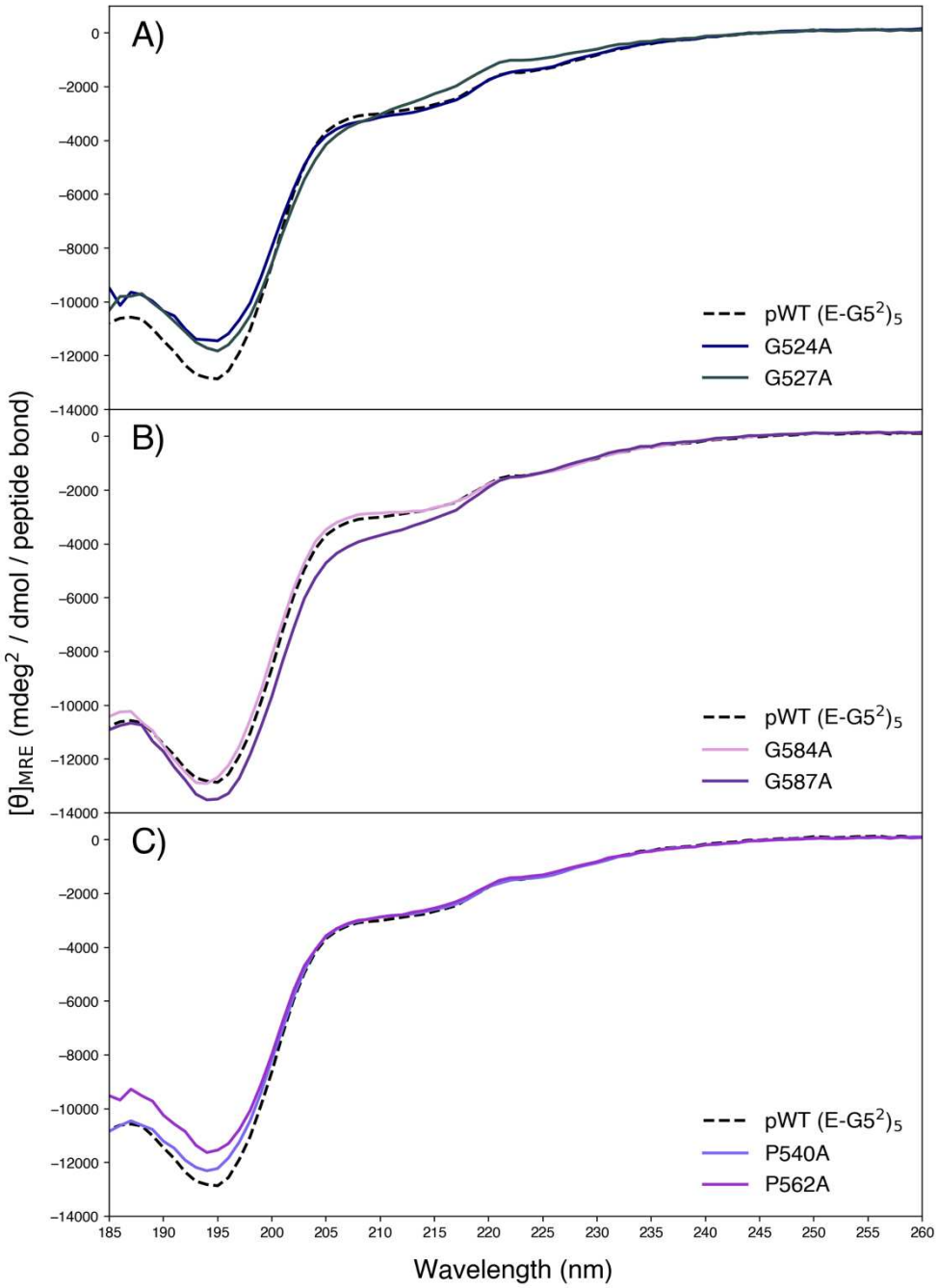

Figure S8. Far-UV CD spectra of the collagen-like motif variants. **A)** G524A and G527A situated in the E sub domain. **B)** G584A and G587A situated in the $G5^2$ sub domain. **C)** P540A and P562A, situated in the E and $G5^2$ sub domains, respectively. pWT $(E\text{-}G5^2)_5$ displayed in every spectrum for visualisation of secondary structure deviation.

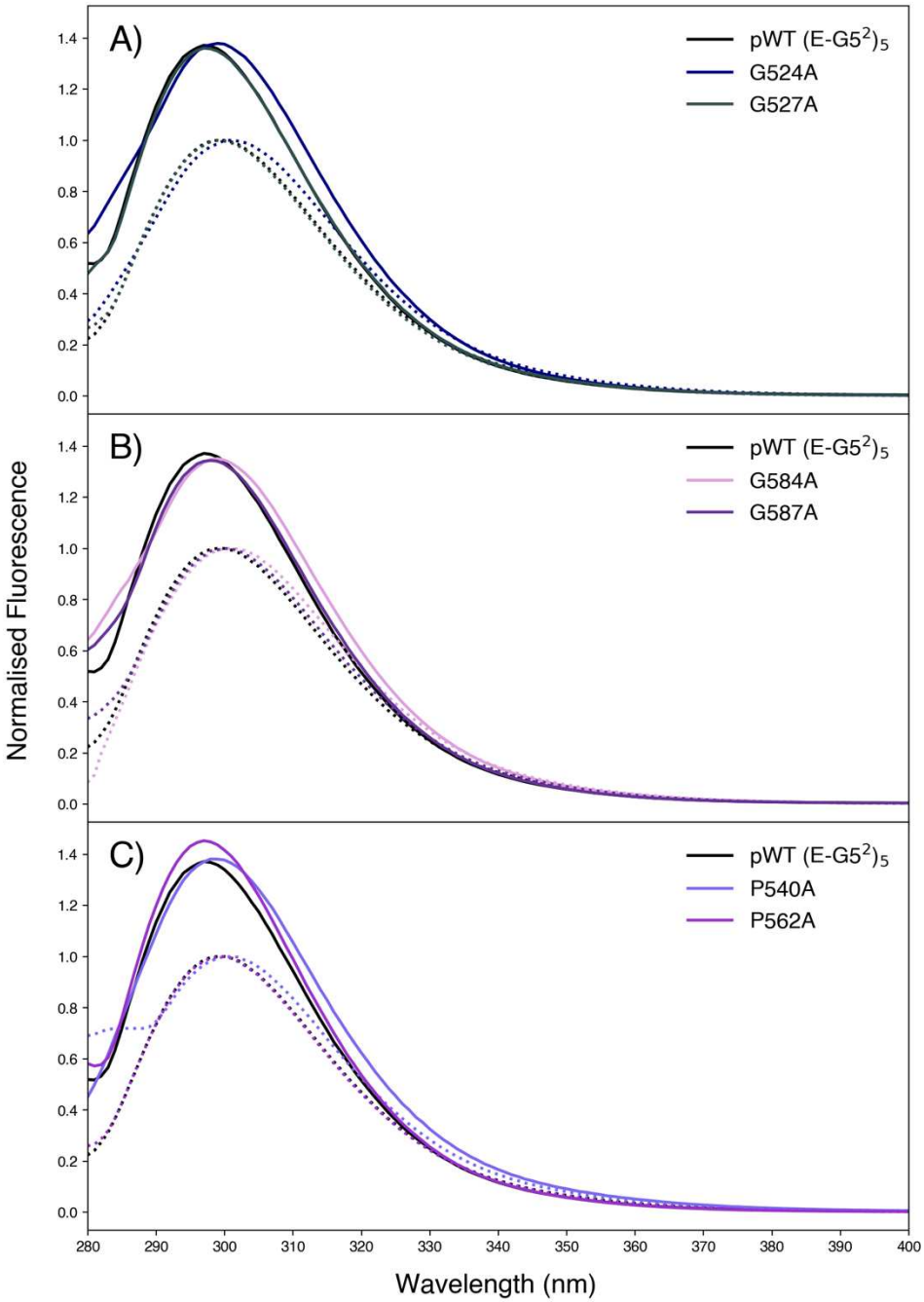

Figure S9. Intrinsic tyrosine fluorescence emission spectra of the collagen-like region variants. **A)** G524A and G527A, **B)** G584A and G587A, and **C)** P540A and P562A. Absence of urea (solid line) and presence of 8 M urea (dashed lines). pWT $(E\text{-}G5^2)_5$ is included for reference. All variants display tertiary structure, which is lost in the presence of chemical denaturant.

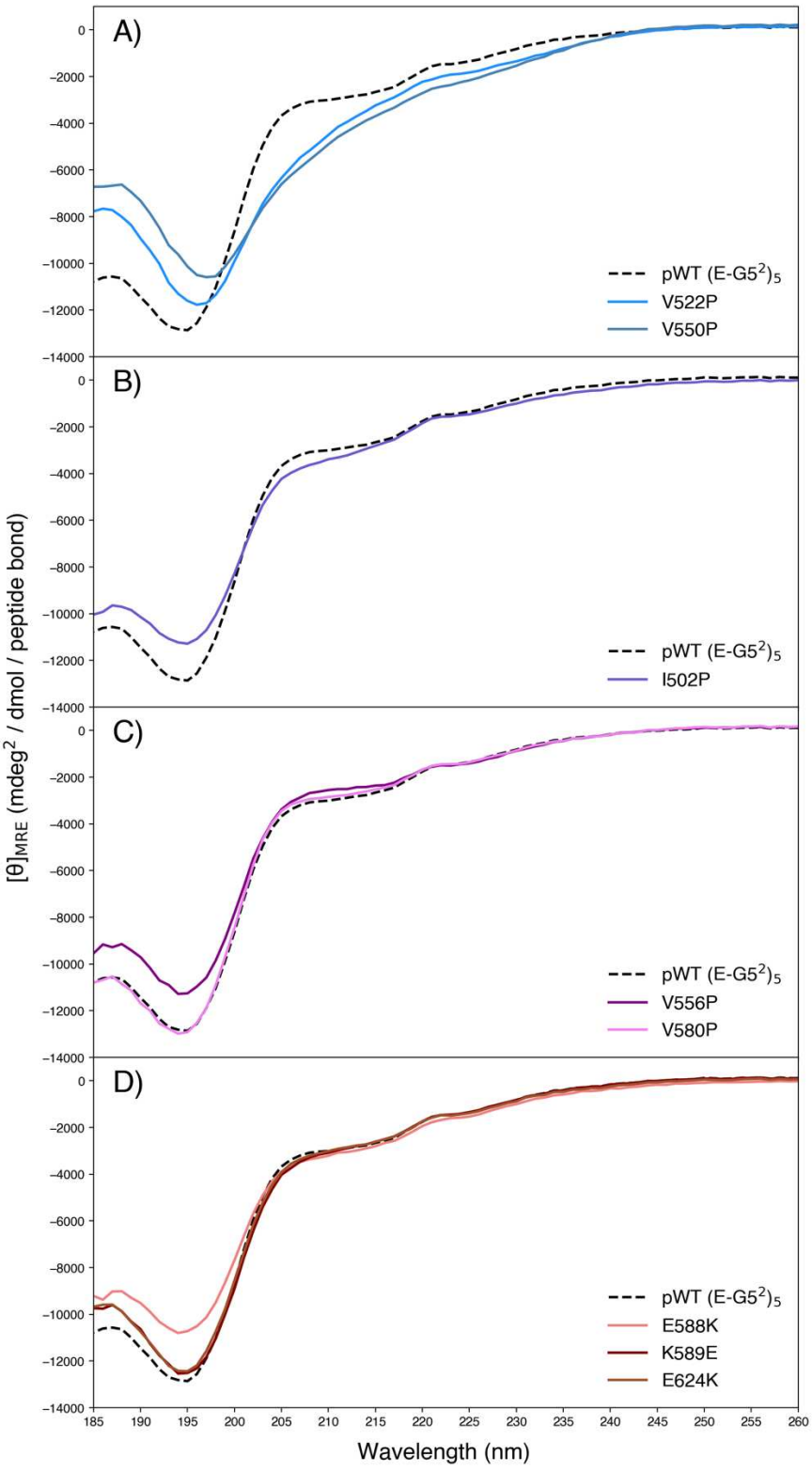

Figure S10. Far-UV CD spectra of the potential 'mechanical clamp' variants. **A)** V522P and V550P, **B)** I502P, **C)** V556P and V580P, and D) E588K, K589E and E624K. These spectra show proline mutations either severely (V522P and V550P) or moderately (I502P and V556P) affected the secondary structure. Charge-reversal variants E588K, K589E and E624K mostly retain secondary structure. pWT $(E\text{-}G5^2)_5$ displayed in every spectrum for visualisation of secondary structure deviation.

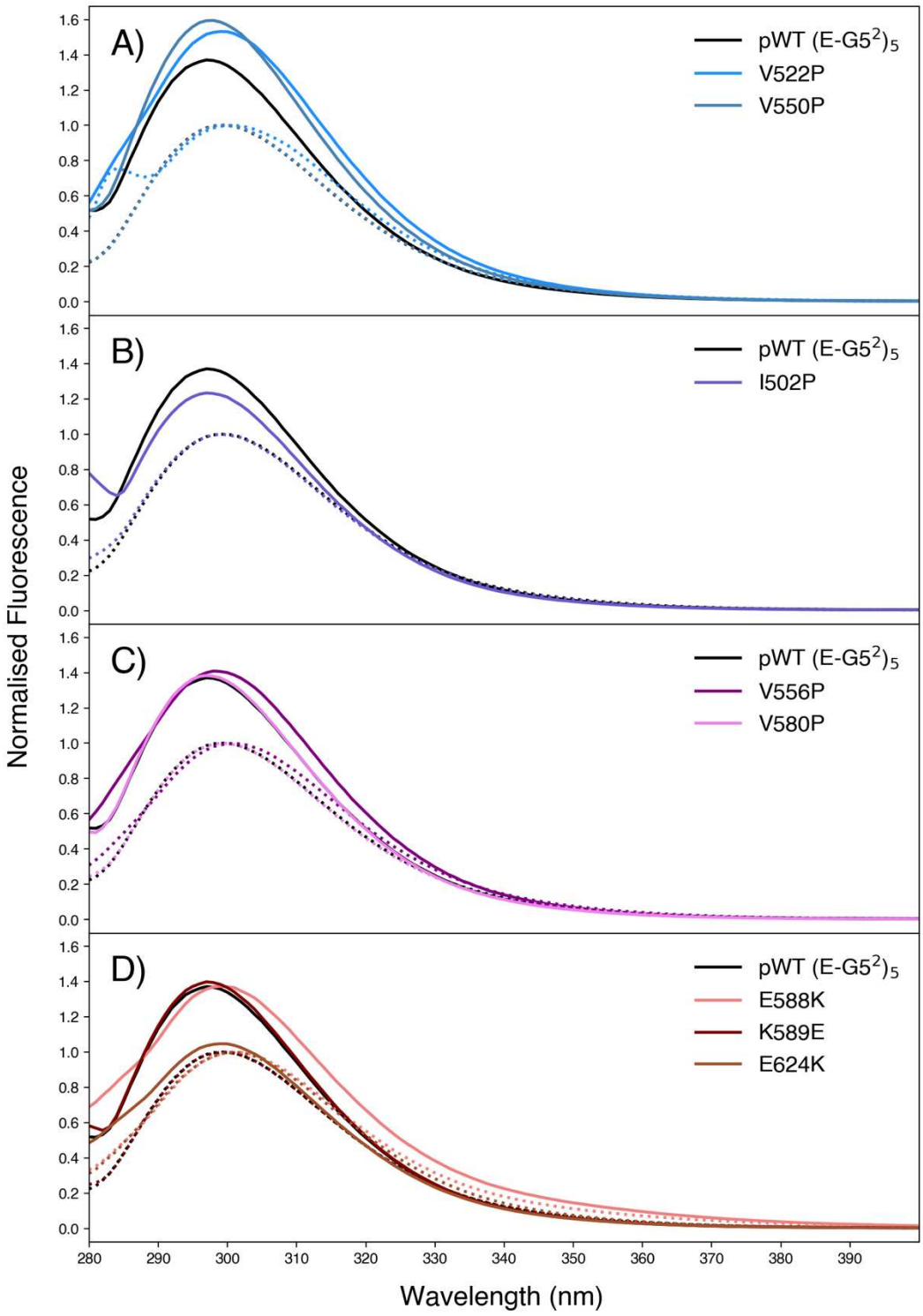

Figure S11. Intrinsic tyrosine fluorescence emission spectra of the potential 'mechanical clamp' proline variants. **A)** V522P and V550P, **B)** I502P, **C)** V556P and V580P, and D) E588K, K589E and E624K. Absence of urea (solid line) and presence of 8 M urea (dashed lines). pWT $(E\text{-}G5^2)_5$ is included for reference. All variants display tertiary structure, which is lost in the presence of chemical denaturant.

