This is a repository copy of *Rapid and Robust Polyprotein Production Facilitates Single-Molecule Mechanical Characterization of β-Barrel Assembly Machinery Polypeptide Transport Associated Domains.*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/91129/

Version: Supplemental Material

Article:

Rapid and Robust Polyprotein Production Facilitates Single Molecule Mechanical Characterization of BamA POTRA Domains

Toni Hoffmann[†,‡,∥,‡], Katarzyna M. Tych[†,‡,‡], Thomas Crosskey[∥], Bob Schiffrin[‡,∥], David J. Brockwell[‡,∥,*] and Lorna Dougan[†,‡,*]

[†]School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, [‡]Astbury Centre for Structural and Molecular Biology, University of Leeds, Leeds, LS2 9JT, [∥]School of Molecular and Cellular Biology, University of Leeds, Leeds, LS2 9JT

Primer	Sequence (5' -> 3') coloured regions encode peptide peptide linker	Comment
Vector_fwd I27-I_rev	GTGAAAGAACTGTGTTGTTAAACGCGTGAGGAATTTTGAAGA AATGGTCGCGCCCACGCTCAGCAGTTCTTTCACTTTCAGATT	Vector backbone
I27-III_fwd I27-III_rev	ACCGTTATTGGTCTGGCGAGCCTGATCGAAGTGGAAAAGCCT CACAATGGTGCCGCTCAGCGCCAGTTCTTTCACTTTCAGATT	Position III insert
I27-V_fwd I27-V_rev	GTTATTACCGGTAGCCTGGCGCTGATCGAAGTGGAAAAGCCT GGTCACAATGCCCAGCGCGCTCAGTTCTTTCACTTTCAGATT	Position V insert
I27-VII_fwd I27-VII_rev	ATTACCGCAGGTGTGAGCCTGCTGATCGAAGTGGAAAAGCCT TCTTCAAAATTCCTCACGCGTTTAACAACACAGTTCTTTCAC	Position VII insert
TmCsp-II_fwd TmCsp-II_rev	CTGAGCGTGGGCGCGACCATTCGTGGCAAAGTGAAATGGTTT GCTCGCCAGACCAATAACGGTTTCCACCACTTTCACATGCGC	Position II insert
TmCsp-IV_fwd TmCsp-IV_rev	GCGCTGAGCGGCACCATTGTGCGTGGCAAAGTGAAATGGTTT CGCCAGGCTACCGGTAATAACTTCCACCACTTTCACATGCGC	Position IV insert
TmCsp-VI_fwd TmCsp-VI_rev	AGCGCGCTGGGCATTGTGACCCGTGGCAAAGTGAAATGGTTT CAGGCTCACACCTGCGGTAATTTCCACCACTTTCACATGCGC	Position VI insert
EcPOTRA2-II_fwd EcPOTRA2-II_rev	CTGAGCGTGGGCGCGACCATTCCGACCATTGCCAGCATTACT GCTCGCCAGACCAATAACGGTCACACCTTCCTGGAACACCAG	Position II insert
EcPOTRA2-IV_fwd EcPOTRA2-IV_rev	GCGCTGAGCGGCACCATTGTGCCGACCATTGCCAGCATTACT CGCCAGGCTACCGGTAATAACCACACCTTCCTGGAACACCAG	Position IV insert
EcPOTRA2-VI_fwd EcPOTRA2-VI_rev	AGCGCGCTGGGCATTGTGACCCCGACCATTGCCAGCATTACT CAGGCTCACACCTGCGGTAATCACACCTTCCTGGAACACCAG	Position VI insert

Table 1. Oligonucleotide primers used to amplify Gibson-compatible cold shock protein (TmCsp), EcPOTRA2 domain and I27 inserts as well as a compatible linear vector backbone.

The DNA and amino acid sequences for the "cassette" and Gibson Assembly approaches.

