Electronic Supplementary Material (ESI) for RSC Advances.
This journal is © The Royal Society of Chemistry 2015

Cross-Strand Disulfides in the Non-Hydrogen Bonding Site of Antiparallel β-sheet (aCSDns): Poised for Biological Switching

Naomi L. Haworth,[A, B] Merridee A. Wouters[C,D,*]

[A] Life and Environmental Sciences, Deakin University, Geelong 3217, Vic, Australia.

[B]Victor Chang Cardiac Research Institute, Sydney 2010, NSW, Australia.

[C] Olivia Newton-John Cancer Research Institute, Heidelberg 3084, Vic, Australia.

[D] School of Cancer Medicine, La Trobe University, Melbourne, 3086, Vic, Australia.

[*]Corresponding author. Email: Merridee.Wouters@onjcri.org.au

Figure S1. Hydrogen bonding diagrams for parallel and antiparallel β-sheet showing: **A** the two different sites (HB and NHB) in antiparallel β-sheet; **B** the single trapezoidal shaped site in parallel β-sheet (which is repeated with a 180º rotation along the chain); **C** sheet shear along the N-C terminal backbone direction (positive shear) in antiparallel β-sheet – this allows the formation of weak hydrogen bonding interactions between the carbonyl oxygens and the H_α atoms (shown in green) resulting in a more stable system; **D** sheet shear in parallel β-sheet where half the hydrogen bonds are compressed (shown in blue) and the remainder are stretched (shown in red), increasing the strain in the system and disfavouring the shear. Note: this figure was previously published in Reference 11: N. L. Haworth and M. A. Wouters, *RSC Adv.*, **2013**, *3*, 24680-24705.