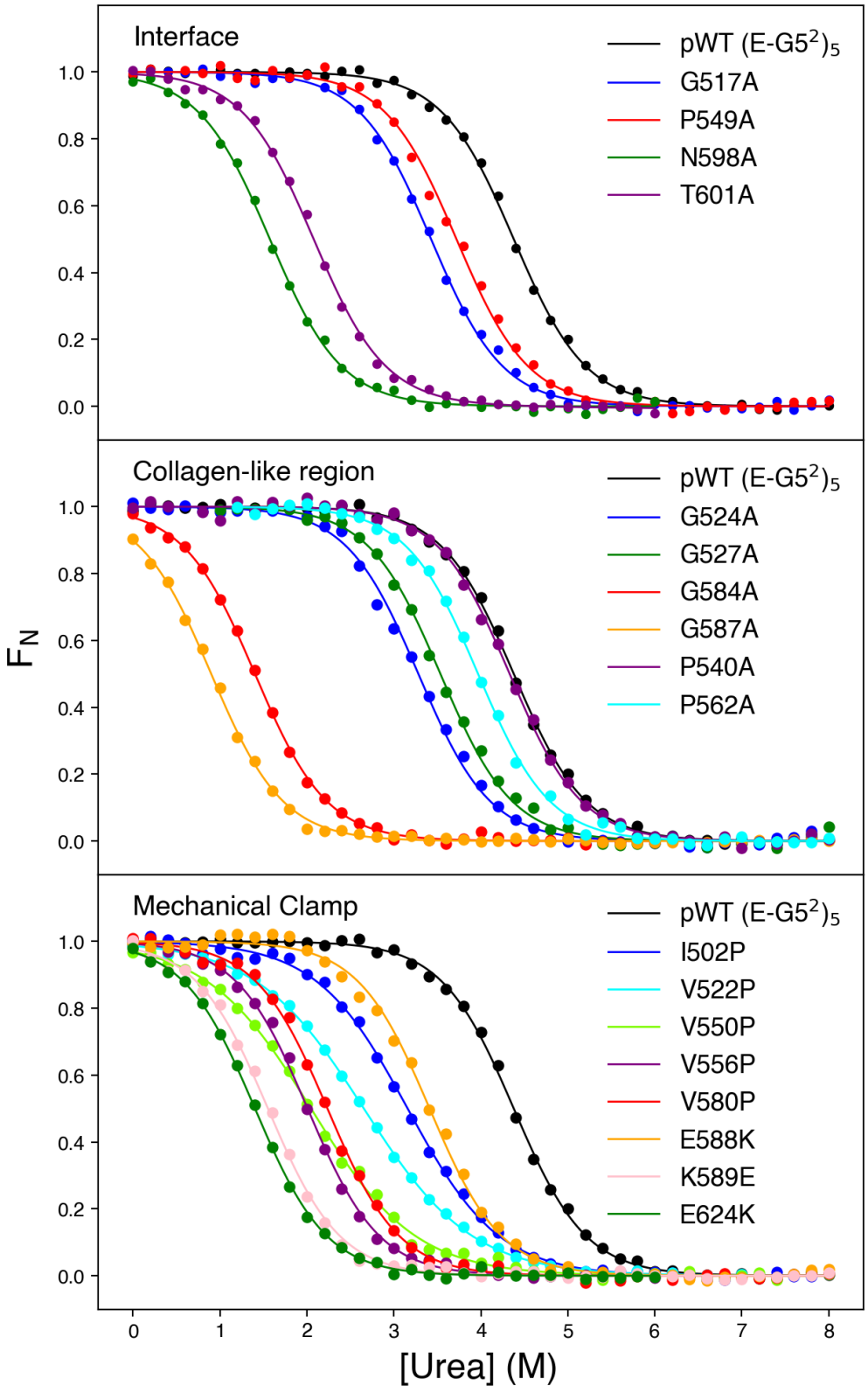

Figure S12. Urea equilibrium denaturation data for interface (top), collagen-like region (middle) and 'mechanical clamp' (bottom) variants obtained using fluorescence emission spectroscopy in phosphate buffered saline, pH 7.4 at 25 °C. Presented as a fraction of folded/unfolded protein (F_N). Lines show two-state to the data (filled circles). pWT $(E\text{-}G5^2)_5$ data included in every plot for comparison. m_{UN} fixed at 6.0 kJ mol^{-1} M^{-1} (corresponding to pWT $(E\text{-}G5^2)_5$) for all variants, with the exception of I502P, V522P, and V550P where the m_{UN} value was allowed to float.

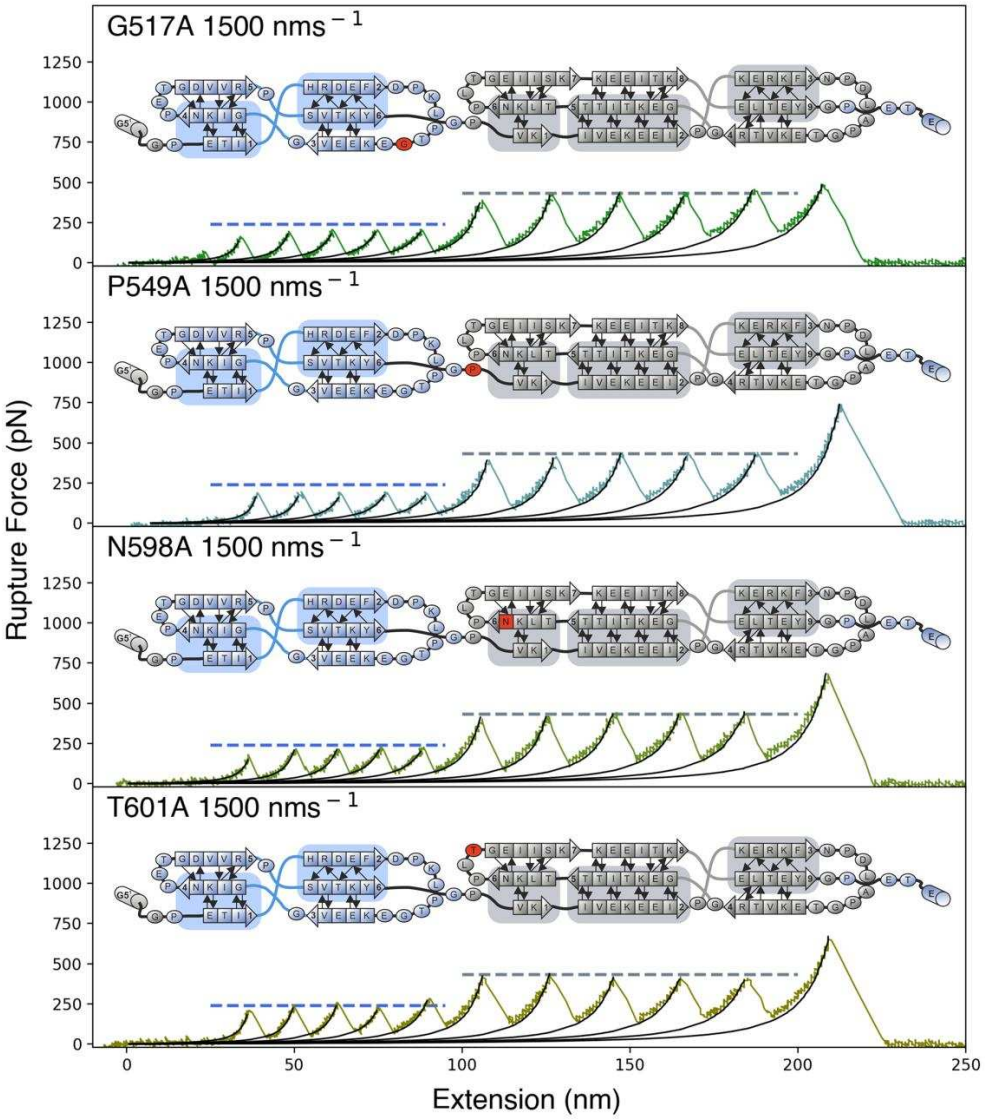

Figure S13. Typical force-distance profiles of the interface variants G517A, P549A, N598A and T601A in phosphate buffered saline, pH 7.4 at a retraction velocity of 1500 nms^{-1} at room temperature. WLC model fitting as black lines. Inset schematic highlights the position of the substituted residue in red. pWT (E-G5^2)$_5$ E and G5^2 average unfolding forces at a retraction velocity of 1500 nms^{-1} in blue and grey dashed lines, respectively.

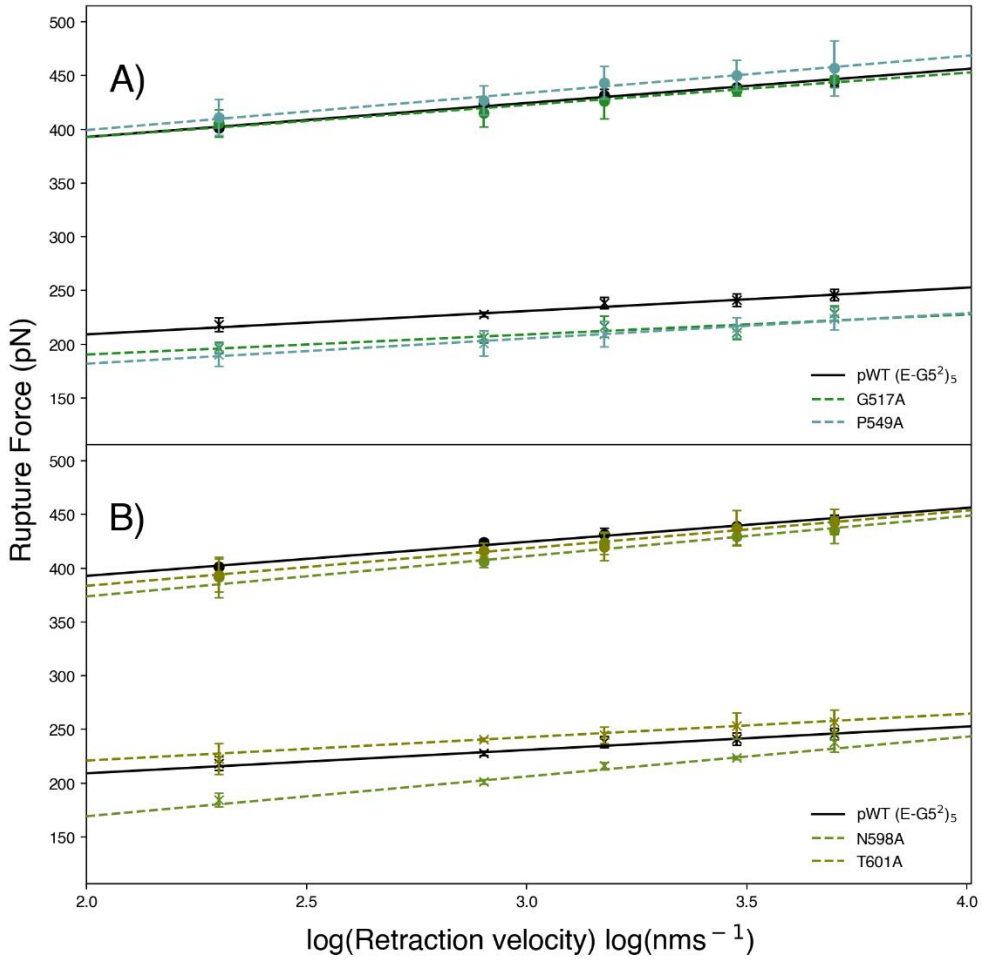

Figure S14. The dependence of rupture force on the logarithm of retraction velocity of the interface variants. E (crosses) and G5^2 (filled circles) subdomains of **A)** G517A and P549A and **B)** N598A and T601A variants. Points and errors are the mean and standard deviation of triplicate datasets. Linear fit is weighted with the inverse of the standard deviation error of the triplicate datasets.

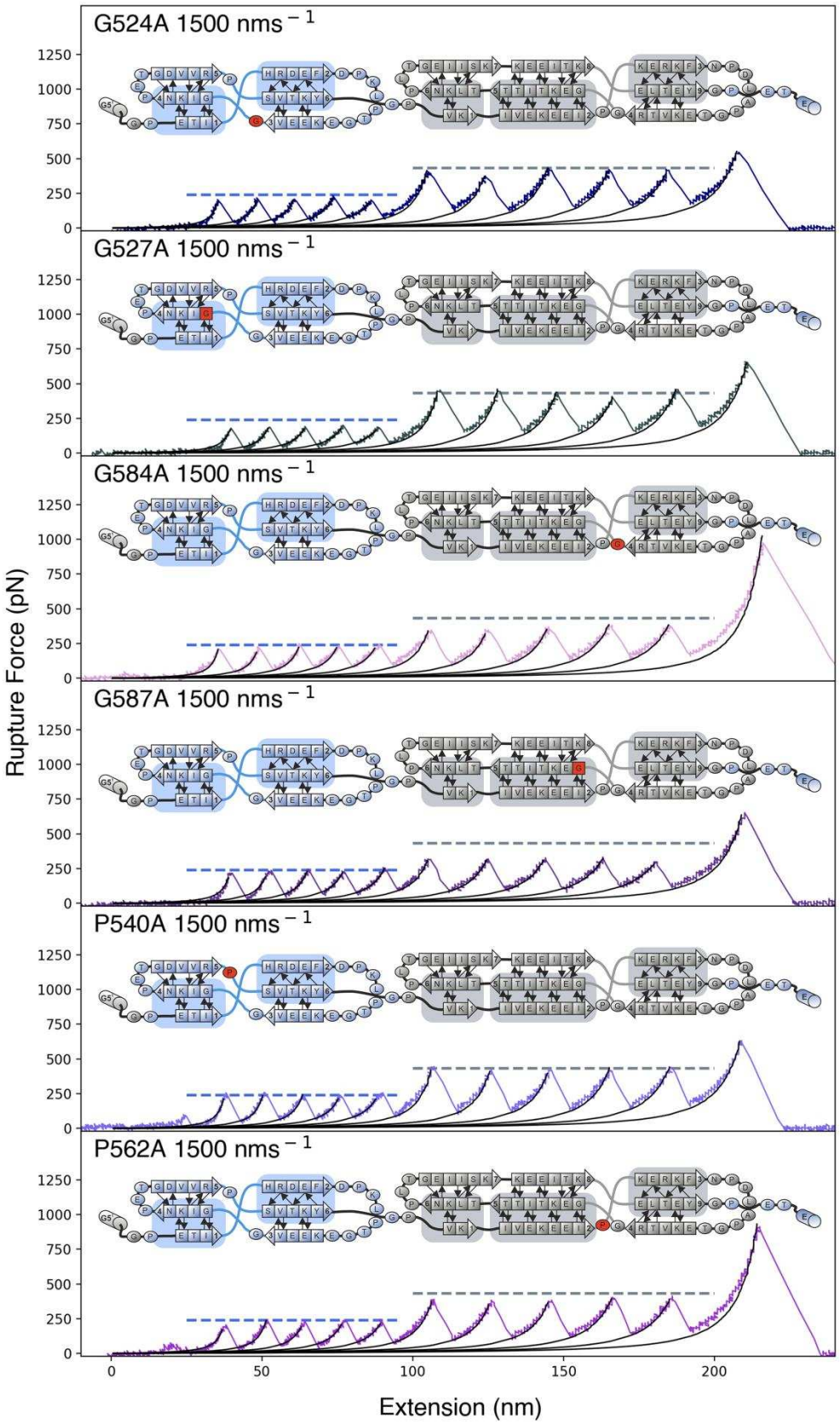

Figure S15. Typical force-distance profiles of the collagen-like motif variants G524A, G527A, G584A, G587A, P540A and P562A in phosphate buffered saline, pH 7.4 at a retraction velocity of 1500 nms^{-1} at room temperature.WLC model fitting as black lines. Inset schematic highlights the position of the substituted residue in red. pWT $(E\text{-}G5^2)_5$ E and $G5^2$ average unfolding forces at a retraction velocity of 1500 nms^{-1} in blue and grey dashed lines, respectively.