(I27-TmCSP)$_3$I27$_{2C}$ "cassette" approach

ATGCATCACCATCACCATCA**CTCGAG**CCTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTG
GTGAAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGCCAGTGGAAGCTGAAAGGACAGCC
TTTGACAGCTTCCCCTGACTCTGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCATAACTCTCAGC
TGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGCTAATGCCAAATCTGCAGCCAATCTGAAAGTGAAAGA**ACT
A**GTA**GAGGCTCGA**CGCGGCAAAGTGAAATGGTTTGATAGCAAAAAAGGCTATGGCTTTATTACCAAAGATGA
AGGCGGCGATGTGTTTGTGCATTGGAGCGCGATTGAAATGGAAGGCTTTAAAACCCTGAAAGAAGGCCAGGTG
GTGGAATTTGAAATTCAGGAAGGCAAAAAAGGCCCGCAGGCGGCGCATGTGAAAGTGGTGGAA**CTTATCGAA
GCGCGC**CTAATAGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAA
ATTGAACTTTCTGAACCTGATGTTCACGGCCAGTGGAAGCTGAAAGGACAGCCTTTGACAGCTTCCCCTGACTC
TGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCATAACTCTCAGCTGGGTATGACAGGAGAGGTT
TCCTTCCAGGCTGCTAATGCCAAATCTGCAGCCAATCTGAAAGTGAAAGAATTG**CTGAGCTCGGCTCGA**CGC
GGCAAAGTGAAATGGTTTGATAGCAAAAAAGGCTATGGCTTTATTACCAAAGATGAAGGCGGCGATGTGTTTG
TGCATTGGAGCGCGATTGAAATGGAAGGCTTTAAAACCCTGAAAGAAGGCCAGGTGGTGGAATTTGAAATTCA
GGAAGGCAAAAAAGGCCCGCAGGCGGCGCATGTGAAAGTGGTGGAA**CTTATCGAAGCAC GGGCC**CTAATAG
AAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGA
ACCTGATGTTCACGGCCAGTGGAAGCTGAAAGGACAGCCTTTGACAGCTTCCCCTGACTCTGAAATCATTGAG
GATGGAAAGAAGCATATTCTGATCCTTCATAACTCTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGC
TAATGCCAAATCTGCAGCCAATCTGAAAGTGAAAGAATTG**CTCAT CCGCGG**ACGTCGCGGCAAAGTGAAATG
GTTTGATAGCAAAAAAGGCTATGGCTTTATTACCAAAGATGAAGGCGGCGATGTGTTTGTGCATTGGAGCGCG
ATTGAAATGGAAGGCTTTAAAACCCTGAAAGAAGGCCAGGTGGTGGAATTTGAAATTCAGGAAGGCAAAAAA
GGCCCGCAGGCGGCGCATGTGAAAGTGGTGGAA**TTGAT TGTACA**AGCTCGTCTAATAGAAGTGGAAAAGCCT
CTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGG
CCAGTGGAAGCTGAAAGGACAGCCTTTGACAGCTTCCCCTGACTCTGAAATCATTGAGGATGGAAAGAAGCAT
ATTCTGATCCTTCATAACTCTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGCTAATGCCAAATCTGC
AGCCAATCTGAAAGTGAAAGAATTGTGTTGTTAA**ACGCGT**

CSP DNA sequence; **I27 DNA sequence**; **linker DNA sequence**; **other sequence (start codon (His)$_6$-tag; (Cys)$_2$ stop codon)**,
restriction site

Restriction sites 5 to 3:

XhoI SpeI BssHII SacI ApaI sacII BsrGI MluI

Amino-acid sequence:

MHHHHHHSSLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQWKLKGQPLTASPDSEIIE

DGKKHILILHNSQLGMTGEVSFQAANAKSAANLKVKEL**VEAR**RGKVKWFDSKKGYGFITK

DEGGDVFVHWSAIEMEGFKTLKEGQVVEFEIQEGKKGPQAAHVKVVE**LIEAR**LIEVEKPL

YGVEVFVGETAHFEIELSEPDVHGQWKLKGQPLTASPDSEIIEDGKKHILILHNSQLGMT

GEVSFQAANAKSAANLKVKEL**LSSAR**RGKVKWFDSKKGYGFITKDEGGDVFVHWSAIEME

GFKTLKEGQVVEFEIQEGKKGPQAAHVKVVE**LIEARA**LIEVEKPLYGVEVFVGETAHFEI

ELSEPDVHGQWKLKGQPLTASPDSEIIEDGKKHILILHNSQLGMTGEVSFQAANAKSAAN

LKVKEL**LIRGR**RGKVKWFDSKKGYGFITKDEGGDVFVHWSAIEMEGFKTLKEGQVVEFEI

QEGKKGPQAAHVKVVE**LIVQAR**LIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQWKLKG