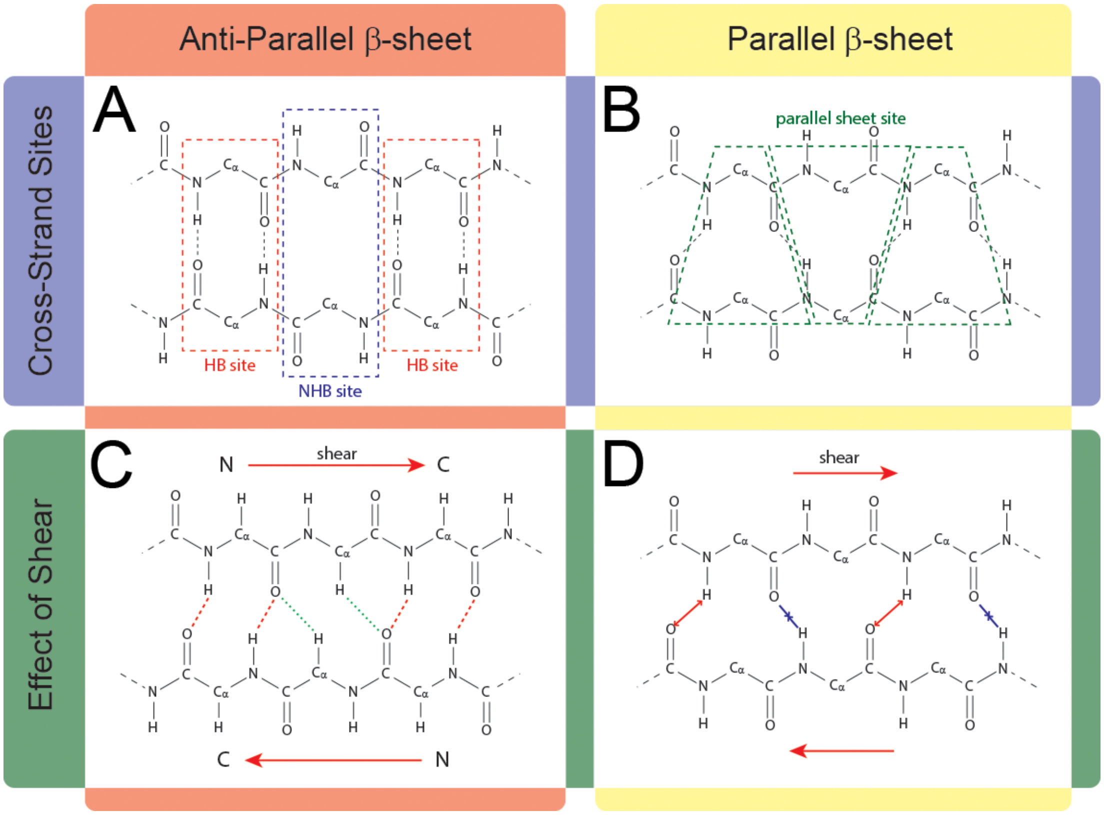

Figure S2. Superimposition of two end-aCSDn structures showing the negligible effect on the disulfide bond of one Cys residue adopting an α backbone conformation. PDB: 1jc1 (redox-sensitive green fluorescent protein) Cys 147 to Cys 204, in which both Cys residues adopt β backbone conformations, is shown in blue. PDB: 2nlr (endoglucanase CelB2) Cys 5 to Cys 31, in which Cys 5 (the lower Cys) adopts an α backbone conformation, is shown in red.

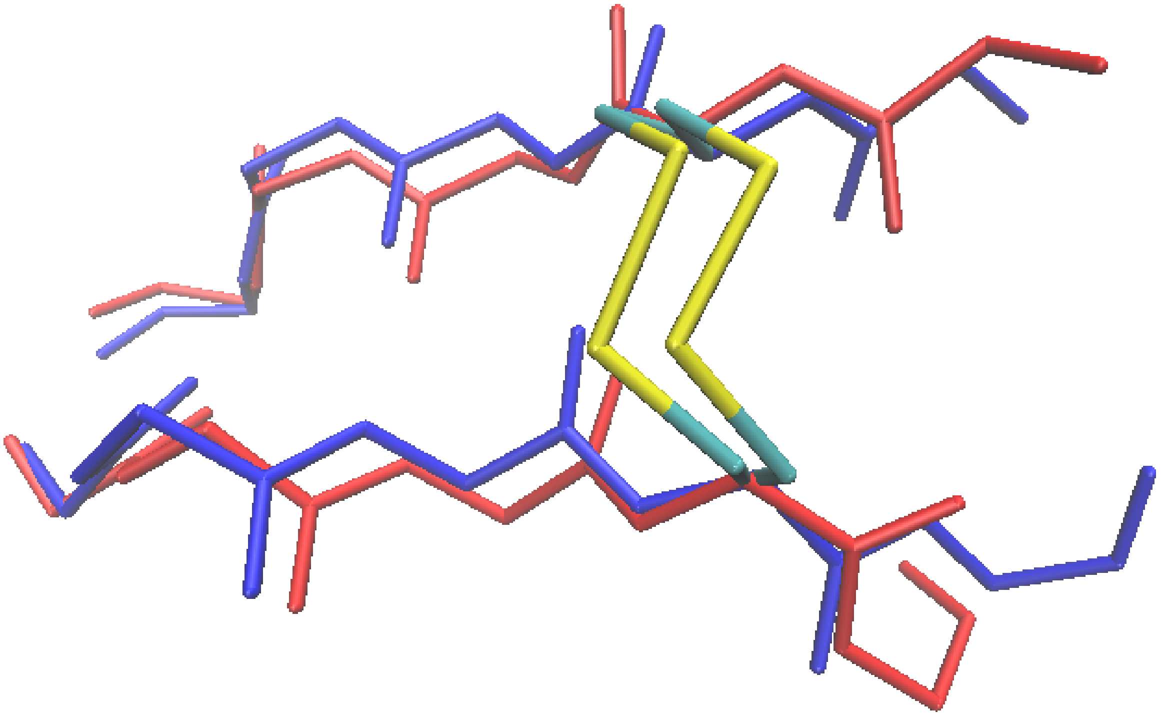

Figure S3. Conformations adopted by right-handed aCSDns which have χ_3 outside the range of 85º to 115º. The majority of these disulfides adopt the staple conformation, with only a very small proportion being right-handed cis disulfides.