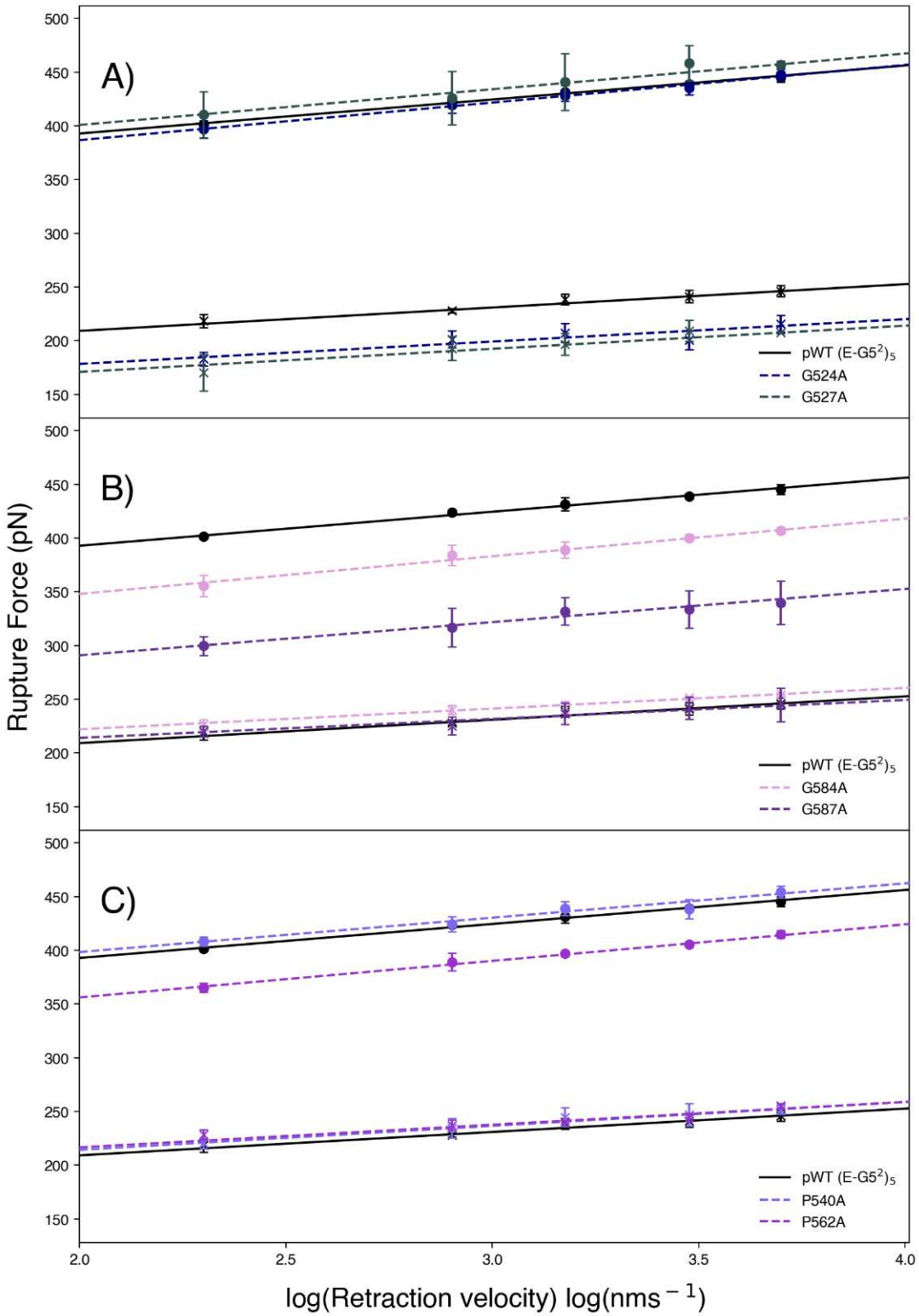

Figure S16. The dependence of rupture force on the logarithm of retraction velocity of the collagen-like motif variants. E (crosses) and G5^2 (filled circles) sub domains of **A)** G524A and G527A, **B)** G584A and G587A and **C)** P540A and P562A variants. Points and errors are the mean and standard deviation of triplicate datasets. Linear fit is weighted with the inverse of the standard deviation error of the triplicate datasets. There is an appreciable decrease in the E sub domain of G524A and G527A and in the G5^2 sub domain of G584A, G587A and P562A.

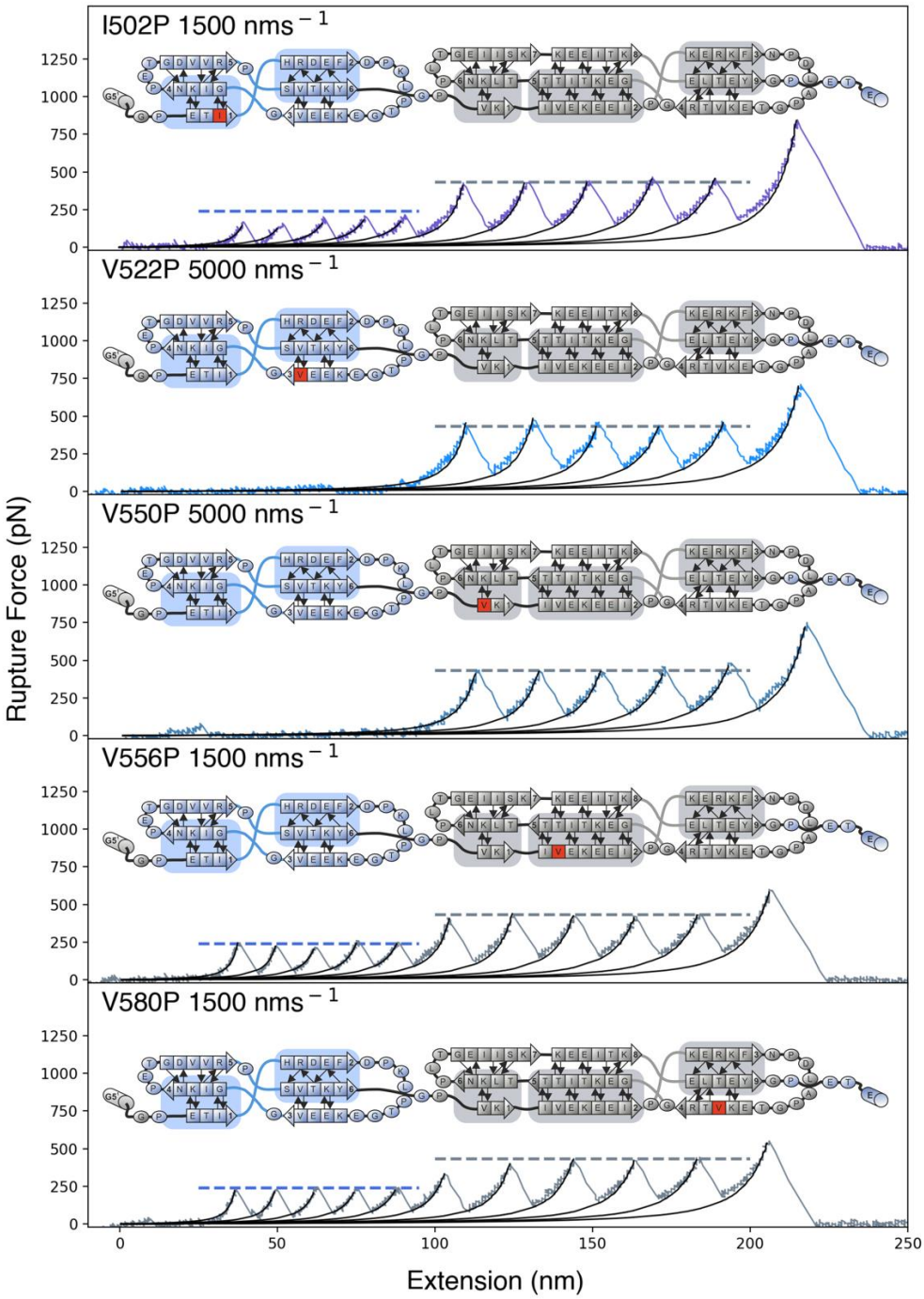

Figure S17. Typical force-distance profiles of the 'mechanical clamp' variants V522P and V550P at a retraction velocity of 5000 nms^{-1} and I502P, V556P and V580P at a retraction velocity of 1500 nms^{-1} in phosphate buffered saline, pH 7.4 at room temperature.WLC model fitting as black lines. Inset schematic highlights the position of the substituted residue in red. pWT $(E\text{-}G5^2)_5$ average unfolding forces at a retraction velocity of 1500 nms^{-1} for E and $G5^2$ in blue and grey dashed lines, respectively. There is an absence of peaks corresponding to E sub domain unfolding for both V522P and V550P.

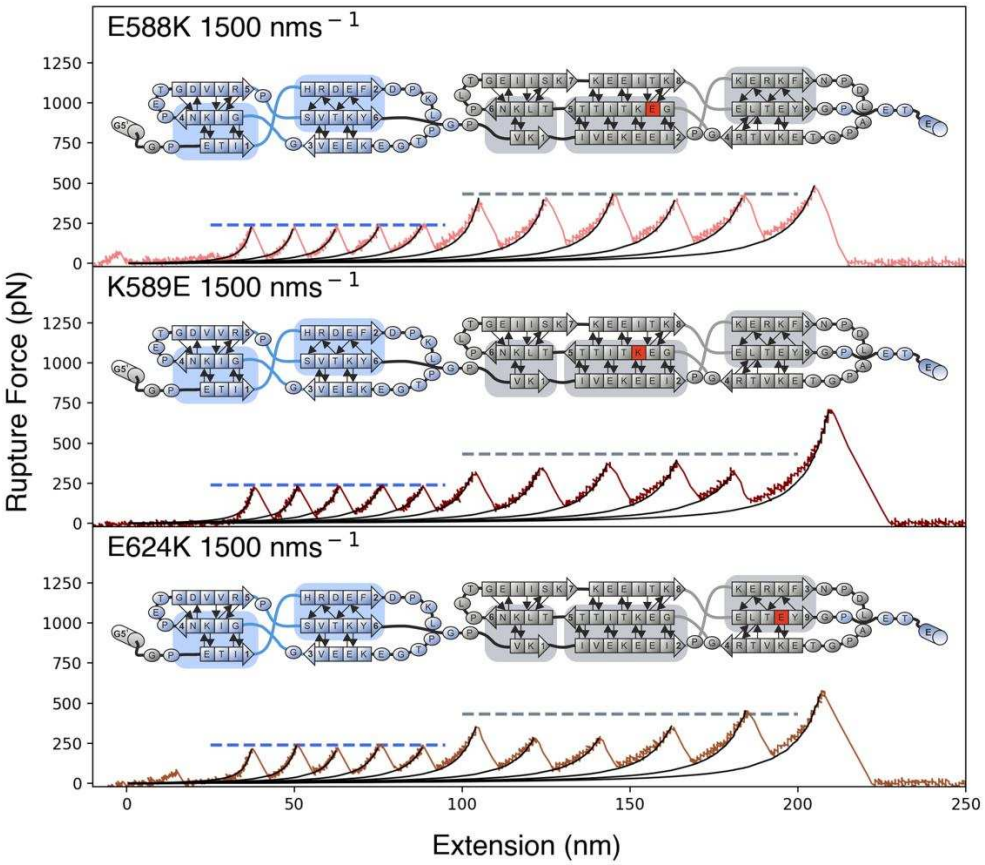

Figure S18. Typical force-distance profiles of the 'mechanical clamp' side-chain interaction variants E588K, K589E and E624K in 1X PBS, pH 7.4 at a retraction velocity of 1500 nms^{-1} at room temperature. WLC model fitting as black lines. Inset schematic highlights the position of the substituted residue in red. pWT $(E\text{-}G5^2)_5$ E and $G5^2$ average unfolding forces at a retraction velocity of 1500 nms^{-1} in blue and grey dashed lines, respectively.

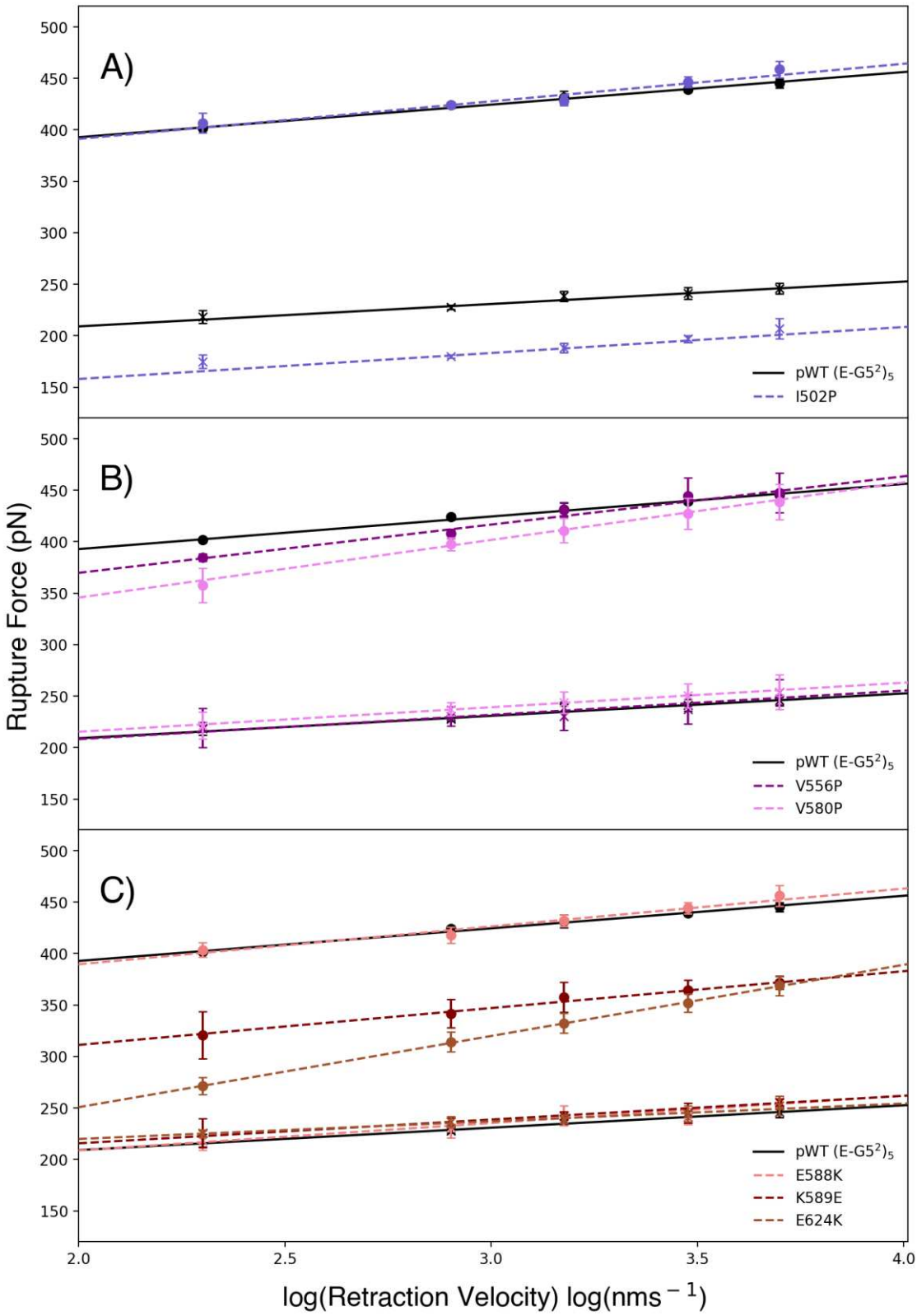

Figure S19. The dependence of rupture force on the logarithm of retraction velocity of the 'mechanical clamp' variants. E (crosses) and G5^2 (filled circles) sub domains of **A)** I502P, **B)** V556P and V580P and **C)** E588K, K589E and E624K variants. Points and errors are the mean and standard deviation of triplicate datasets. Linear fit is weighted with the inverse of the standard deviation error of the triplicate datasets.