QPLTASPDSEIIEDGKKHILILHNSQLGMTGEVSFQAANAKSAANLKVKEL**CC**

(I27-TmCSP)₃I27₂C "Gibson" approach

ATGCATCACCATCACCATCACTCGAGCCTGATCGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTG
AAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGCCAGTGGAAGCTGAAAGGACAGCCTTTGAC
AGCTTCCCCTGACTCTGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCATAACTCTCAGCTGGGTATGA
CAGGAGAGGTTTCCTTCCAGGCTGCTAATGCCAAATCTGCAGCCAATCTGAAAGTGAAAGAACTGCTGAGCGTGGG
CGCGACCATTCGCGGCAAAGTGAAATGGTTTGATAGCAAAAAAGGCTATGGCTTTATTACCAAAGATGAAGGCGG
CGATGTGTTTGTGCATTGGAGCGCGATTGAAATGGAAGGCTTTAAACCCTGAAAGAAGGCCAGGTGGTGGAATTT
GAAATTCAGGAAGGCAAAAAAGGCCCGCAGGCGGCGCATGTGAAAGTGGTGGAAACCGTTATTGGTCTGGCGAGC
CTGATCGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTT
CTGAACCTGATGTTCACGGCCAGTGGAAGCTGAAAGGACAGCCTTTGACAGCTTCCCCTGACTCTGAAATCATTGA
GGATGGAAAGAAGCATATTCTGATCCTTCATAACTCTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGCT
AATGCCAAATCTGCAGCCAATCTGAAAGTGAAAGAACTGGCGCTGAGCGGCACCATTGTGCGCGGCAAAGTGAAA
TGGTTTGATAGCAAAAAAGGCTATGGCTTTATTACCAAAGATGAAGGCGGCGATGTGTTTGTGCATTGGAGCGCGA
TTGAAATGGAAGGCTTTAAACCCTGAAAGAAGGCCAGGTGGTGGAATTTGAAATTCAGGAAGGCAAAAAAGGCC
CGCAGGCGGCGCATGTGAAAGTGGTGGAAGTTATTACCGGTAGCCTGGCGCTGATCGAAGTGGAAAAGCCTCTGTA
CGGAGTAGAGGTGTTTGTTGGTGAAACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGCCAGTGG
AAGCTGAAAGGACAGCCTTTGACAGCTTCCCCTGACTCTGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCC
TTCATAACTCTCAGCTGGGTATGACAGGAGAGGTTTCCTTCCAGGCTGCTAATGCCAAATCTGCAGCCAATCTGAA
AGTGAAAGAACTGAGCGCGCTGGGCATTGTGACCCGCGGCAAAGTGAAATGGTTTGATAGCAAAAAAGGCTATGG
CTTTATTACCAAAGATGAAGGCGGCGATGTGTTTGTGCATTGGAGCGCGATTGAAATGGAAGGCTTTAAACCCTG
AAAGAAGGCCAGGTGGTGGAATTTGAAATTCAGGAAGGCAAAAAAGGCCCGCAGGCGGCGCATGTGAAAGTGGT
GGAAATTACCGCAGGTGTGAGCCTGCTGATCGAAGTGGAAAAGCCTCTGTACGGAGTAGAGGTGTTTGTTGGTGAA
ACAGCCCACTTTGAAATTGAACTTTCTGAACCTGATGTTCACGGCCAGTGGAAGCTGAAAGGACAGCCTTTGACAG
CTTCCCCTGACTCTGAAATCATTGAGGATGGAAAGAAGCATATTCTGATCCTTCATAACTCTCAGCTGGGTATGACA
GGAGAGGTTTCCTTCCAGGCTGCTAATGCCAAATCTGCAGCCAATCTGAAAGTGAAAGAACTGTGTTGTTAAACGC
GT

CSP DNA sequence; I27 DNA sequence; linker DNA sequence, other sequence (start codon (His)₆-tag; (Cys)₂ stop codon)

Amino-acid sequence:

MHHHHHHSSLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQWKLKGQPLTASPDSEIIEDGKKHILILHNSQLGMTGEVS
FQAANAKSAANLKVKELLSVGATIRGKVKWFDSKKGYGFITKDEGGDVFVHWSAIEMEGFKTLKEGQVVEFEIQEGKK
GPQAAHVKVVETVIGLASLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQWKLKGQPLTASPDSEIIEDGKKHILILHNS
QLGMTGEVSFQAANAKSAANLKVKELALSGTIVRGKVKWFDSKKGYGFITKDEGGDVFVHWSAIEMEGFKTLKEGQV
VEFEIQEGKKGPQAAHVKVVEVITGSLALIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQWKLKGQPLTASPDSEIIEDG
KKHILILHNSQLGMTGEVSFQAANAKSAANLKVKELSALGIVTRGKVKWFDSKKGYGFITKDEGGDVFVHWSAIEMEG
FKTLKEGQVVEFEIQEGKKGPQAAHVKVVEITAGVSLLIEVEKPLYGVEVFVGETAHFEIELSEPDVHGQWKLKGQPLTA
SPDSEIIEDGKKHILILHNSQLGMTGEVSFQAANAKSAANLKVKELCC