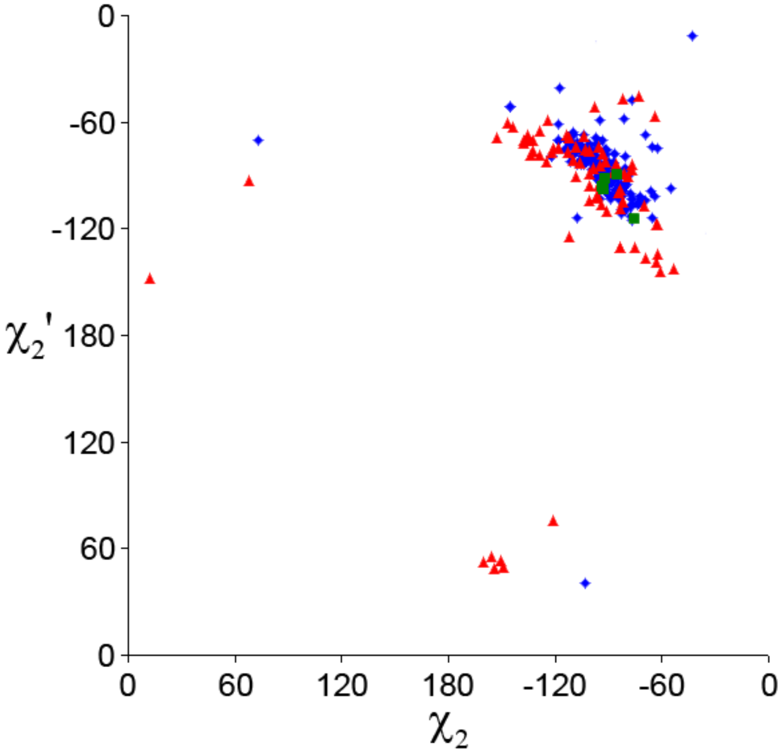

Figure S4. **A** Definitions of the residue backbone vectors $\mathbf{b_1}$ and $\mathbf{b_2}$ and the strand backbone vector $\mathbf{b_2}$-$\mathbf{b_1}$ ($\mathbf{b_2}$+$\mathbf{b_1}$ for parallel β-sheet). **B** Definition of interresidue shear showing the projection of the C_α positions onto the strand backbone vector with the shear being the distance between the projected points. **C** Definition of the vector used as the central axis of the dihedral angle in the calculation of interresidue twist. **D** Definition of interresidue twist as viewed along the direction of the green arrow shown in part **C**. Note: this figure was previously published in the Supplementary Material of Reference 11: N. L. Haworth and M. A. Wouters, *RSC Adv.*, **2013**, *3*, 24680-24705.