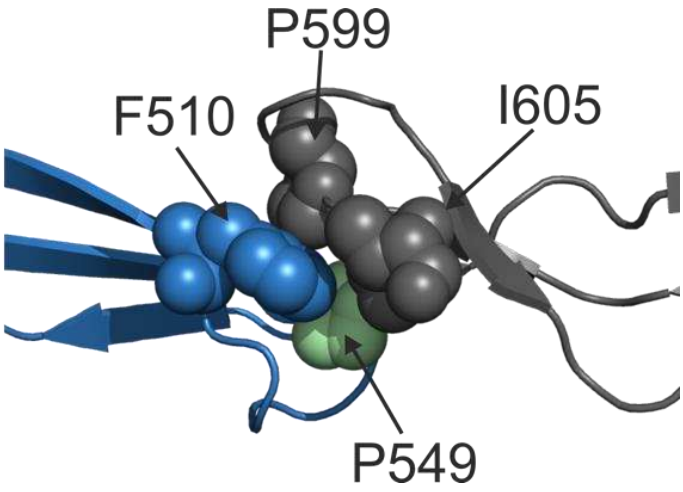

Figure S20. Important residues in the E-G5^2 pseudohydrophobic interface. The pseudohydrophobic core found at the interface comprises of F510, P549, P599 and I605.

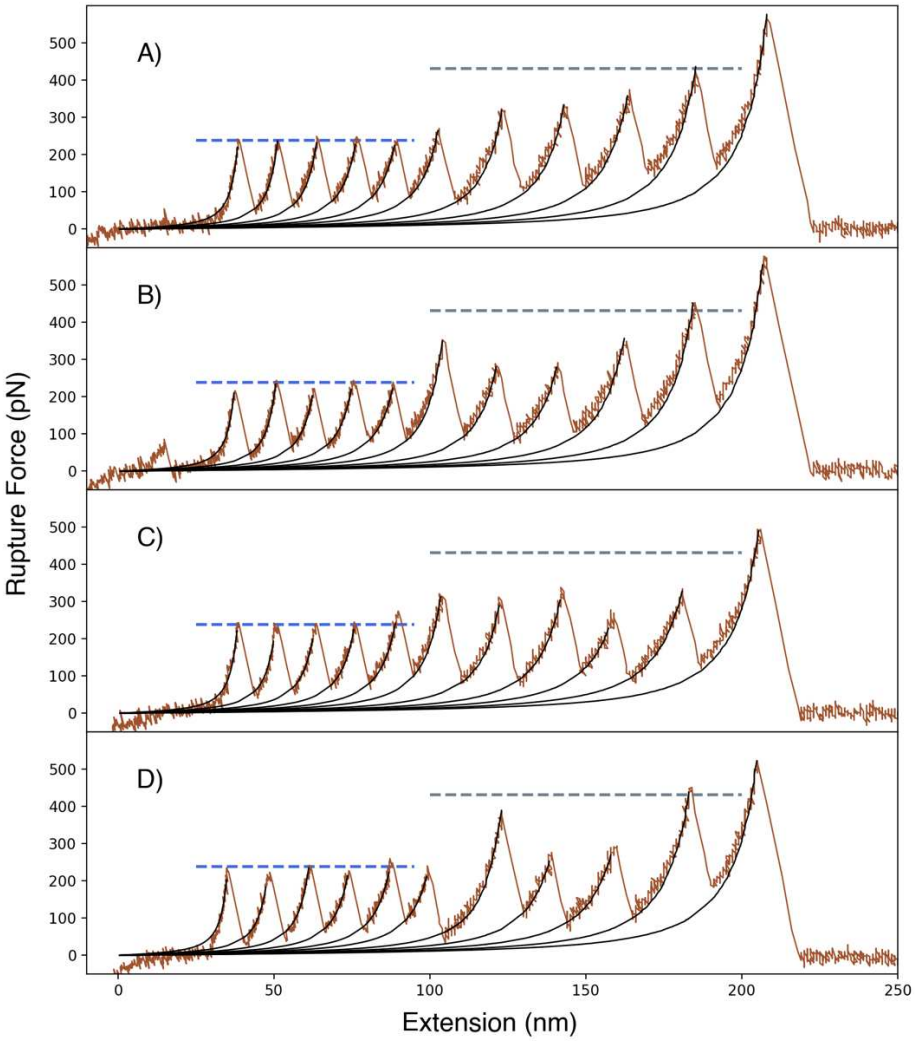

Figure S21. Examples of typical force-distance profiles of E624K in phosphate buffered saline, pH 7.4 at a retraction velocity of 1500 nms^{-1}. WLC model fitting as black lines. pWT (E-G5^2)$_5$ E and G5^2 average unfolding forces at a retraction velocity of 1500 nms^{-1} in blue and grey dashed lines, respectively. A-D) These force-distance profiles highlight the heterogeneity of G5^2 unfolding force, which varies both between and within force-distance profiles. The E sub domains are unchanged from pWT (E-G5^2)$_5$.

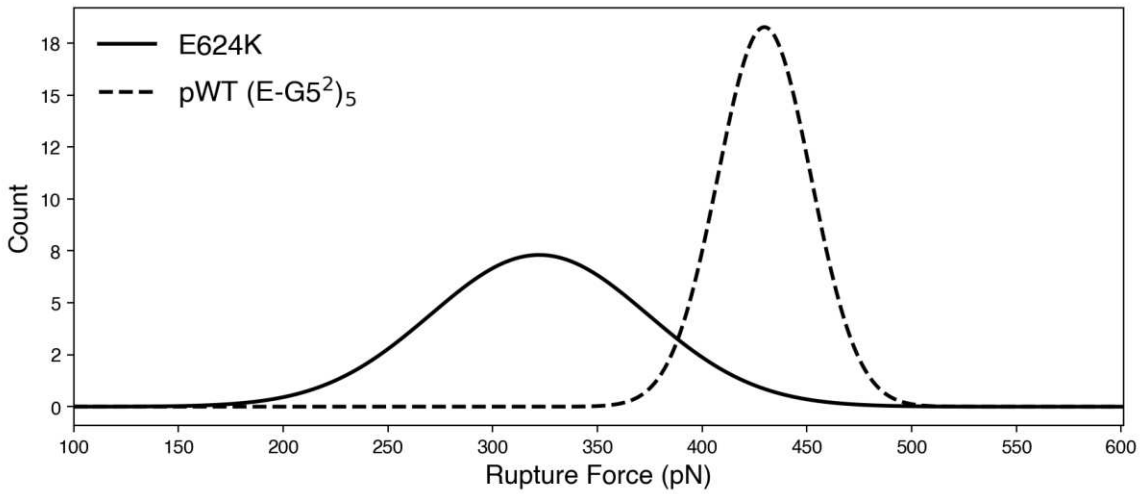

Figure S22. Visualisation of the increase of FWHM of E624K $(E\text{-}G5^2)_5$ $G5^2$ rupture force Gaussian fit in comparison to pWT $(E\text{-}G5^2)_5$ at a retraction velocity of 1500 nms^{-1}. For this replicate, the FWHM is 122.2 and 52.4 pN for E624K and pWT $(E\text{-}G5^2)_5$ $G5^2$ subdomain unfolding, respectively.

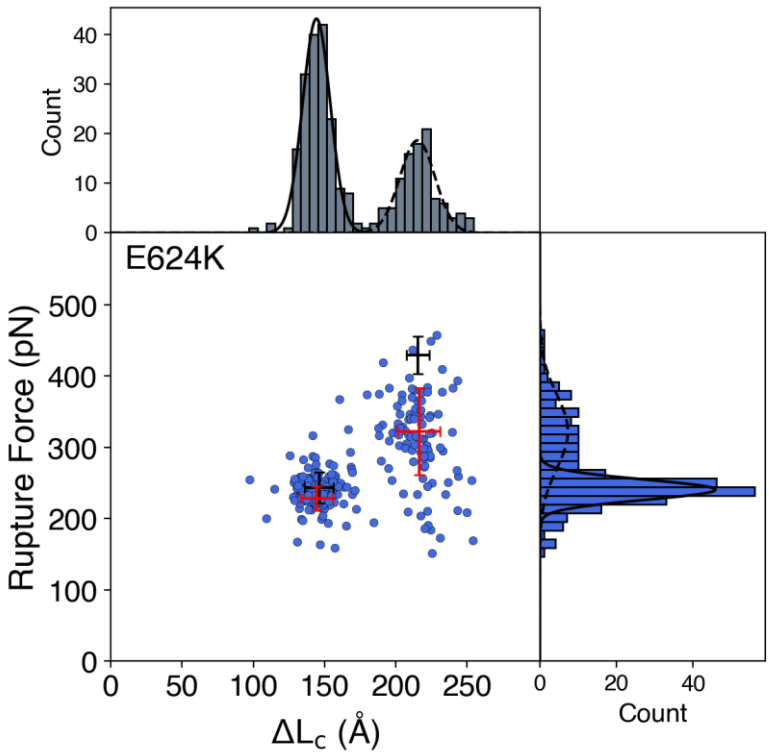

Figure S23. Example E624K $(E\text{-}G5^2)_5$ scatterplot with associated histograms at retraction velocity of 1500 nms^{-1} for one repeat. E and $G5^2$ Gaussian fits in black solid and dashed lines, respectively. Red scatterplot crosshairs are the mode and the FWHM from the corresponding marginal histogram Gaussian distribution fits. pWT $(E\text{-}G5^2)_5$ cross hairs from one repeat at the corresponding retraction velocity displayed in black for reference. Both the $\Delta F_{U,G5}$ and $\Delta L_{C,G5}$ distribution have a significantly wider FWHM for the E624K $(E\text{-}G5^2)_5$ $G5^2$ sub domain vs pWT $(E\text{-}G5^2)_5$.

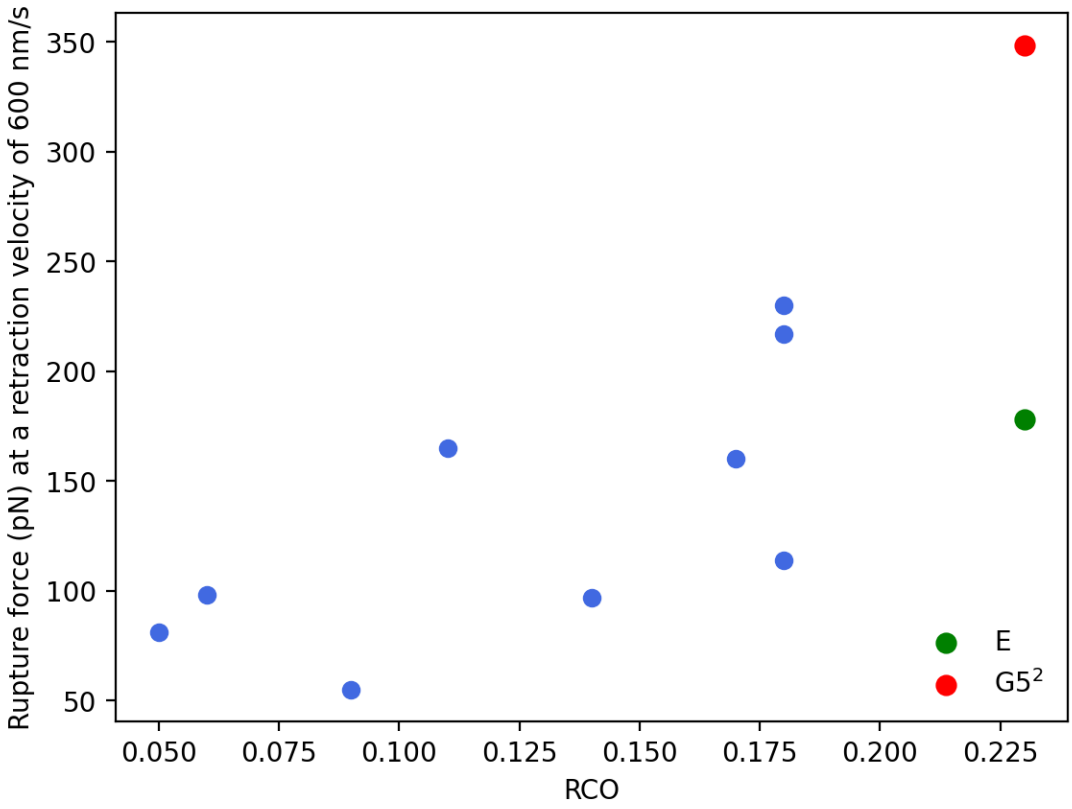

Figure S24. Scatterplot of RCO-rupture force at 600 nms^{-1}. There is a moderate positive correlation between RCO and rupture force at 600 nms^{-1} (R = 0.74). Plot includes proteins Fe-pfRD, I27, Top7, Tn 3FNIII, I1, 13FNIII, C2B, 10FNIII, C2A (data taken from reference [2]), and E and $G5^2$ of SasG. E and $G5^2$ display an identical RCO, but $G5^2$ displays a significantly greater mechanical strength. Protein RCO determined using Baker laboratory online calculator (available at https://depts.washington.edu/bakerpg/contact_order/ [3]).