Single-molecule AFM experiments on (I27-TmCsp)$_3$-I27)^{GA}. Stretching (I27-TmCsp)$_3$-I27)^{GA} at a constant pulling speed of 600 nms^{-1} results in a sawtooth-like unfolding pattern, an example of which is shown in Figure S1A (upper trace). Each peak in the sawtooth pattern corresponds to the unfolding of a single domain of either TmCsp or I27. For comparison, an example force-extension trace is also shown from a previous study in which the chimeric polyprotein (I27-TmCsp)$_3$-I27 had been produced using the standard method (Figure S1A, lower trace). The inter-peak distance (x_{p2p}), defined as the distance from one unfolding peak to the same force value on the following curve, and the peak unfolding force (F_U) was then measured for each unfolding event and frequency histograms constructed for both x_{p2p} and F_U. The x_{p2p} frequency histogram (Figure S1C) displays a bimodal distribution centered at 18.4 (±0.8) nm and 23.4 (±0.9) nm for the original construct and 18.5 (±0.9) nm and 23.7 (±0.8) nm for the GA construct corresponding, in both constructs, to increases in contour length (ΔL_C) upon unfolding of each domain of 23.5 nm and 28.0 nm, respectively. The spacing between peaks reflects the number of amino-acids in the force-resistant native structure and thus allows each unfolding event to be assigned to a particular protein domain type (TmCsp and I27 exhibit x_{p2p} values of around 18.5 and 23.5 nm, respectively). The F_U frequency distribution also displays a bimodal distribution centered at 72 (±2) pN and 173 (±3) pN for the original construct and 75 (±3) pN and 180 (±3) pN for the Gibson assembly construct at a pulling speed of 600 nms^{-1} (Figure S1C), reflecting the different mechanical strengths of the immunoglobulin-like I27 domain and the OB-fold of TmCsp.