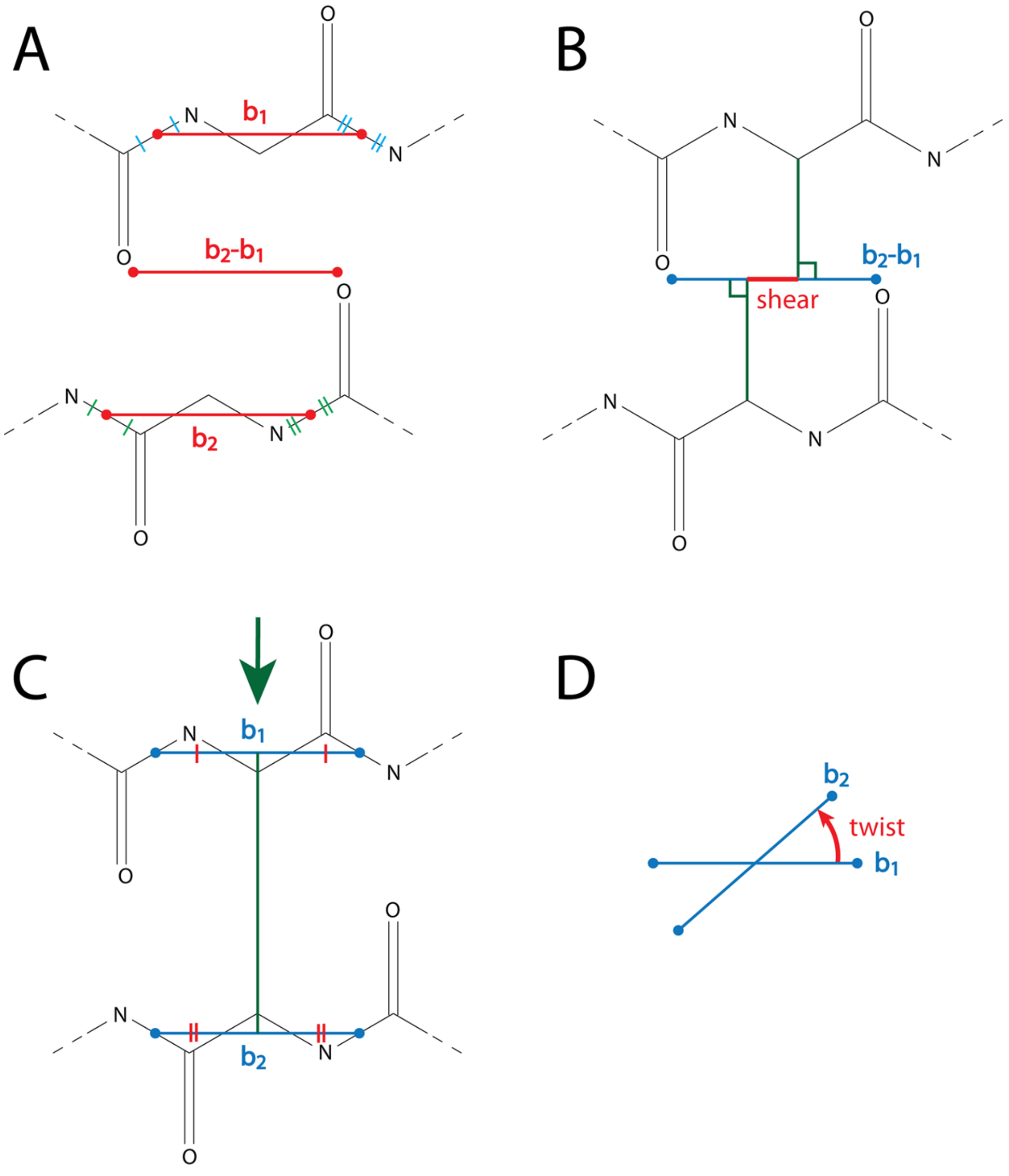

Table S1. A complete list of disulfide properties and protein fold data for all aCSDn clusters. Cases where proof of redox activity is indicated by "Yes" in column 11 are described in more detail in Table 3. Also see Table 3 for more detailed descriptions of column headings. *Access to this disulfide is impeded along both edges. †The disulfide in this structure is in a mixed oxidation state.