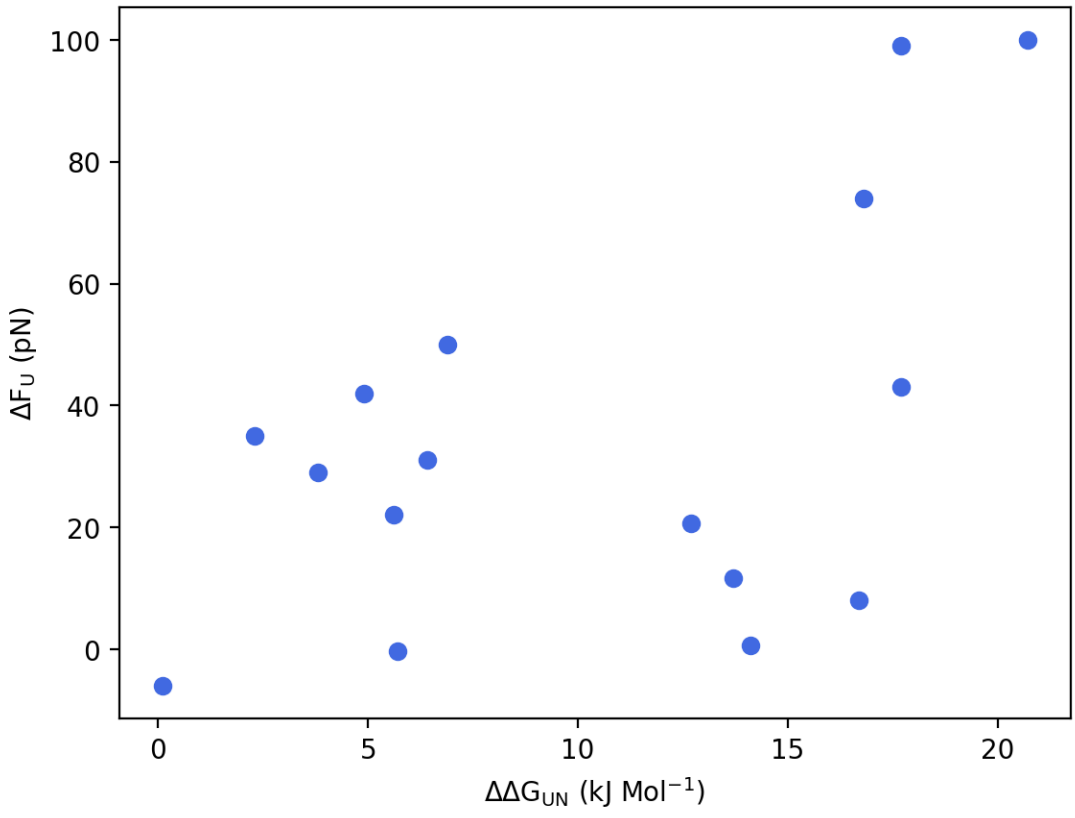

Figure S25. Scatterplot of $\Delta\Delta G_{UN}$-ΔF_U. There is no significant correlation between mechanical strength and thermodynamic stability. One datum is displayed per variant i.e. if the mutation is located in the E sub domain, the ΔF_U of the E sub domain ($\Delta F_{U,E}$) vs $\Delta\Delta G_{UN}$ is displayed.

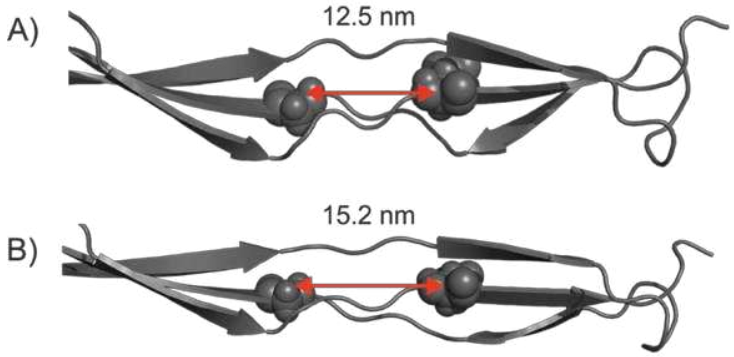

Figure S26. MD trajectory snapshots of the collagen-like region of $G5^2$ elongating and untwisting during the forced-unfolding simulations on E-$G5^2$.Measurement place markers on G587 and E621. A) structure at zero external force displays a native collagen-like structure measuring at 12.5 nm and B) elongated and slightly untwisted collagen-like region measuring at 15.2 nm prior to global unfolding.

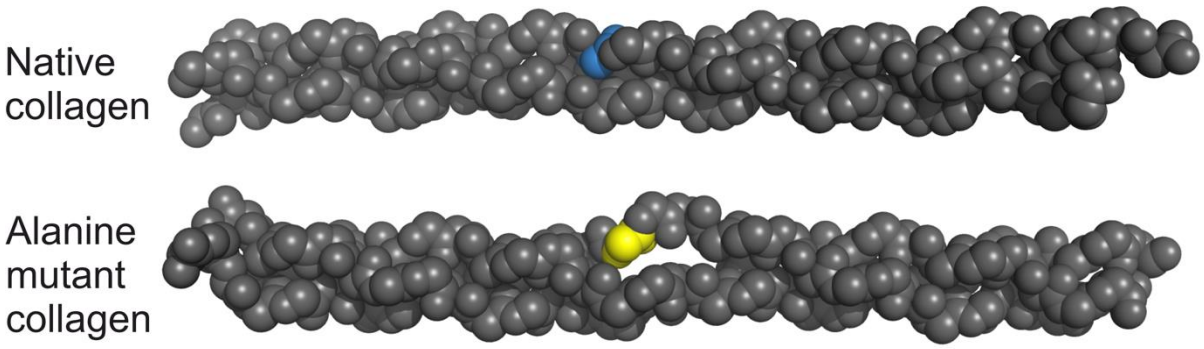

Figure S27. Glycine to alanine mutations in collagen. Native collagen structure (PDB: 1BKV[4]) with one of the repeating glycine residues (of $X_{aa}Y_{aa}GlyX_{aa}Y_{aa}Gly$) in blue and mutant collagen structure (PDB: 1CAG[5]) with the same glycine mutated to alanine (yellow). This produces local untwisting and a kink at the site of substitution due to accommodation of the side chain methyl group of alanine in the interior of the triple helix.

Location of substitution	Construct	Sub Domain Location	ΔG_{UN} (kJ mol^{-1})	m_{UN} (kJ mol^{-1} M^{-1})	$\Delta\Delta G_{UN}$ (kJ mol-1)
-	*$G5^2$	-	11.7 ± 0.7	4.2 ± 0.2	-
-	*E-$G5^2$	-	26.2 ± 0.8	5.9 ± 0.2	-
-	pWT (E-$G5^2$)$_5$	-	26.1 ± 0.8	6.0 ± 0.2	-
I	G517A (E-$G5^2$)$_5$	E	20.6 ± 0.1	**	5.6 ± 0.8
	P549A (E-$G5^2$)$_5$	GP linker	22.3 ± 0.2	**	3.8 ± 0.8
	N598A (E-$G5^2$)$_5$	$G5^2$	9.4 ± 0.2	**	16.7 ± 0.9
	T601A (E-$G5^2$)$_5$	$G5^2$	12.5 ± 0.2	**	13.7 ± 0.8
CLM	G524A (E-$G5^2$)$_5$	E	19.7 ± 0.2	**	6.4 ± 0.8
	G527A (E-$G5^2$)$_5$	E	21.2 ± 0.2	**	4.9 ± 0.8
	G584A (E-$G5^2$)$_5$	$G5^2$	8.4 ± 0.2	**	17.7 ± 0.9
	G587A (E-$G5^2$)$_5$	$G5^2$	5.4 ± 0.5	**	20.7 ± 0.9
	P540A (E-$G5^2$)$_5$	E	26.0 ± 0.1	**	0.1 ± 0.8
	P562A (E-$G5^2$)$_5$	$G5^2$	23.8 ± 0.2	**	2.3 ± 0.8
MC	I502P (E-$G5^2$)$_5$	E	14.9 ± 0.7	4.7 ± 0.2	11.2 ± 1.1
	V522P (E-$G5^2$)$_5$	E	10.6 ± 0.4	4.0 ± 0.1	15.6 ± 0.9
	V550P (E-$G5^2$)$_5$	$G5^2$	8.5 ± 0.9	4.1 ± 0.3	17.7 ± 0.9
	V556P (E-$G5^2$)$_5$	$G5^2$	12.0 ± 0.2	**	14.1 ± 0.8
	V580P (E-$G5^2$)$_5$	$G5^2$	13.4 ± 0.2	**	12.7 ± 0.8
	E588K (E-$G5^2$)$_5$	$G5^2$	20.5 ± 0.2	**	5.7 ± 0.8
	K589E (E-$G5^2$)$_5$	$G5^2$	9.3 ± 0.3	**	16.8 ± 0.9
	E624K (E-$G5^2$)$_5$	$G5^2$	8.4 ± 0.2	**	17.7 ± 0.9

Table S1. Thermodynamic parameters of monomeric $G5^2$, E-$G5^2$, pWT (E-$G5^2$)$_5$, and variants thereof obtained using equilibrium denaturation monitored by intrinsic tyrosine fluorescence emission. ΔG_{UN} and m_{UN} displayed errors are the errors of the fit. $\Delta\Delta G_{UN}$ calculated from $\Delta G_{UN,pWT}$ - $\Delta G_{UN,Mut}$ with the propagated error. I: interface, CLM: collagen-like motif and MC: 'Mechanical clamp'. * values taken from reference[6]. **m_{UN} value fixed at 6.0 kJ mol^{-1} M^{-1}. m_{UN} value left to float for I502P, V522P, and V550P as spectroscopic evidence suggests partial unfolding or complete unfolding of the E sub domain in these variants. All data presented in this table obtained in phosphate buffered saline at 25 °C.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average ΔL_C (Å)
200	135	213.3	224.1	146.5	146.3
	218	224.2		146.3	
	116	217.6		146.0	
800	151	225.7	227.9	147.3	146.0
	150	229.4		146.0	
	171	228.6		144.8	
1500	248	233.6	238.5	151.1	147.7
	266	243.8		146.1	
	120	238.1		145.9	
3000	108	234.9	241.2	150.7	147.7
	217	245.9		147.1	
	84	242.9		145.5	
5000	79	240.3	246.1	149.1	146.4
	130	250.4		146.7	
	100	247.6		143.5	

Table S2. Summary of rupture force and ΔL_C statistics for pWT $(E\text{-}G5^2)_5$ E subdomain mechanical unfolding in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode rupture force and ΔLC are obtained from the Gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average ΔL_C (Å)
200	53	399.6	401.6	216.3	215.4
	81	401.6		215.0	
	41	403.6		214.7	
800	114	423.9	423.9	215.2	214.8
	83	421.1		215.0	
	66	426.6		214.0	
1500	235	426.5	431.5	215.3	214.8
	104	429.7		215.6	
	40	438.4		213.4	
3000	79	437.0	439.0	216.7	213.8
	188	438.7		214.0	
	44	441.4		210.7	
5000	42	441.1	445.2	216.7	214.7
	60	444.7		213.3	
	68	449.8		214.1	

Table S3. Summary of rupture force and ΔL_C statistics for pWT $(E\text{-}G5^2)_5$ $G5^2$ subdomain mechanical unfolding in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode rupture force and ΔL_C are obtained from the Gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average ΔL_C (Å)
200	213 102 188	241.5 219.0 228.1	229.5	154.5 149.1 151.0	151.5
800	117 94 143	267.6 236.1 248.6	250.7	153.8 149.0 154.8	152.5
1500	87 114 107	276.3 256.6 260.1	258.4	154.7 148.8 147.9	150.5
3000	79 116 127	285.5 260.0 254.9	266.8	157.5 150.0 152.0	153.2
5000	71 131 121	293.8 266.3 281.4	280.5	157.0 153.2 151.5	153.9

Table S4. Summary of rupture force and ΔL_C statistics for SasG E subdomain mechanical unfolding in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode rupture force and ΔL_C are obtained from the Gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average ΔL_C (Å)
200	87 117 218	388.6 374.5 386.9	383.3	220.9 215.3 218.2	218.2
800	133 89 243	422.0 387.7 401.0	403.6	218.7 217.7 218.4	218.3
1500	82 106 134	433.4 396.2 420.6	408.4	222.8 217.2 214.4	218.1
3000	102 91 143	428.1 404.1 417.0	416.4	222.4 217.0 217.1	218.8
5000	124 163 154	450.5 413.3 437.2	433.7	224.5 220.4 216.1	220.3

Table S5. Summary of rupture force and ΔL_C statistics for SasG G^5 subdomain mechanical unfolding in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode rupture force and ΔL_C are obtained from the Gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Variant	Expected Mass (Da)	Measured Mass (Da)
pWT $(E\text{-}G5^2)_5$	76095.0	76095.7 ± 0.9
G517A $(E\text{-}G5^2)_5$	76165.2	76163.8 ± 0.3
P549A $(E\text{-}G5^2)_5$	75964.9	75965.1 ± 0.5
N598A $(E\text{-}G5^2)_5$	75879.9	75880.3 ± 0.4
T601A $(E\text{-}G5^2)_5$	75944.9	75943.9 ± 0.3
G524A $(E\text{-}G5^2)_5$	76165.2	76165.6 ± 1.0
G527A $(E\text{-}G5^2)_5$	76165.2	76163.6 ± 2.0
G584A $(E\text{-}G5^2)_5$	76165.2	76166.0 ± 1.0
G587A $(E\text{-}G5^2)_5$	76165.2	76164.0 ± 0.6
P540A $(E\text{-}G5^2)_5$	75964.9	75963.1 ± 0.9
P562A $(E\text{-}G5^2)_5$	75964.9	75964.8 ± 0.3
I502P $(E\text{-}G5^2)_5$	76014.8	76014.4 ± 0.8
V522P $(E\text{-}G5^2)_5$	76085.0	76084.2 ± 1.6
V550P $(E\text{-}G5^2)_5$	76085.0	76083.7 ± 2.9
V556P $(E\text{-}G5^2)_5$	76085.0	76085.5 ± 1.8
V580P $(E\text{-}G5^2)_5$	76085.0	76085.8 ± 2.0
E588K $(E\text{-}G5^2)_5$	76090.3	76090.7 ± 0.5
K589E $(E\text{-}G5^2)_5$	76099.8	76099.6 ± 0.5
E624K $(E\text{-}G5^2)_5$	76090.3	76092.4 ± 3.1

Table S6. Table of pWT $(E\text{-}G5^2)_5$ and variants thereof, the expected mass and the measured mass by LC-MS.

Group	Construct	Sub Domain Location	E ΔL_C Range (Å)	$G5^2$ ΔL_C Range or modal value** (Å)
I	G517A $(E\text{-}G5^2)_5$	E	142.6-150.6	213.9-216.0
	P549A $(E\text{-}G5^2)_5$	GP linker	145.0-154.4	213.9-229.9
	N598A $(E\text{-}G5^2)_5$	$G5^2$	141.7-150.4	213.0-216.1
	T601A $(E\text{-}G5^2)_5$	$G5^2$	144.5-150.4	211.5-219.1
CLM	G524A $(E\text{-}G5^2)_5$	E	142.6-150.3	213.3-217.4
	G527A $(E\text{-}G5^2)_5$	E	143.3-150.5	212.6-217.7
	G584A $(E\text{-}G5^2)_5$	$G5^2$	143.5-148.5	213.6-217.4
	G587A $(E\text{-}G5^2)_5$	$G5^2$	140.7-147.9	210.9-217.0
	P540A $(E\text{-}G5^2)_5$	E	145.0-153.4	213.4-217.2
	P562A $(E\text{-}G5^2)_5$	$G5^2$	143.0-151.5	211.9-216.9
MC	I502P $(E\text{-}G5^2)_5$	E	140.5-149.5	212.7-216.0
	V522P* $(E\text{-}G5^2)_5$	E	-	216.9**
	V550P* $(E\text{-}G5^2)_5$	$G5^2$	-	215.4**
	V556P $(E\text{-}G5^2)_5$	$G5^2$	143.8-164.1	209.5-219.0
	V580P $(E\text{-}G5^2)_5$	$G5^2$	144.5-148.4	213.4-217.3
	E588K $(E\text{-}G5^2)_5$	$G5^2$	144.0-158.9	212.4-219.9
	K589E $(E\text{-}G5^2)_5$	$G5^2$	143.6-147.6	210.0-216.8
	E624K $(E\text{-}G5^2)_5$	$G5^2$	143.0-145.9	214.2-218.5

Table S7. Table of ΔL_C values for each variant mechanically unfolded in phosphate buffered saline, pH 7.4 at room temperature. ΔL_C values are the range of the Gaussian fit modal values of the triplicate repeats, with the exception of V522P and V550P. *V522P and V550P modal values at 5000 nms^{-1}, and the corresponding $G5^2$ ΔL_C** is the modal value from the single repeat Gaussian fits. All ΔL_C values are consistent with the pWT $(E\text{-}G5^2)_5$ corresponding domain. I: interface, CLM: collagen-like motif and MC: 'Mechanical clamp'.