The distinct mechanical phenotypes of TmCsp and I27 can be observed in a scatter plot (Figure S1C, dark red and yellow symbols) that combines the data for x_{p2p} and F_U and shows two clear populations of events at a pulling speed of 600 nms^{-1}. For comparison, we also show the distribution of x_{p2p} and F_U measured in a previous study in which the chimeric polyprotein (I27-TmCsp)$_3$-I27 had been produced using the standard method (Figure S1C, light red and orange symbols). Excellent agreement can be found between for x_{p2p} and F_U measured for the constructs made using each method. We followed the same procedure to obtain force-extension traces at three other pulling velocities; 100, 200 (Figure S1C (lower)), and 2000 nms^{-1} (Figure S2 and Table S2). At each pulling velocity we completed three experiments to measure F_U for each unfolding peak and constructed three histograms of F_U. We found that the pulling speed dependence of F_U for TmCSP and I27 were the same for both constructs, (I27-TmCsp)$_3$-I27^{GA} and (I27-TmCsp)$_3$-I27 (Figure S2).
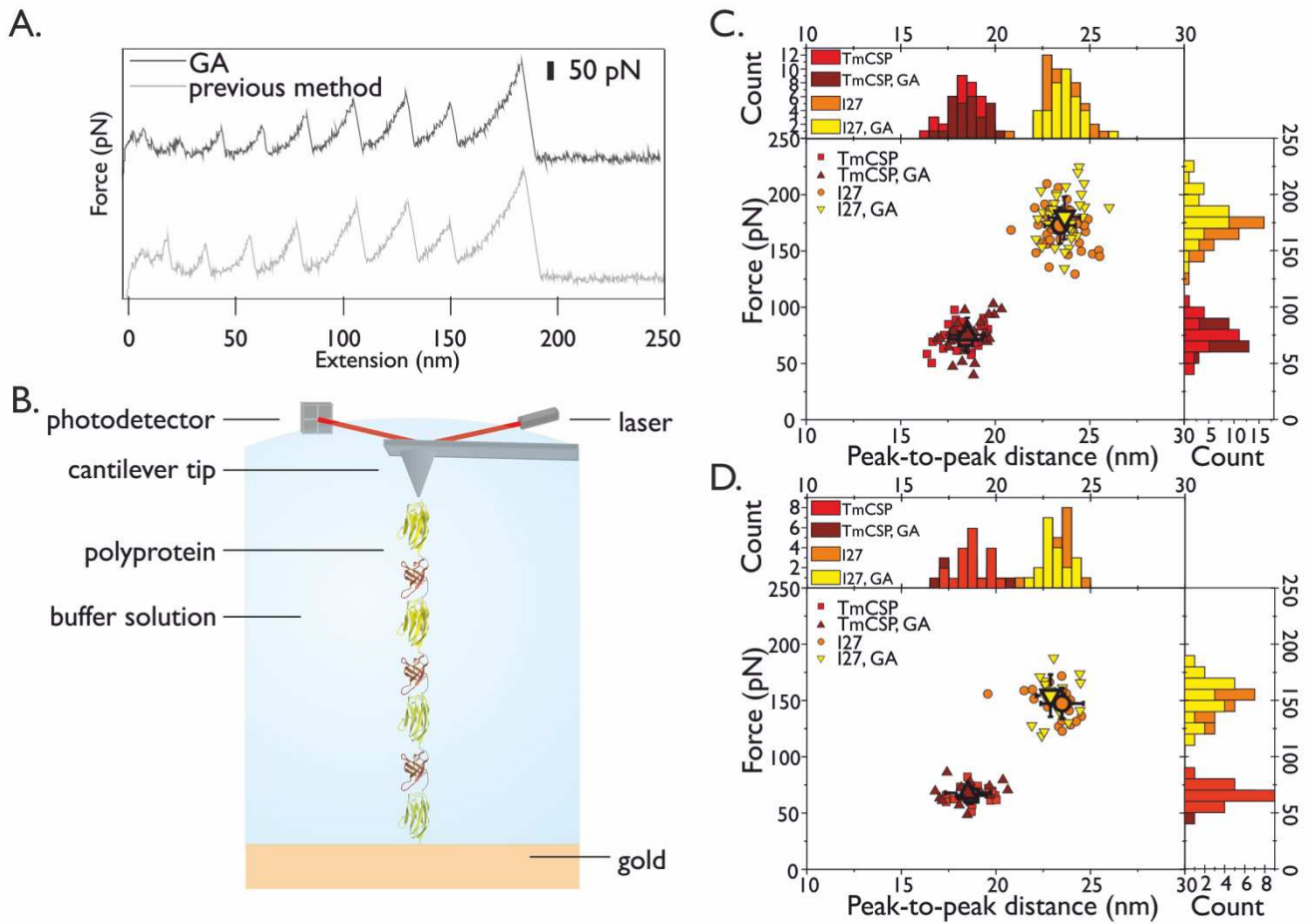
Figure S1. Mechanical unfolding of the (I27-TmCsp)$_3$-I27^{GA} polyprotein made using Gibson Assembly, compared with (I27-TmCsp)$_3$-I27 constructed using the classical cassette approach. (A) Example sawtooth force-extension profiles that result from the mechanical unfolding at 600 nms^{-1} of (I27-TmCsp)$_3$-I27^{GA} (upper trace) compared with (I27-TmCsp)$_3$-I27 (lower trace). (B) Schematic showing a single (I27-TmCsp)$_3$-I27 molecule (TmCsp and I27, red and yellow, respectively) attached to a gold surface (bottom) and the tip of an AFM cantilever (top). (C) (upper) The scatter plots of TmCsp- and I27- specific unfolding forces and inter-peak distances at 600 nms^{-1} for (I27-TmCsp)$_3$-I27^{GA} with 28 TmCsp unfolding events (dark red upward-pointing triangles) and 37 I27 unfolding events (yellow