Cluster Name	Representative disulfide	No Ox	TTE /kJ mol^{-1}		LonE	LatE	Conf.	Non str?	Taxonomic Conservation	Proof of redox activity?	Fold	Position within fold	Example of fold without disulfide
			Ave	σ									
Duffy binding protein #1	1zro A 25-36	12	11.9	1.7	true/end	0	rst	No	*Plasmodium*		Duffy binding domain-like	On N-terminal β-hairpin lying between helical sub-domains	Not seen but unlikely to be impossible
Duffy binding protein #2	2xu0 A 477-483	1	12.2	0.0	true	0	rst	Na	*Plasmodium falciparum*		Duffy binding domain-like	On C-terminal β-hairpin	Not seen - no other structures with C-terminal β-hairpin
ERO1-like protein α #1	3ahq A 37-46	1	11.7	0.0	true	1	rst	Na	Metazoa, ichthyosporea, choanoflagellida		ERO1-like	Connecting strands 1 and 2 (3 strand β-sheet involving N-terminal β-hairpin)	Not seen - N-terminal β-hairpin missing in other structures
ERO1-like protein α #2	3ahq A 35-48	1	37.0	0.0	end	1	rcs	Na	Metazoa, ichthyosporea, choanoflagellida		ERO1-like	Connecting strands 1 and 2 (3 strand β-sheet involving N-terminal β-hairpin)	Not seen - N-terminal β-hairpin missing in other structures
Immunoglobulin heavy chain variable domain (VH)	1ai1 H 35-52	2	13.4	1.4	true	1*	rst	Yes	Mammalia		Immunoglobulin-like β-sandwich	CC' - outside (N-terminal end)	VH mouse cluster 7.2 (1j1p)
CD4, 1st C2-set domain	1cdy A 130-159	23	18.4	10.8	true	0/1/2	rst/lcs	Yes	Primates + muroids	Yes	Immunoglobulin-like β-sandwich	CF - inside (central)	CD4 3rd C2 set domain (1cdy)
lir-1 and lir-2, ligand binding domains	1ugn A 134-144	11	15.1	2.7	true/end	0/1	rst	Yes	Haplorrhini		Immunoglobulin-like β-sandwich	CC' - outside (central)	ICAM-2 (1zxq)
Intercellular adhesion molecules (ICAM) 1, 3 and 5	2oz4 A 404-430	1	12.7	0.0	true	1*	rst	Yes	Euteleostomi		Immunoglobulin-like β-sandwich	CF - inside (central)	ICAM-2 (1zxq)
Chitinase, N-terminal domain	1eib A 115-120	27	13.2	1.2	true	1	rst	Yes	Bacteria, arthropoda + viruses	2wlz A 115-120†	Immunoglobulin-like β-sandwich	FG - outside (N-terminal end)	Bacterial chitobiase, C-terminal domain (1qba)
Cytokine receptors #1	2c4f U 186-209	50	14.4	2.4	true	0/1	rst	Yes	Euteleostomi + viruses	Yes	Immunoglobulin-like β-sandwich	FG - inside (opposite end to β-hairpin)	IL2RB (2b5i)
Type 1 cytokine receptors #2	2b5i B 10-20	65	14.1	4.4	true	1	rst	Yes	Metazoa	1ern A 28-38	Immunoglobulin-like β-sandwich	AB - inside (central)	Tissue Factor (2c4f)
Type 1 cytokine receptors #3	1n26 A 146-157	5	21.3	3.5	end	1	rst	Yes	Euteleostomi		Immunoglobulin-like β-sandwich	DE - inside (central - at point where C' strand is transitioning to become D strand)	Tissue Factor (2c4f)
Type 1 cytokine receptors #4	1cd9 B 148-191	5	13.5	1.8	true	1/2	rst	Yes	Euteleostomi		Immunoglobulin-like β-sandwich	CF - inside (N-terminal end)	Tissue Factor (2c4f)
Type 1 cytokine receptors #5	3di2 B 54-62	3	12.8	1.0	true	1	rst	Yes	Euteleostomi		Immunoglobulin-like β-sandwich	CC' - outside (C-terminal end, well before C' strand transitions to form D strand)	Tissue Factor (2c4f)
Interleukin-12, p40 domain	3d85 D 148-171	6	12.9	0.9	true	1	rst	Yes	Euteleostomi	3d87 D 148-171	Immunoglobulin-like β-sandwich	DE - inside (on C-terminal side of AB disulfide)	Tissue Factor (2c4f)
Interferon γ receptor, α chain	1fyh E 178-183	2	31.2	21.9	true	1	rst/lst	No	Euteleostomi + viruses		Immunoglobulin-like β-sandwich	DE - inside (short hairpin at N-terminal end)	Not seen - hairpin seems to be being held together by disulfide
Actinoxanthin-like	1noa A 37-47	6	13.8	1.5	true	2	rst	No	Actinomycetales		Immunoglobulin-like β-sandwich	CC' - outside (C-terminal end)	Not seen but unlikely to be impossible
Pilus chaperones (PapD N-terminal domain-like)	1z9s A 98-137	14	12.8	0.9	true	1/2	rst	Yes	Bacteria	1p5u A 98-137†	Immunoglobulin-like β-sandwich	FG - outside (central)	FimC (3bwu)
DsbD, N-terminal domain	1l6p A 103-109	2	14.3	0.8	true	1	rst	Yes	Bacteria	Yes	Immunoglobulin-like β-	FG - outside (near	Immunoglobulin