Description	Forward Sequence	Reverse Sequence
MCS BsaI site removal	5'- ATAGGGAGACAACGGTTTC-3'	5'- AGTGAGTCGTATTAATTTC-3'
AmpR gene BsaI site removal	5'-AGCGTGGGTCCCGCGGTATCA-3'	5'- CACCGGCTCCAGATTTATCAG-3'

Table S8. Primers to remove BsaI sites from pET14b to create pET14b$\Delta bsaI$. Substituted bases in red text.

Mutant	Forward Sequence	Reverse Sequence
I502P	*	*
G517A	5'-TTACCTACCGCTGAGAAGGAA-3'	5'-CTTCGGATCAAACTCATCACGATG-3'
V522P	5'-GAAGGAAGAACCTCCTGGTAAGCCGG-3'	5'-TCACCGGTAGGTAACTTC-3'
G524A	5'-GAAGTTCCTGCTAAGCCGGGTATTAAAA-3'	5'-TTCCTTCTCACCGGTAGG-3'
G527A	5'-GGTAAGCCGGCTATTAAAAACC-3'	5'-AGGAACTTCTTCCTTCTC-3'
V550P	5'-ATACGGTCCTCCTAAAGGCGAC-3'	5'-TTGGTAACGCTGTCAACAG-3'
V556P	5'-CGACAGTATTCCGGAAAAAGAGGAAATC-3'	5'-CCTTTAACAGGACCGTATTTG-3'
P540A	5'-TGTGGTTCGTGCGCCTGTTGA-3'	5'-TCGCCTGTCTCAGGGTTTTTAATAC-3'
P549A	5'-CAAATACGGTGCTGTTAAAGGC-3'	5'-GTAACGCTGTCAACAGGC-3'
P562A	5'-AGAGGAAATCGCGTTCGAAAAAG-3'	5'-TTTTCCACAATACTGTCG-3'
V580P	5'-CACCGAAAAACCGACCCGTGAGGGTC-3'	5'-CCCGGTGCTAAATCAGGA-3'
G584A	5'-ACCCGTGAGGCTCAAAAAGGTG-3'	5'-CACTTTTTCGGTCCCCGG-3'
G587A	5'-GGTCAAAAAGCTGAGAAGACCATTAC-3'	5'-CTCACGGGTCACTTTTTC-3'
E588K	5'-TCAAAAAGGTAAGAAGACCATTACAAC-3'	5'-CCCTCACGGGTCACTTTT-3'
K589E	5'-AAAAGGTGAGGAGACCATTACAACC-3'	5'-TGACCCTCACGGGTCACT-3'
N598A	5'-TACACTGAAAGCCCCGCTGACC-3'	5'-GGGGTTGTAATGGTCTTC-3'
T601A	5'-AAACCCGCTGGCCGGCGAGAT-3'	5'-TTCAGTGTAGGGGTTGTAATGGTCTTCTC-3'
E624K	5'-CGAACTGACCAAATACGGCCC-3'	5'-TTAATAGGGTCCTTTGTGATCTC-3'
E-G5^2 Δ*BsaI*	5'-ACGGCCCGGAAACCTGTTGTT-3'	5'-ATTCGGTCAGTTCGTTAATAGGGTCC-3'

Table S9. Primers for mutagenesis of E-G5^2 cassette. Substituted bases in red text. *I502P was created during amplification of cassettes using primers found in **Table S11**.

Component	Forward Sequence	Reverse Sequence
pET14b destination vector	5'-GGTACT**GGTCTC**G*GACG*AAAGGAAGCTGAGTTGGC-3'	5'-GGTACT**GGTCTC**G*GACG*AAAGGAAGCTGAGTTGGCTG-3'
Cassette 1	5'-GGTACT**GGTCTC**G*ATCA*TCATCACAGCAGCGGCCCGGAAACCATCGCACCGGGCCATC-3'	5'-GGTACT**GGTCTC**G*CCCA*CGCTCAGGGTTTCCGGGCCGTATTC-3'
Cassette 2	5'-GGTACT**GGTCTC**G*TGGG*CGCGACCATTGGCCCGGAAACCATCGCACCGGGCCATC-3'	5'-GGTACT**GGTCTC**G*CAGA*CCAATAACGGTGGTTTCCGGGCCGTATTC-3'
Cassette 3	5'-GGTACT**GGTCTC**C*TCTG*GCGAGCGGCCCGGAAACCATCGCACCGGGCCATC-3'	5'-GGTACT**GGTCTC**C*GTGC*CGCTCAGCGCGGTTTCCGGGCCGTATTC-3'
Cassette 4	5'-GGTACT**GGTCTC**C*GCAC*CATTGTGGGCCCGGAAACCATCGCACCGGGCCATC-3'	5'-GGTACT**GGTCTC**C*GGCT*ACCGGTAATAACGGTTTCCGGGCCGTATTC-3'
Cassette 5	5'-GGTACT**GGTCTC**G*AGCC*TGGCGGGCCCGGAAACCATCGCACCGGGCCATC-3'	5'-GGTACT**GGTCTC**G*CGTC*AACAACAGGTTTCCGGGC-3'

Table S10. Primers for the creation of the linearised pET14b$\Delta bsaI$ destination vector and E-G5[2] cassettes for subsequent Golden Gate assembly. Bases in bold represent the BsaI recognition site. Underlined and italicised are the four bases, which become complementary overhangs post BsaI digestion.

Component	Forward Sequence
I502P Cassette 1	5'-GGTACT**GGTCTC**G*ATCA*TCATCACAGCAGCGGCCCGGAAACCCCCGCACCGGGCCATC-3'
I502P Cassette 2	5'-GGTACT**GGTCTC**G*TGGG*CGCGACCATTGGCCCGGAAACCCCCGCACCGGGCCATC-3'
I502P Cassette 3	5'-GGTACT**GGTCTC**C*TCTG*GCGAGCGGCCCGGAAACCCCCGCACCGGGCCATC-3'
I502P Cassette 4	5'-GGTACT**GGTCTC**C*GCAC*CATTGTGGGCCCGGAAACCCCCGCACCGGGCCATC-3'
I502P Cassette 5	5'-GGTACT**GGTCTC**G*AGCC*TGGCGGGCCCGGAAACCCCCGCACCGGGCCATC-3'

Table S11. Forward primers utilised in the creation of I502P cassettes for subsequent Golden Gate assembly. Bases in bold represent the BsaI recognition site. Underlined and italicised are the four bases, which become complementary overhangs post BsaI digestion. Red text displays the bases changed to substitute isoleucine for proline.

Construct	Process	Forward Sequence	Reverse Sequence
H_6-MBP-TEV-SasG	pMAL-c5X-TEV linearisation and amplification (destination vector)	5'-GAATTCCCTGCAGGTAATTAAATAAGC-3'	5'-GGATCCCTGAAAGTACAGGTTT-3'

SasG GA-ready insert	5'- *GAAAACCTGTACTTTCA* *GGGATCC*GCACCTAA GACCATCACCGAG-3'	5'- *GCTTATTTAATTACCTGCA* *GGGAATTC TTAGCAGCA* GGTCTCCGGACCATACT CG-3'

Table S12. DNA primers utilised during Q5® PCR for pMAL-c5X-MBP-TEV-SasG (native SasG) construct creation. Complementary stretches for annealing during GA are italicised.

Variant	Retraction velocity (nms^{-1})	n	Modal Rupture Force (pN)	SD (pN)
V522P				
	200	87	402.6	± 51.1
	5000	114	459.7	± 58.6
V550P				
	200	49	408.4	± 59.3
	5000	149	454.1	± 62.9

Table S13. Table displaying the number of events, modal rupture force of 1 repeat of $G5^2$ unfolding in V522P and V550P at retraction velocities of 200 and 5000 nms^{-1}.Carried out in phosphate buffered saline, pH 7.4 at room temperature.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	44	240.7	219.0	144.8	148.2
	154	211.2		148.7	
	50	205.2		151.0	
800	99	241.0	230.4	143.9	149.1
	129	227.2		146.7	
	57	223.0		156.5	
1500	90	245.4	230.5	146.3	152.2
	150	227.5		146.2	
	82	218.7		164.1	
3000	28	255.0	237.0	146.8	149.1
	91	227.6		150.8	
	56	246.3		149.7	
5000	86	267.8	254.0	143.8	148.4
	73	246.7		149.1	
	50	247.6		152.4	

Table S14. V556P E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	27	389.1	384.8	215.8	216.9
	177	382.0		216.7	
	62	383.4		218.2	
	61	413.0		214.7	

800	105	404.6	408.0	215.6	216.5
	127	406.5		219.0	
1500	86	438.8	431.0	214.4	215.5
	133	428.3		214.6	
	206	425.7		217.5	
3000	104	463.7	444.5	215.9	215.7
	134	440.7		216.5	
	78	429.1		214.7	
5000	42	469.8	447.5	209.5	213.6
	78	436.8		214.6	
	66	435.9		216.7	

Table S15. V556P G5[2] sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	202	216.8	221.5	145.7	145.4
	97	235.9		144.5	
	176	211.6		145.9	
800	176	233.2	237.3	146.5	146.6
	117	244.8		145.8	
	157	234.0		147.5	
1500	223	237.7	243.8	145.8	146.7
	124	255.8		146.2	
	130	237.8		148.0	
3000	128	242.6	250.6	148.4	147.3
	102	261.4		146.0	
	210	239.8		147.4	
5000	145	244.0	254.1	147.6	146.7
	81	273.5		144.6	
	160	244.7		148.0	

Table S16. V580P E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	125	353.8	357.8	214.3	215.9
	62	376.0		217.3	
	112	343.5		216.1	
800	76	403.9	397.6	214.3	214.5
	73	397.4		214.0	
	101	391.5		215.1	
	131	400.9		212.9	

1500	77 91	423.7 407.7	410.8	214.1 215.8	214.3
3000	87 68 163	418.1 444.8 418.9	427.3	213.4 214.4 215.4	214.4
5000	87 70 128	430.4 458.1 427.3	438.6	214.9 216.0 214.8	215.2

Table S17. V580P $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	Repeat	FWHM (pN)	Average (pN)	pWT $(E\text{-}G5^2)_5$ $G5^2$ FWHM* (pN)
200	1 2 3	101.1 122.7 113.8	112.5	56.2
800	1 2 3	52.0 85.5 64.8	67.5	58.3
1500	1 2 3	78.3 60.1 74.4	71.0	54.1
3000	1 2 3	70.7 90.9 66.5	76.0	60.5
5000	1 2 3	86.4 82.3 87.4	85.4	58.6

Table S18. FWHM values for V580P $G5^2$ sub domain mechanical unfolding in phosphate buffered saline, pH 7.4 at room temperature. FWHM from the gaussian fits of the triplicate repeats and average is the mean of these. SD: standard deviation (sample). The values were generally higher than for pWT $(E\text{-}G5^2)_5$, which may indicate deviation in the kinetic parameters of mechanical unfolding.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	110 42 56	188.1 176.0 184.8	183.0	144.5 148.9 142.6	145.3
800	103 75 113	206.7 204.6 192.5	201.3	143.3 148.8 146.6	146.2
1500	155 148 95	212.8 197.1 211.7	207.2	145.6 147.1 143.6	145.5

3000	115 135 83	214.2 196.9 204.1	200.5	146.3 146.4 146.5	146.4
5000	74 96 94	224.7 211.3 211.0	215.7	146.8 150.3 146.1	147.7

Table S19. G524A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	70 67 20	388.0 402.2 400.7	397.0	217.4 215.5 213.3	215.4
800	66 107 72	413.0 427.9 417.5	419.5	214.1 215.7 213.8	214.5
1500	127 147 55	427.9 422.9 433.7	428.2	216.3 215.3 213.5	215.0
3000	77 140 58	431.9 431.4 442.2	435.2	216.2 215.1 213.7	215.0
5000	45 110 66	445.2 443.9 453.8	447.6	216.8 215.1 214.9	215.6

Table S20. G524A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	79 67 38	175.0 165.7 170.4	170.4	145.9 143.8 144.6	144.8
800	115 64 119	197.1 183.4 198.0	192.8	143.0 143.3 150.5	145.6
1500	167 127 69	205.7 190.1 195.0	196.9	145.0 144.8 147.5	145.8
	152	210.0		145.1	

3000	129	199.9	209.3	143.3	146.3
	62	218.1		150.4	
5000	123	214.8	207.4	146.6	147.0
	123	204.3		146.0	
	67	202.9		148.5	

Table S21. G527A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	45	420.9	410.2	215.4	216.4
	46	403.3		216.2	
	47	406.4		217.7	
800	37	442.2	425.9	212.6	214.5
	41	412.0		216.0	
	108	423.4		215.0	
1500	58	459.9	440.9	214.6	214.6
	62	436.3		215.1	
	71	426.5		214.2	
3000	40	476.3	458.5	214.6	214.7
	33	448.6		213.9	
	46	450.5		215.7	
5000	34	478.0	457.0	213.3	213.5
	65	444.2		212.7	
	58	448.8		214.3	

Table S22. G527A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	127	221.9	226.0	144.1	144.8
	128	231.6		145.5	
	187	224.4		144.6	
800	144	235.2	240.2	145.5	145.2
	117	242.5		146.1	
	145	242.9		143.9	
1500	162	242.7	242.6	146.7	146.1
	174	248.2		145.6	
	145	236.9		146.1	
3000	134	249.1	251.5	146.5	145.4
	119	250.8		146.2	

	151	252.3		143.5	
	175	255.2		145.5	
5000	94	254.7	254.4	148.5	146.7
	109	253.4		146.1	

Table S23. G584A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
	61	344.4		215.7	
200	105	358.8	355.5	214.9	215.1
	143	363.3		214.5	
	68	373.4		215.9	
800	109	386.2	384.0	214.7	215.0
	100	392.3		214.3	
	133	388.4		215.3	
1500	129	381.8	389.0	216.5	215.9
	86	396.8		215.8	
	56	398.6		215.7	
3000	95	398.0	400.1	214.8	214.9
	51	403.7		214.3	
	93	404.0		215.6	
5000	98	407.4	407.1	217.4	215.5
	66	409.9		213.6	

Table S24. G584A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
	53	216.3		146.2	
200	81	218.2	220.0	144.5	143.8
	97	225.6		140.7	
	77	215.7		146.9	
800	62	230.9	225.2	145.1	145.5
	111	229.0		144.5	
	43	227.9		145.0	
1500	94	234.4	236.6	144.7	144.3
	73	247.4		143.4	
	57	230.9		144.8	
3000	72	243.8	241.8	147.9	145.6
	68	250.8		144.2	
	59	228.5		144.7	
5000	78	246.3	244.8	144.1	144.2
	115	259.6		143.8	

Table S25. G587A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	44	300.2	299.6	216.4	215.5
	62	290.5		214.7	
	74	308.1		215.5	
800	54	300.5	316.7	210.9	214.0
	44	313.7		217.0	
	79	335.9		214.3	
1500	26	321.5	331.9	212.9	212.6
	55	328.1		212.9	
	48	345.9		212.2	
3000	37	318.5	333.8	217.6	214.1
	53	329.9		211.9	
	48	353.0		212.8	
5000	53	327.6	339.7	214.6	214.6
	57	328.4		215.0	
	58	363.0		214.1	