downward-pointing triangles) and for (I27-TmCsp)$_3$-I27 with 37 TmCsp unfolding events (red squares) and 51 I27 unfolding events (orange circles), are shown in combination with their respective distribution histograms. (D) The scatter plots at 200 nms^{-1} for (I27-TmCsp)$_3$-I27^{GA} with 15 TmCsp unfolding events (dark red upward-pointing triangles) and 19 I27 unfolding events (yellow downward-pointing triangles) and for (I27-TmCsp)$_3$-I27 with 19 TmCsp unfolding events (red squares) and 21 I27 unfolding events (orange circles). Error bars on the median values for each data set (large open symbols) indicate the standard deviation.
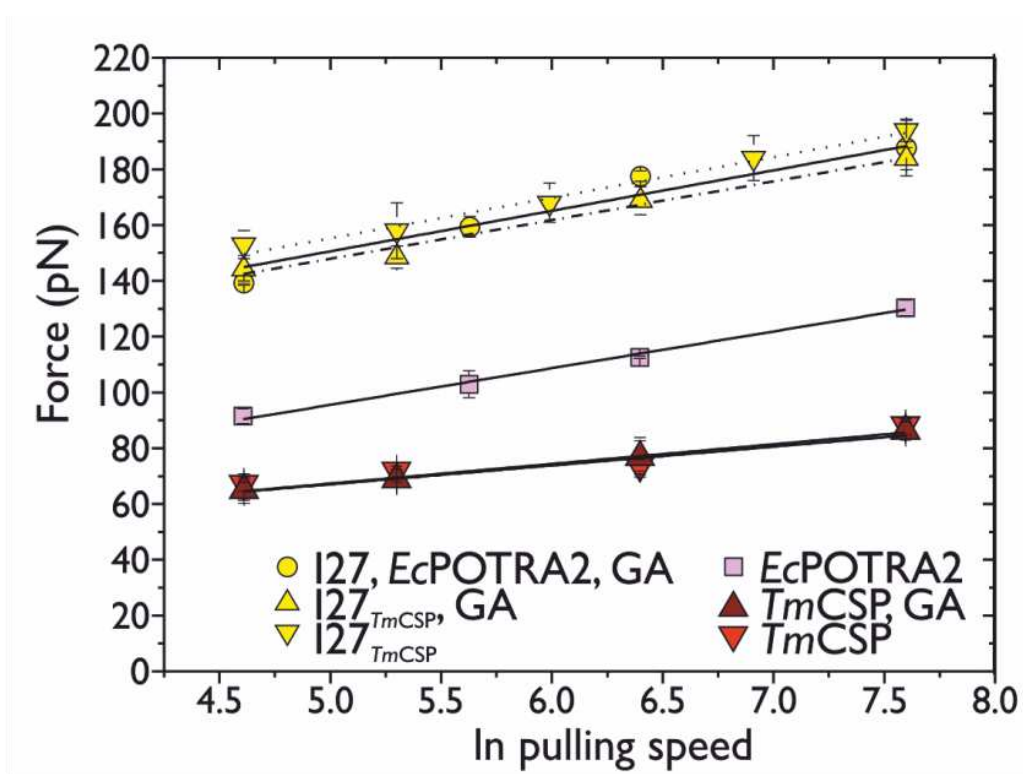
Figure S2. Comparison of the pulling speed dependence of the 3 different polyprotein constructs; the (I27-TmCsp)$_3$-I27^{GA} polyprotein constructed using the Gibson assembly method (I27 is depicted as yellow upward triangles and TmCsp is depicted as dark red upward triangles), for the (I27-TmCsp)$_3$-I27) polyprotein made using the cassette approach and published previously[31] (I27 is depicted as yellow downward triangles and TmCsp is depicted as red downward triangles) and the (I27-EcPOTRA2)$_3$-I27^{GA} polyprotein constructed using the Gibson assembly method (I27 is depicted as yellow circles and EcPOTRA2 is depicted as purple squares). Each set of data points at a given pulling speed show the median value of the unfolding force for I27, TmCsp and EcPOTRA2 from three experiments completed under the same conditions. The error bars indicate the standard deviation between the three experiments. Solid and dashed lines are the best fits to the data, where the solid line represents the fit to the I27$_{POTRA}$, GA data, the dotted-dashed line represents the fit to the I27$_{TmCSP}$, GA data and the dashed line indicates the fit to the I27$_{TmCSP}$ data. Solid lines show the best fits to the data for TmCSP and EcPOTRA2.