											sandwich	β-hairpin)	(1bww)
Ephrin receptors, 1st Fn-III domain	3fl7 A 369-376	1	17.7	0.0	true	1	rst	Yes	Chordata		Immunoglobulin-like β-sandwich	CC' - outside (near β-hairpin)	Immunoglobulin heavy chain variable domain (VH) mouse cluster 7.3 (1ai1)
β-1,3-Xylanase	2cov D 413-418	4	15.3	0.1	true	1	rst	Yes	Proteobacteria		Immunoglobulin-like β-sandwich	CC' - outside (near β-hairpin)	Interleukin-6 receptor α chain, domain 2 (1n26)
Integrin α, thigh domain	3ije A 478-535	3	27.5	23.7	true	1	rst/lst	Yes	Metazoa		Immunoglobulin-like β-sandwich	BE - outside (central)	Immunoglobulin (1bww)
Integrin α, calf-1 domain	3ije A 668-681	3	13.8	0.9	true	1	rst	Yes	Metazoa		Immunoglobulin-like β-sandwich	C'C'' - inside (central)	Immunoglobulin (1bww)
Integrin α, calf-2 domain	3ije A 822-884	3	14.7	1.2	true/end	1	rst	Yes	Metazoa	Yes	Immunoglobulin-like β-sandwich	C'C'' - inside (central)	Immunoglobulin (1bww)
UL18 protein	3d2u A 238-253	2	14.7	0.5	true	1	rst	Yes	Human cytomegalovirus		Immunoglobulin-like β-sandwich	DE - inside (central)	Immunoglobulin (1bww)
Invasion-associated protein B #1	3dtd A 54-63	12	12.5	0.3	true	0	rst	Na	Proteobacteria		Immunoglobulin-like β-sandwich	Connecting strands A-1 and A-2 - inside (near β-hairpin)	Not seen - not in main fold
Invasion-associated protein B #2	3dtd A 124-129	12	12.6	0.2	true	1/2	rst	Yes	Proteobacteria		Immunoglobulin-like β-sandwich	DE - inside (near β-hairpin)	Immunoglobulin (1bww)
Complement α chain-like proteins #1	2hr0 B 1367-1436	12	16.5	9.8	true	1	rst/lst/lcs	Yes	Metazoa + plants	2wii B 1367-1436	Immunoglobulin-like β-sandwich	BE - outside (central)	Immunoglobulin (1bww)
Muscle-specific kinase receptor	2iep A 98-112	2	12.0	0.1	true	1	rst	Yes	Vertebrata		Immunoglobulin-like β-sandwich	FG - outside (central)	Titin (1g1c)
Spondin-1/Reelin	2zou A 156-182	8	13.5	1.0	end	0	rst	Yes	Metazoa		Immunoglobulin-like β-sandwich	FG - outside (opposite end to β-hairpin)	Immunoglobulin (1bww)
New antigen receptor (IgNAR), variable domain	2coq A 29-89	1	11.2	0.0	end	0	rst	Yes	Orectolobiformes		Immunoglobulin-like β-sandwich	CF - inside (N-terminal end)	Immunoglobulin heavy chain variable domain (VH) mouse cluster 7.3 (1ai1)
Anti-VSG immunoglobulin heavy chain, variable domain cAbAn35	1yc7 A 32-99	6	15.0	3.7	true/end	0/1	rst	Yes	*Camelus dromedarius*		Immunoglobulin-like β-sandwich	CF - inside (N-terminal end)	Immunoglobulin heavy chain variable domain (VH) mouse cluster 7.3 (1ai1)
Monoclonal Antibody 2D12.5 Fab	1nc2 C214-D221	1	11.7	0.0	end	0	rst	Na	na		Immunoglobulin-like β-sandwich	Intermolecular	na
Putative membrane anchored protein	3lso A 52-196	2	11.8	1.7	true	2	rst	Yes	*Corynebacterium*		Immunoglobulin-like β-sandwich	Connecting strands A-1 and G	Immunoglobulin heavy chain variable domain (VH) mouse cluster 7.3 (1ai1)
T cell receptor 5c.c7, β chain	3qjh D 69-75	5	15.5	2.1	end	0/1	rst	Yes	Muridae		Immunoglobulin-like β-sandwich	DE - inside (β-hairpin end)	T-cell antigen receptor (1oga)
Diphtheria toxin, C-terminal domain	1f0l A 461-471	9	12.3	0.8	true	1	rst	Yes	*Corynebacterium*		Common fold of diphtheria toxin/transcription factors/cytochrome f	Connecting strands 8 and 9 (additional strands between D and E of Ig fold) - inside (central)	Bacterial chitobiase, N-terminal domain (1qba)
FimH (Minor pilus subunit)	1klf B 3-44	29	15.2	2.0	true	1/2*	rst	Yes	Proteobacteria	1tr7 A 3-44	Common fold of diphtheria toxin/ transcription factors/cytochrome f	AB - inside (N terminal end)	Caf1 (1pv5)
Minor pilus subunits	1klf B 161-187	37	14.1	3.3	true/end	0/1	rst	Yes	Proteobacteria		Common fold of diphtheria toxin/ transcription factors/cytochrome f	A1B2 - inside (N terminal end)	Caf1 (1pv5)
F17-type fimbrial adhesins	2bsc A 53-110	14	12.3	0.6	true (A-bent)	0	rst	Yes	γ proteobacteria		Common fold of diphtheria toxin/transcription factors/cytochrome f	Connecting strands 5 and 10 (C to additional strand between D and E of Ig fold) - outside (N-terminal end)	FimH (1klf)
Rel/Dorsal transcription factors, DNA binding domain	1lei B 116-121	3	35.1	2.0	true	0/1*	lcs	Yes	Metazoa	1ooa A 116-121	Common fold of diphtheria toxin/transcription factors/cytochrome f	Connecting strands 6 and 7 (short β-hairpin at start of C' in Ig labelling) - inside (central)	p65 subunit of NF-kappa B, N-terminal domain (2ram)