Table S26. G587A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	45	215.3	220.4	148.6	146.5
	47	221.3		145.0	
	65	224.6		145.8	
800	83	224.6	234.6	147.6	147.8
	56	237.4		149.8	
	91	241.8		146.0	
1500	89	247.3	244.6	153.4	150.4
	76	234.9		152.1	
	133	251.4		145.6	
3000	74	229.6	247.1	147.9	147.6
	73	249.5		147.4	
	131	244.7		147.4	
5000	86	246.7	251.5	147.0	147.1
	108	254.5		147.5	
	197	253.2		146.6	

Table S27. P540A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	49	410.0	408.3	216.1	216.3
	21	411.4		215.6	
	55	403.4		217.2	
800	78	416.7	424.5	216.0	214.9
	49	426.9		214.1	
	49	429.9		214.7	
1500	114	434.9	439.0	217.0	216.2
	83	435.8		216.1	
	99	446.4		215.4	
3000	70	428.2	438.3	214.9	215.2
	57	440.7		215.3	
	84	445.9		215.6	
5000	87	448.1	454.2	215.5	214.6
	56	458.5		213.4	
	147	456.0		214.8	

Table S28. P540A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	95	223.3	228.4	144.4	144.6
	96	229.6		143.0	
	56	232.4		146.5	
800	168	242.8	237.2	143.3	146.9
	142	234.4		145.8	
	73	234.5		151.5	
1500	144	239.1	240.4	147.7	147.4
	107	240.5		145.5	
	92	241.5		149.0	
3000	197	251.3	244.6	146.8	147.4
	171	245.4		144.9	
	92	243.8		150.4	
5000	142	258.0	255.4	147.6	147.3
	158	254.2		144.8	
	72	254.0		149.5	

Table S29. P562A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	47	365.5	365.3	216.3	216.3
	46	361.1		215.7	
	45	369.2		216.9	
800	84	398.6	389.2	214.1	215.0
	66	384.1		213.6	
	50	384.9		217.3	
1500	101	398.7	396.8	215.4	215.3
	60	395.3		214.3	
	63	396.5		216.2	
3000	73	407.5	405.3	215.1	214.6
	60	404.6		213.3	
	63	403.8		215.4	
5000	55	411.0	414.9	211.9	213.4
	57	415.3		213.2	
	45	418.3		215.1	

Table S30. P562A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	89	192.9	196.6	148.8	146.8
	78	193.9		143.9	
	41	202.9		147.7	
800	87	209.8	206.7	145.7	147.0
	53	200.2		147.5	
	69	210.0		147.8	
1500	64	227.2	216.5	143.3	145.4
	83	207.5		146.7	
	70	214.8		146.1	
3000	81	221.8	210.8	147.4	146.0
	67	212.0		143.7	
	69	209.6		146.8	
5000	100	235.5	228.7	149.6	147.6
	81	221.0		142.6	
	46	229.4		150.6	

Table S31. G517A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	105	391.2	405.5	216.0	214.9
	33	415.3		214.9	
	33	409.9		213.9	
800	66	422.4	415.4	215.2	215.1
	29	400.5		214.6	
	63	423.4		215.3	
1500	70	440.3	426.1	214.9	214.5
	58	408.7		214.6	
	82	429.2		214.1	
3000	69	435.9	436.0	216.0	215.2
	40	431.2		215.3	
	66	440.8		214.1	
5000	90	445.9	446.6	215.8	215.2
	55	439.2		215.4	
	50	454.6		214.4	

Table S32. G517A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	69	185.2	190.2	145.0	148.3
	64	202.5		145.6	
	60	183.0		154.4	
800	131	189.5	200.8	148.7	148.0
	143	212.9		145.6	
	82	200.2		149.7	
1500	61	196.9	209.5	146.9	147.1
	156	220.5		146.9	
	219	211.2		147.4	
3000	90	208.2	215.1	148.4	149.4
	120	226.0		146.1	
	71	211.0		153.8	
5000	71	219.7	223.8	147.1	150.1
	64	235.4		146.9	
	44	216.1		156.5	

Table S33. P549A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	20	411.6	411.1	213.9	219.3
	41	427.5		214.1	
	82	394.1		229.9	
800	70	422.2	426.9	214.9	215.2
	76	442.1		215.1	
	81	416.3		215.7	
1500	70	431.6	443.4	214.8	215.3
	78	460.3		214.3	
	136	438.1		216.7	
3000	38	452.2	450.4	215.4	218.3
	97	463.4		215.0	
	76	435.6		224.5	
5000	43	452.7	456.8	214.6	218.1
	54	484.2		214.8	
	88	433.5		224.9	

Table S34. P549A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	74	191.5	184.3	144.7	146.3
	47	181.4		147.3	
	43	179.8		146.9	
800	42	203.5	201.4	150.4	148.3
	101	199.8		148.2	
	82	200.9		146.1	
1500	24	218.8	216.0	141.7	144.9
	64	216.8		145.5	
	71	212.3		147.5	
3000	78	225.0	223.6	146.5	146.4
	73	224.7		147.0	
	53	222.4		145.9	
5000	68	246.4	238.2	148.4	147.5
	116	240.3		147.8	
	53	227.8		146.3	

Table S35. N598A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	126	404.6	391.6	214.9	214.4
	65	400.3		213.8	
	50	370.1		214.4	
800	89	411.4	405.9	214.8	214.0
	96	402.6		214.3	
	82	403.5		213.0	
1500	73	435.1	423.4	216.1	215.1
	60	418.9		213.2	
	97	416.3		215.9	
3000	129	432.5	429.7	214.6	214.5
	61	435.8		213.6	
	53	420.8		215.4	
5000	82	448.2	435.5	215.2	213.8
	84	435.4		213.3	
	62	422.9		213.0	

Table S36. N598A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	80	232.6	222.5	144.5	146.8
	61	228.7		150.4	
	48	206.2		145.6	
800	98	239.1	240.8	145.9	145.7
	161	242.5		146.3	
	100	240.9		145.0	
1500	105	249.3	244.6	147.8	145.9
	97	248.8		144.9	
	111	235.6		145.1	
3000	59	263.2	253.3	145.4	146.7
	107	256.8		149.1	
	81	239.9		145.5	
5000	84	259.2	256.9	146.6	147.2
	80	266.7		148.2	
	75	244.6		146.7	

Table S37. T601A E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	32	405.8	393.4	214.1	216.6
	54	398.2		216.5	
	39	376.3		219.1	
800	57	419.9	416.3	214.8	215.1
	132	421.1		214.4	
	59	407.9		216.1	
1500	75	429.8	420.0	213.9	213.8
	59	424.4		214.3	
	57	405.7		213.2	
3000	46	444.3	437.6	211.5	213.5
	58	449.5		214.4	
	67	418.9		214.8	
5000	49	447.1	443.2	213.5	213.2
	47	452.6		212.4	
	53	430.1		213.8	

Table S38. T601A $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	70	212.3	218.4	146.2	152.5
	21	213.8		158.9	
	20	229.1		152.4	
800	59	220.4	230.7	146.1	146.6
	49	231.6		147.5	
	86	240.0		146.1	
1500	117	231.7	242.6	148.6	148.8
	64	247.2		146.2	
	40	249.0		151.6	
3000	66	237.7	241.9	148.2	147.5
	32	234.6		145.4	
	41	249.3		149.0	
5000	41	252.1	256.1	144.0	151.5
	29	254.2		157.5	
	48	262.0		152.9	

Table S39. E588K E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	89	408.0	403.5	216.6	218.6
	49	407.3		217.7	
	52	395.3		221.7	
800	60	410.5	417.9	212.4	214.5
	48	425.6		215.5	
	58	417.6		215.5	
1500	133	437.9	431.9	215.2	215.7
	60	426.2		214.7	
	53	431.5		217.2	
3000	68	447.9	444.3	214.1	216.5
	31	438.0		215.4	
	52	446.9		219.9	
5000	45	457.7	456.2	218.3	218.6
	62	465.5		218.2	
	79	445.5		219.4	

Table S40. E588K $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	95	221.2	225.8	146.5	145.7
	133	241.7		145.4	
	118	214.6		145.2	
800	130	234.9	236.5	145.2	144.6
	127	240.5		145.1	
	145	234.2		143.6	
1500	118	240.0	242.3	144.1	143.9
	112	247.4		144.2	
	152	239.6		143.4	
3000	119	241.5	245.3	145.2	145.9
	120	254.8		147.6	
	88	235.8		144.8	
5000	120	253.8	255.5	145.1	145.7
	104	259.2		145.4	
	106	253.4		146.6	

Table S41. K589E E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)

200	96 125 102	308.3 346.9 306.7	320.6	216.8 212.4 215.0	214.7
800	51 48 76	334.1 357.9 333.4	341.8	212.4 211.4 214.1	212.6
1500	71 103 94	348.7 374.6 350.0	357.7	214.2 213.3 213.7	213.8
3000	47 58 45	367.0 372.5 353.2	364.2	215.3 210.0 215.5	213.6
5000	88 54 67	370.5 378.4 366.1	371.7	213.8 213.3 214.5	213.9

Table S42. K589E G5[2] sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
200	67 166 142	221.8 228.7 224.9	225.1	142.7 143.0 143.5	143.1
800	141 235 247	231.4 242.2 235.8	236.5	143.1 143.6 143.4	143.4
1500	170 178 195	234.8 240.6 242.4	239.3	144.3 144.3 144.7	144.4
3000	103 117 191	232.1 247.5 241.3	244.4	144.6 145.9 145.0	145.2
5000	83 170 145	241.2 257.2 257.5	252.0	143.9 144.6 144.4	144.3

Table S43. E624K E sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

Speed (nms^{-1})	n	Mode Rupture Force (pN)	Average (pN)	Mode ΔL_C (Å)	Average (Å)
	47	264.5		214.9	

200	101 108	268.9 280.6	271.3	217.0 216.6	216.2
800	96 145 175	315.2 323.1 303.9	314.0	216.8 218.5 216.5	217.3
1500	93 101 134	341.2 322.2 333.5	332.3	215.8 216.1 216.6	216.2
3000	65 84 110	343.6 361.0 350.8	351.8	216.9 216.8 216.4	216.7
5000	40 90 92	357.9 373.5 375.3	368.9	214.2 216.2 214.3	214.9

Table S44. E624K $G5^2$ sub domain mechanical unfolding statistics. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode is obtained from the gaussian fits to the histograms. Average is the mean of the mode values at each speed.

	E624K – Rupture Force $G5^2$			pWT $(E\text{-}G5^2)_5$ $G5^2$
Speed (nms^{-1})	Repeat	FWHM (pN)	Average (pN)	FWHM (pN)
200	1 2 3	97.1 115.7 126.1	113.0	56.2
800	1 2 3	119.7 105.2 127.4	117.4	58.3
1500	1 2 3	100.8 122.2 102.2	108.4	54.1
3000	1 2 3	126.5 133.8 107.0	122.4	60.5
5000	1 2 3	126.8 122.9 113.9	121.2	58.6

Table S45. Summary of rupture force FWHM statistics for E624K vs pWT $(E\text{-}G5^2)_5$ $G5^2$ sub domain mechanical unfolding. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode FWHM is obtained from the gaussian fits to the rupture force histograms. Average is the mean of the mode values at each speed. The FWHM is significantly larger for E624K than pWT $(E\text{-}G5^2)_5$ at every retraction velocity.

	E624K - ΔL_C $G5^2$			pWT $(E\text{-}G5^2)_5$ $G5^2$
Speed (nms^{-1})	Repeat	FWHM (Å)	Average (Å)	Average FWHM (Å)

200	1 2 3	39.2 31.8 29.9	33.6	12.3
800	1 2 3	22.5 29.1 31.5	27.7	14.9
1500	1 2 3	20.7 29.5 26.0	25.4	14.4
3000	1 2 3	18.7 23.4 30.5	24.2	14.3
5000	1 2 3	31.3 25.7 29.6	28.9	16.8

Table S46. Summary of ΔL_C FWHM statistics for E624K vs pWT $(E\text{-}G5^2)_5$ $G5^2$ sub domain mechanical unfolding. Carried out in phosphate buffered saline, pH 7.4 at room temperature. n is the number of unfolding events. Mode FWHM is obtained from the gaussian fits to the ΔL_C histograms. Average is the mean of the mode values at each speed. The FWHM is significantly larger for E624K than pWT $(E\text{-}G5^2)_5$ at every retraction velocity.

Native SasG gene construction and protein over-expression:

A gene encoding the *E. coli* optimised native SasG sequence, a gift from Professor Jennifer Potts, University of Sydney, Australia) was inserted into a linearised pMAL-c5X-MBP-TEV vector (utilised in our lab[7]) using Gibson Assembly®[8] to form an MBP-TEV-SasG fusion to address poor yields of SasG production. Briefly, PCR-based (using Q5® DNA polymerase (NEB, Hertfordshire, UK)) amplification and linearisation (using primers outlined in Supplementary Table S12) for the SasG gene and pMAL-c5X-MBP-TEV destination vector, respectively, was performed. The purified SasG gene cassette and PCR-linearised pMAL-c5X-MBP-TEV were then assembled using the Gibson Assembly assembly kit (NEB, Hertfordshire, UK) following the manufacturers protocol, using a 5:1 molar ratio of SasG cassette to linearised destination vector. The reaction mixture was incubated at 50 °C for 20 minutes. 2 µl of the product was transformed into NEB® 5-α competent *E. coli* cells (NEB, Hertfordshire, UK). The sequence of each assembled variant was confirmed by Sanger Sequencing (Twist Bioscience, Cambridge, UK). The final construct comprises an N-terminal his tagged MBP, followed by a TEV cleavage, the SasG gene and C-terminally capped with two cysteine residues for protein immobilisation to gold substrate for SMFS experiments.

HHHHHH-MBP-TEV-SasG-CC

Plasmid encoding the MBP-TEV-SasG construct was transformed into *E. coli* BL21 (DE3) (Agilent, California, USA) and expressed in an identical fashion to pWT $(E\text{-}G5^2)_5$ (except of the omission of chloramphenicol). After 24 hours, the cells were harvested and the cell pellet resuspended in lysis buffer (20 mM Na_2HPO_4/ NaH_2PO_4, 8 M urea, 500 mM NaCl, 5 mM imidazole, 2 mM DTT, 1mM PMSF, 2 mM benzamidine and a rice grain of lysozyme and DNaseI). After cell disruption by sonication the cleared lysate was applied to a lab-packed 200 ml pre-charged Ni Sepharose™ Fast Flow column (Cytiva, Massachusetts, USA). Wash/refolding buffer (20 mM Na_2HPO_4-NaH_2PO_4, 500 mM NaCl, 10 mM imidazole, 2 mM DTT, 1mM PMSF, 2 mM benzamidine) was applied until baseline A_{280} was achieved and the protein eluted using wash buffer supplemented with 50 mM imidazole. The protein of interest was collected, concentrated and diluted to give a final imidazole concentration of 5 mM. A 1 ml HisTrap™ FF column (Cytiva, MA, USA) was pre-equilibrated with the wash/refolding buffer and 20 mg of lab-made H_6-TEV protease was bound to it. Once washed with the wash/refolding buffer, the construct was loaded and cycled through the TEV-bound column for 15 hours, 25 °C at a slow flowrate. The flow through containing the protein without H_6-MBP-TEV (SasG) was collected and concentrated prior to dilution in 20 mM Tris-HCl, 1 mM EDTA and 2 mM DTT to give a final NaCl concentration of 50 mM. Subsequently, this was loaded onto 2 × 5 ml HiTrap SP HP column (Cytiva) stacked atop of 4 × 5 ml HiTrap Q HP columns (Cytiva) pre-equilibrated with AEX wash buffer (20 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA and 2 mM DTT). The columns were then washed until baseline A_{280} was reached. Following removal of the cation exchange columns, the protein was eluted at a 50-185 mM NaCl gradient over 15 column volumes. Subsequent purification was performed with gel filtration using a 320 ml Hiload Superdex™ 26/600 75 column (Cytiva) in phosphate buffered saline, pH 7.4. Once centrifuged at 16 ×g for 15 minutes to remove any insoluble material or aggregates, further purification was performed by gel filtration using a 320 ml Hiload Superdex™ 26/600 75 column (Cytiva) pre-

equilibrated with phosphate buffered saline, pH 7.4. Successful isolation was confirmed by LC-MS (Supplementary Table S6). The recombinant protein was snap frozen in liquid nitrogen and stored at -80 °C. The final protein sequence of SasG is found in Supplementary Information.