6

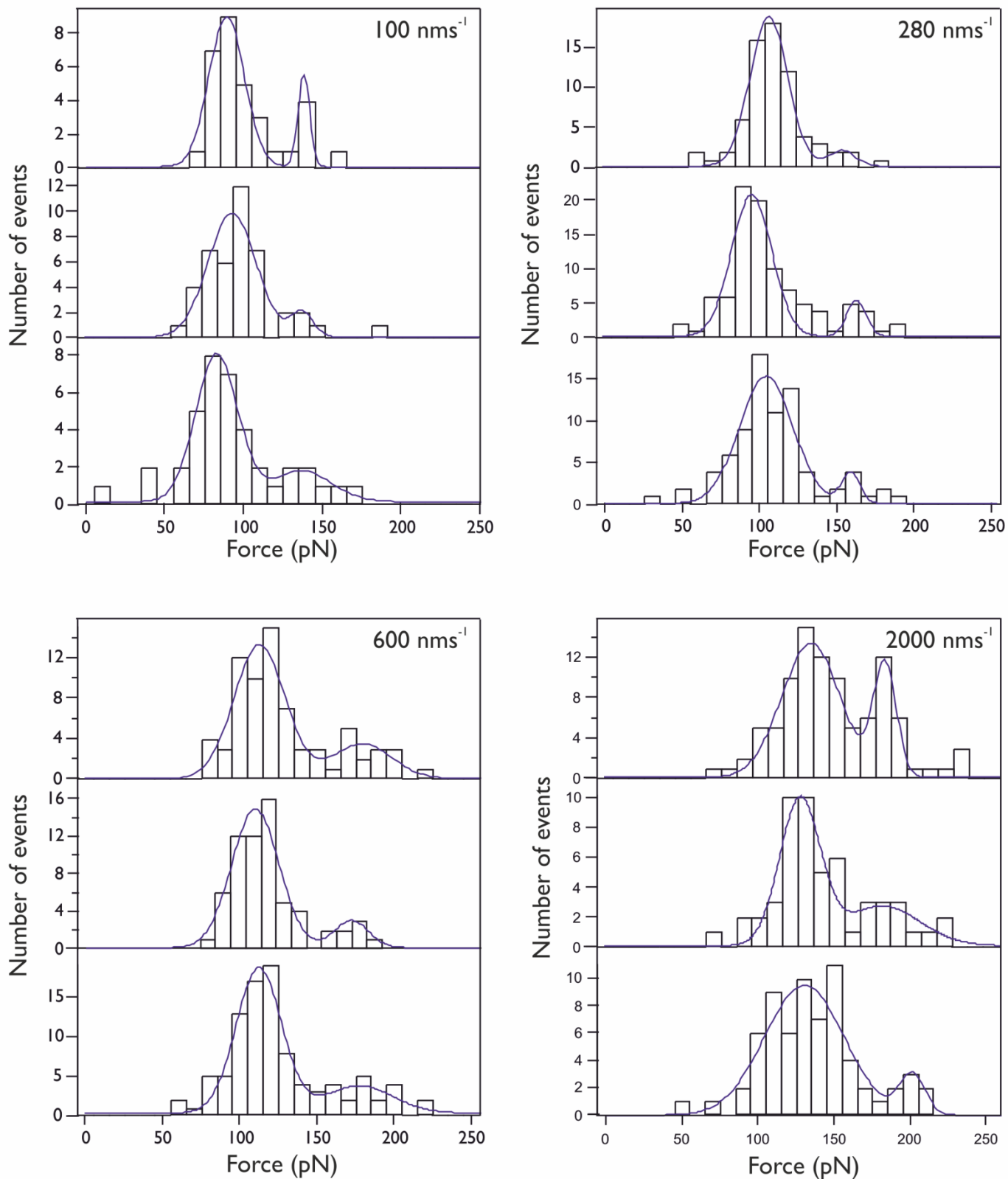**Figure S3**: Unfolding force histograms for experiments conducted in triplicate at pulling speeds of 100, 280, and 2000 nm s[-1] for the (I27- EcPOTRA2)$_3$-I27 polyprotein. The histograms show a clear separation in the distributions of the forces resulting from the mechanical unfolding of I27 and the EcPOTRA2. Gaussian fits to histograms for each data set are used to obtain a measure of the unfolding forces.
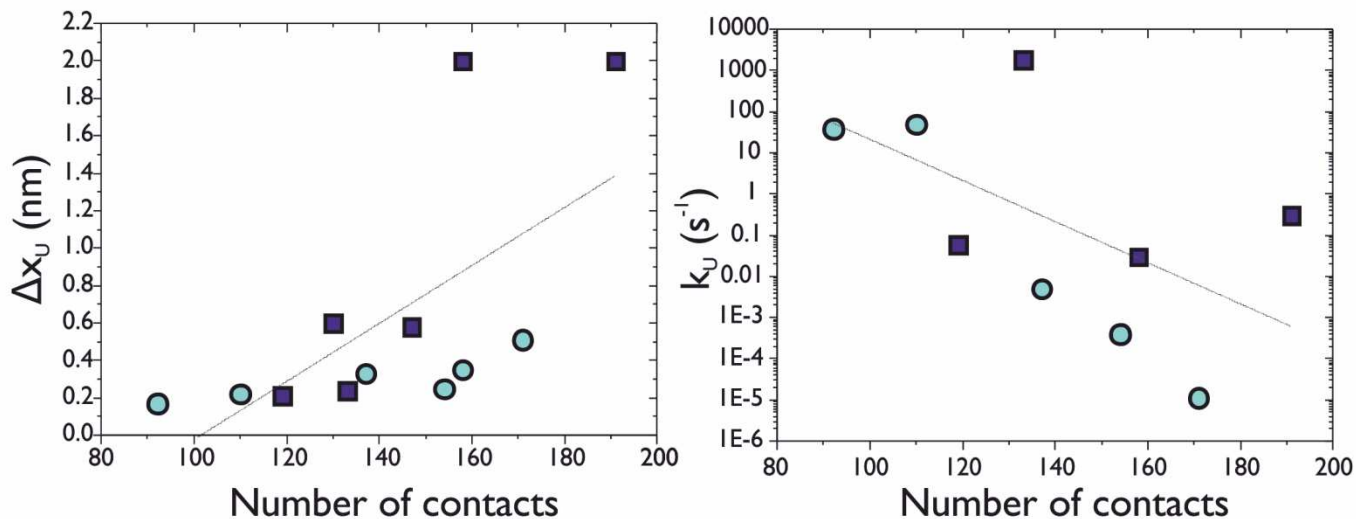
7

Figure S4: The number of inter-residue contacts does not correlate with the distance to the unfolding transition state, Δx_U and the unfolding rate, k_U when all twelve of the $\alpha + \beta$ proteins which have been studied using single molecule force spectroscopy are included. On the left is the number of inter-residue contacts plotted against Δx_U and on the right the number of inter-residue contacts plotted against k_U. Light blue circles denote proteins which possess proximal, parallel N- and C-terminal β-strands. $\alpha + \beta$ proteins which lack this structural feature are shown as dark blue squares.
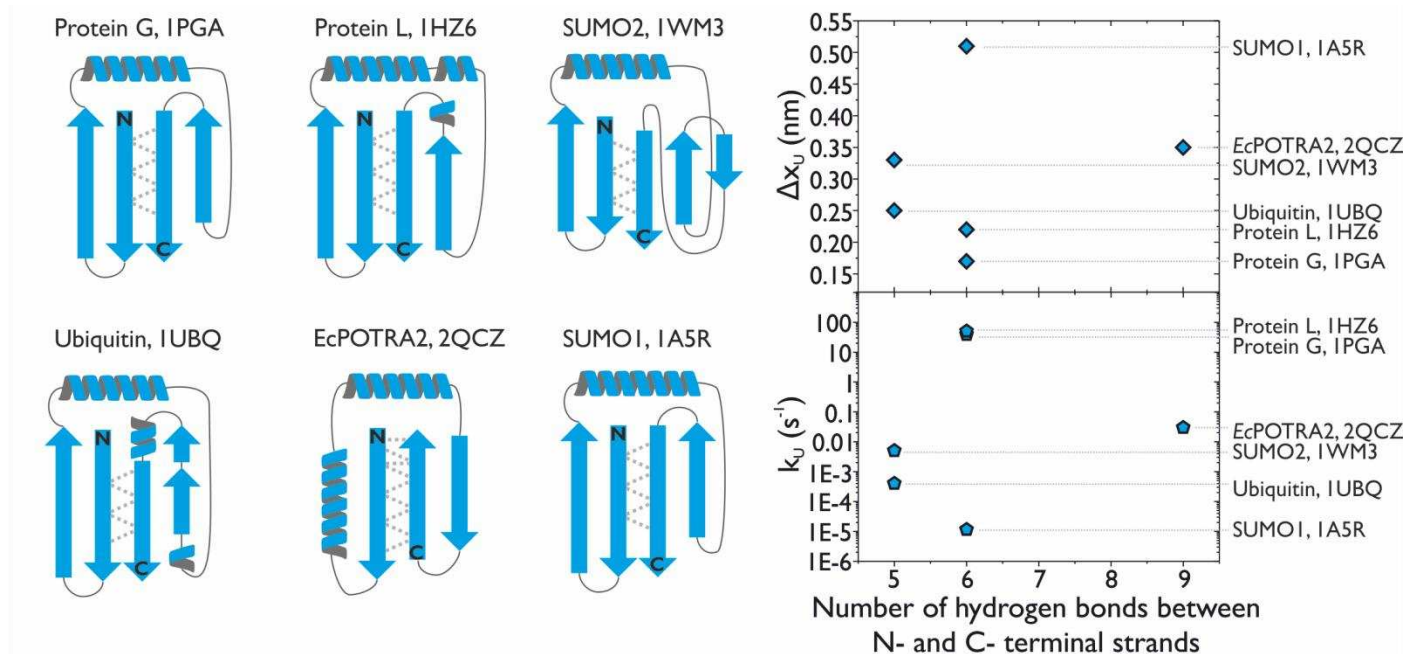
Figure S5: Schematics showing the topology of 6 α + β proteins, protein G, protein L, SUMO2, Ubiquitin, EcPOTRA2 and SUMO1 and the number of hydrogen bonds between the N- and C-terminal strands (grey dashed lines). (B) The number of hydrogen bonds between the N- and C-terminal strands is protein G (6), protein L (6), SUMO2 (5), Ubiquitin (5), EcPOTRA2 (9) and SUMO1 (6). The number of hydrogen bonds between the N- and C-terminal strands against Δx_U (upper) and k_U (lower) are shown.

Speed [nms^{-1}]	# Csp Events	# I27 Events	Median unfolding force TmCsp [pN] (± SD)	Average [pN] (±SD)	Median unfolding force I27 [pN] (± SD)	Average [pN] (±SD)
100	11	13	67 (± 13)		143 (± 14)	
	15	19	68 (± 9)	66 (± 5)	152 (± 18)	149 (± 6)
	33	41	61 (± 20)		152 (± 20)	
200	18	20	68 (± 14)		157 (± 20)	
	36	46	71 (± 12)	69 (± 1)	160 (± 16)	157 (± 2)
	12	17	69 (± 16)		155 (± 20)	
600	48	53	82 (± 14)		186 (± 22)	
	26	34	75 (± 16)	76 (± 6)	181 (± 20)	183 (± 2)
	13	21	72 (± 20)		181 (± 31)	
2000	43	44	91 (± 18)		199 (± 26)	
	50	55	83 (± 15)	87 (± 4)	190 (± 20)	194 (± 4)
	29	41	87 (± 13)		193 (± 36)	

Table S2. Summary of mechanical unfolding data for (I27-TmCsp)$_3$-I27^{GA}

Speed [nms^{-1}]	# Events	Median unfolding force POTRA [pN] (± Gaussian width)	Average [pN] (±SD)	Median unfolding force I27 [pN] (± Gaussian width)	Average [pN] (±SD)
100	24	90 (± 16)		139 (± 5)	
	31	95 (± 22)	91 (± 3)	140 (± 10)	139 (± 1)
	39	90 (± 19)		139 (± 27)	
280	86	108 (± 17)		155 (± 12)	
	96	96 (± 18)	103 (± 5)	164(± 9)	159 (± 4)
	80	104 (± 25)		159 (± 9)	
600	72	113 (± 23)		180 (± 27)	
	64	112 (± 22)	113 (± 1)	175 (± 16)	178 (± 2)
	96	113 (± 21)		178 (± 31)	
2000	96	133 (± 26)		181 (± 10)	
	53	127 (± 19)	130 (± 3)	180 (± 34)	188 (± 10)
	67	131 (± 36)		202 (± 12)	

Table S3. Summary of mechanical unfolding data for (I27-EcPOTRA2)$_3$-I27