Mechanical ϕ-value analysis:

The theoretical basis for calculating mechanical ϕ-values is outlined by Best and colleagues [9], and has been verified using simulations [10,11]. Mechanical ϕ-value analysis is distinct from solution studies as mechanical unfolding studies are vectorial (i.e. single-molecule force spectroscopy measures the kinetic stability of a protein region proximal to the extension points). Briefly, equilibrium denaturation is utilised to measure the effects of the mutation on the ground state (i.e. under no external force), and then the barrier to forced denaturation (i.e. the activation energy of unfolding) is measured through mechanical unfolding experiments.

Mechanical unfolding is carried out at multiple speeds to give rise to a "dynamic force spectrum" (unfolding force versus log retraction velocity (ln *v*), as exemplified in Figure 3 and Supplementary Figures S4, S13, S15, and S18). Mechanical unfolding experiments are typically kinetically described by a two-state model: a folded and unfolded state separated by a single transition barrier. The exponential increase in the unfolding rate constant in the presence of force (k_u^F) can be described using the analytical equation derived by Evans and Ritchie [12]:

$$k_u^F = Aexp^{-\frac{(\Delta G_{TS}^F - Fx_u)}{(k_BT)}} = k_u^0 exp^{\frac{Fx_u}{k_BT}}$$

As the unfolding rate is assumed to depend on force, we are able to solve the most probable unfolding force of a protein pulled at constant velocity. The relationship between unfolding force and loading rate depends on x_u (described as the distance from the native well to the TS barrier). This is described by (the Bell-Evans-Ritchie model, shown below), where the rupture force (F_R) is proportional to the logarithm of loading rate (R)[12,13]:

$$F_R = \left(\frac{k_BT}{x_u}\right)\ln\left(\frac{Rx_u}{k_u^0 k_BT}\right)$$

k_u^0 is the mechanical unfolding rate constant under zero force.

If the mutation results in an unchanged x_u (i.e. the transition state position is not altered by mutation) it is possible to measure the change in activation energy from comparing the mechanical unfolding kinetics of the WT ($k_u^{0,WT}$) versus the mutant ($k_u^{0,MUT}$) mechanical activation energy at zero force. When x_u values are unchanged between WT and mutant, the log ratio of rate constants is observed as:

$$ln\left(\frac{k_u^{0,WT}}{k_u^{0,MUT}}\right) = ln\left(\frac{v^{WT}}{v^{MUT}}\right) + ln\left(\frac{C(f,x_u,T)^{WT}}{C(f,x_u,T)^{MUT}}\right)$$

If the force spectra of proteins have similar slopes (*m*, on the dynamic force spectrum), they will have near identical x_u values. Because function $C(f,x_u,T)$ depends on the elastic properties of the protein (x_u and T, will be identical between WT and mutant), the change in activation free energy at any force, *f*, can be calculated from:

$$\Delta\Delta G_{TS-N}^{WT-MUT} = -RT\ln(v_{WT} - v_{MUT})$$

As force spectra of proteins with similar x_u values will have similar slopes (*m*, on the dynamic force spectrum), the mechanical activation free energy at any force at a given retraction velocity, can be calculated:

$$\Delta\Delta G_{TS-N}^{WT-MUT} = -RT(f_{WT} - f_{MUT})/m$$

DNA and protein sequences

DNA encoding H_6 displayed in green, start codon in dark blue, stop codon in red, linkers in orange and C-terminal double cysteine in light blue.

pWT (E-G5^2)$_5$:

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCCGGAAACCATCGCACCGGGCCATCG
TGATGAGTTTGATCCGAAGTTACCTACCGGTGAGAAGGAAGAAGTTCCTGGTAAGCCGGGTATTA
AAAACCCTGAGACAGGCGATGTGGTTCGTCCGCCTGTTGACAGCGTTACCAAATACGGTCCTGTTA
AAGGCGACAGTATTGTGGAAAAAGAGGAAATCCCGTTCGAAAAAGAACGCAAATTTAATCCTGAT
TTAGCACCGGGCACCGAAAAAGTGACCCGTGAGGGTCAAAAAGGTGAGAAGACCATTACAACCCC
TACACTGAAAAACCCGCTGACCGGCGAGATCATTAGCAAGGGTGAGAGTAAGGAAGAGATCACA
AAGGACCCTATTAACGAACTGACCGAATACGGCCCGGAAACCCTGAGCGTGGGCGCGACCATTGG
CCCGGAAACCATCGCACCGGGCCATCGTGATGAGTTTGATCCGAAGTTACCTACCGGTGAGAAGG
AAGAAGTTCCTGGTAAGCCGGGTATTAAAAACCCTGAGACAGGCGATGTGGTTCGTCCGCCTGTTG
ACAGCGTTACCAAATACGGTCCTGTTAAAGGCGACAGTATTGTGGAAAAAGAGGAAATCCCGTTC
GAAAAAGAACGCAAATTTAATCCTGATTTAGCACCGGGCACCGAAAAAGTGACCCGTGAGGGTCA
AAAAGGTGAGAAGACCATTACAACCCCTACACTGAAAAACCCGCTGACCGGCGAGATCATTAGCA
AGGGTGAGAGTAAGGAAGAGATCACAAAGGACCCTATTAACGAACTGACCGAATACGGCCCGGA
AACCACCGTTATTGGTCTGGCGAGCGGCCCGGAAACCATCGCACCGGGCCATCGTGATGAGTTTG
ATCCGAAGTTACCTACCGGTGAGAAGGAAGAAGTTCCTGGTAAGCCGGGTATTAAAAACCCTGAG
ACAGGCGATGTGGTTCGTCCGCCTGTTGACAGCGTTACCAAATACGGTCCTGTTAAAGGCGACAGT
ATTGTGGAAAAAGAGGAAATCCCGTTCGAAAAAGAACGCAAATTTAATCCTGATTTAGCACCGGG
CACCGAAAAAGTGACCCGTGAGGGTCAAAAAGGTGAGAAGACCATTACAACCCCTACACTGAAA
AACCCGCTGACCGGCGAGATCATTAGCAAGGGTGAGAGTAAGGAAGAGATCACAAAGGACCCTA
TTAACGAACTGACCGAATACGGCCCGGAAACCGCGCTGAGCGGCACCATTGTGGGCCCGGAAACC
ATCGCACCGGGCCATCGTGATGAGTTTGATCCGAAGTTACCTACCGGTGAGAAGGAAGAAGTTCC
TGGTAAGCCGGGTATTAAAAACCCTGAGACAGGCGATGTGGTTCGTCCGCCTGTTGACAGCGTTAC
CAAATACGGTCCTGTTAAAGGCGACAGTATTGTGGAAAAAGAGGAAATCCCGTTCGAAAAAGAAC
GCAAATTTAATCCTGATTTAGCACCGGGCACCGAAAAAGTGACCCGTGAGGGTCAAAAAGGTGAG
AAGACCATTACAACCCCTACACTGAAAAACCCGCTGACCGGCGAGATCATTAGCAAGGGTGAGAG
TAAGGAAGAGATCACAAAGGACCCTATTAACGAACTGACCGAATACGGCCCGGAAACCGTTATTA
CCGGTAGCCTGGCGGGCCCGGAAACCATCGCACCGGGCCATCGTGATGAGTTTGATCCGAAGTTA
CCTACCGGTGAGAAGGAAGAAGTTCCTGGTAAGCCGGGTATTAAAAACCCTGAGACAGGCGATGT
GGTTCGTCCGCCTGTTGACAGCGTTACCAAATACGGTCCTGTTAAAGGCGACAGTATTGTGGAAAA
AGAGGAAATCCCGTTCGAAAAAGAACGCAAATTTAATCCTGATTTAGCACCGGGCACCGAAAAAG
TGACCCGTGAGGGTCAAAAAGGTGAGAAGACCATTACAACCCCTACACTGAAAAACCCGCTGACC
GGCGAGATCATTAGCAAGGGTGAGAGTAAGGAAGAGATCACAAAGGACCCTATTAACGAACTGA
CCGAATACGGCCCGGAAACCTGTTGTTGA

The amino acid sequences below are after N-terminal methionine excision or cleavage with TEV protease (SasG). Amino acid substitutions coloured in red and linker residues in orange.

SasG:

GSAPKTITELEKKVEEIPFKKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGVIISKGEPKEEITK
DPINELTEYGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETIT
PGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKFNPDL
APGTEKVTREGQKGEKTITTPTLKNPLTGVIISKGEPKEEITKDPINELTEYGPETITPGHRDEFDPKLPTG
EKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFKKERKFNPDLAPGTEKVTREGQK
GEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETITPGHRDEFDPKLPTGEKEEVPGKPGIKNP
ETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLT
GEIISKGESKEEITKDPINELTEYGPETITPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVT
KYGPVKGDSIVEKEEIPFKKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITK
DPINELTEYGPETITPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETC
C

pWT (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGT
IVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

G517A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTAEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTAEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTAEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGT
IVGPETIAPGHRDEFDPKLPTAEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTAEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

P549A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGAVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGAVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGAVKGDSIVEKEEI
PFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSG
TIVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGAVKGDSIVEKEEIPFEK
ERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAG
PETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGAVKGDSIVEKEEIPFEKERK
FNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

N598A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKAPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKAPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKAPLTGEIISKGESKEEITKDPINELTEYGPETALSGT
IVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKGEKTITTPTLKAPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKGEKTITTPTLKAPLTGEIISKGESKEEITKDPINELTEYGPETCC

T601A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLAGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLAGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLAGEIISKGESKEEITKDPINELTEYGPETALSGT
IVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLAGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKGEKTITTPTLKNPLAGEIISKGESKEEITKDPINELTEYGPETCC

G524A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPAKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPAKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPAKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGT
IVGPETIAPGHRDEFDPKLPTGEKEEVPAKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPAKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

G527A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPAIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPAIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPAIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGT
IVGPETIAPGHRDEFDPKLPTGEKEEVPGKPAIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPGKPAIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

G584A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREAQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREAQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREAQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGT
IVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREAQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP

ETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREAQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

G587A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKAEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKAEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREGQKAEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGT
IVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKAEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKAEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

P540A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPAVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPAVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPAVDSVTKYGPVKGDSIVEKEEI
PFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSG
TIVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPAVDSVTKYGPVKGDSIVEKEEIPFEK
ERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAG
PETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPAVDSVTKYGPVKGDSIVEKEEIPFEKERK
FNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

P562A (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIAFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIAFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEI
AFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSG
TIVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIAFEK
ERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAG
PETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIAFEKERK
FNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

I502P (E-G5^2)$_5$

GSSHHHHHHSSGPETPAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETPAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETPAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEI
PFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSG
TIVGPETPAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFE
KERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLA
GPETPAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKER
KFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

V522P (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEPPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIV
EKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPE
TLSVGATIGPETIAPGHRDEFDPKLPTGEKEEPPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKE
EIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETTVI
GLASGPETIAPGHRDEFDPKLPTGEKEEPPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFE
KERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGTIV

GPETIAPGHRDEFDPKLPTGEKEEPPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERK
FNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGPETI
APGHRDEFDPKLPTGEKEEPPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKFNPD
LAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

V550P (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPPKGDSIV
EKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPE
TLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPPKGDSIVEKE
EIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETTVI
GLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPPKGDSIVEKEEIPFE
KERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGTIV
GPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPPKGDSIVEKEEIPFEKERK
FNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGPETI
APGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPPKGDSIVEKEEIPFEKERKFNPD
LAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

V556P (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIP
EKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPE
TLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIPEKE
EIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETTVI
GLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIPEKEEIPFE
KERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGTIV
GPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIPEKEEIPFEKERK
FNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGPETI
APGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIPEKEEIPFEKERKFNPD
LAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

V580P (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKPTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKPTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETTV
IGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPF
EKERKFNPDLAPGTEKPTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGTI
VGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKPTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKPTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

E588K (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGKKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGKKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREGQKGKKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGT
IVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKGKKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKGKKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

K589E (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEETITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEETITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETTV
IGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPF
EKERKFNPDLAPGTEKVTREGQKGEETITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETALSGTI
VGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKGEETITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKGEETITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETCC

E624K (E-G5^2)$_5$

GSSHHHHHHSSGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSI
VEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTKYGP
ETLSVGATIGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEK
EEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTKYGPETT
VIGLASGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIP
FEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTKYGPETALSGT
IVGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKE
RKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTKYGPETVITGSLAGP
ETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKF
NPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTKYGPETCC

Supporting information references:

[1] J.N. Milstein, J.-C. Meiners, Worm-Like Chain (WLC) Model, in: Encyclopedia of Biophysics, Springer Berlin Heidelberg, 2013: pp. 2757–2760.

[2] T. Hoffmann, K.M. Tych, M.L. Hughes, D.J. Brockwell, L. Dougan, Towards design principles for determining the mechanical stability of proteins, Phys. Chem. Chem. Phys. 15 (2013) 15767–15780.

[3] K.W. Plaxco, K.T. Simons, D. Baker, Contact order, transition state placement and the refolding rates of single domain proteins11Edited by P. E. Wright, J Mol Biol. 277 (1998) 985–994.

[4] R.Z. Kramer, J. Bella, P. Mayville, B. Brodsky, H.M. Berman, Sequence dependent conformational variations of collagen triple-helical structure, Nat. Struct. Biol. 6 (1999) 454–457.

[5] J. Bella, M. Eaton, B. Brodsky, H.M. Berman, Crystal and molecular structure, of a collagen-like peptide at 1.9 Å resolution, Science (1979). 266 (1994) 75–81.

[6] D.T. Gruszka, C.A.T.F. Mendonça, E. Paci, F. Whelan, J. Hawkhead, J.R. Potts, J. Clarke, Disorder drives cooperative folding in a multidomain protein, Proc. Natl. Acad. Sci. U.S.A. 113 (2016) 11841–11846.

[7] R. Schiffrin, Roles of periplasmic chaperones and BamA in outer membrane protein folding, University of Leeds, 2016.

[8] D.G. Gibson, L. Young, R.-Y. Chuang, J.C. Venter, C.A. Hutchison, H.O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases, Nat Methods. 6 (2009) 343–345.

[9] R.B. Best, S.B. Fowler, J.L. Toca-Herrera, J. Clarke, A simple method for probing the mechanical unfolding pathway of proteins in detail, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 12143–12148.

[10] R.B. Best, S.B. Fowler, J.L.T. Herrera, A. Steward, E. Paci, J. Clarke, Mechanical unfolding of a titin Ig domain: structure of transition state revealed by combining atomic force microscopy, protein engineering and molecular dynamics simulations., J. Mol. Biol. 330 (2003) 867–877.

[11] S.P. Ng, R.W.S. Rounsevell, A. Steward, C.D. Geierhaas, P.M. Williams, E. Paci, J. Clarke, Mechanical unfolding of TNfn3: The unfolding pathway of a fnIII domain probed by protein engineering, AFM and MD simulation, J. Mol. Biol. 350 (2005) 776–789.

[12] G.I. Bell, Models for the specific adhesion of cells to cells, Science (1979). 200 (1978) 618–627.

[13] E. Evans, K. Ritchie, Dynamic strength of molecular adhesion bonds, Biophys J. 72 (1997) 1541–1555.