T-Box transcription factors, TBX4 and TBX5	2x6v A202-B202	1	15.2	0.0	true	2	rst	Na	Vertebrates		Common fold of diphtheria toxin/transcription factors/cytochrome f	Homodimer interface	na
Ceruloplasmin/Coagulation factors/Hephaestin	2j5w A 155-181	8	20.4	11.2	true/end	0/1	rst/lst/lcs	Yes	Chordata		Cupredoxin-like	Connecting strands 7 and 8 - outside (C-terminal end)	Multicopper oxidase CueO (1kv7)
Plastocyanin [g]	1tef A 30-69	4	15.1	1.7	true	1	rst	Yes	Engineered		Cupredoxin-like	Connecting strands 3 and 5 - outside	Plastocyanin (1plc)
Pilus chaperones (PapD C-terminal domain-like)	1n0l A 207-212	18	15.8	1.4	true	1	rst	Yes	Bacteria		C2 domain-like	GH - inside (short β-hairpin)	FimC (3bwu)
WWC1 and WWC2 proteins, C2 domain	2z0u A 112-122	2	15.5	0.1	true	1	rst	Yes	Euteleostomi		C2 domain-like	FG - outside (near β-hairpin)	Synaptogamin I (1uow)
Neurophysin II	2hnu A 74-79	47	15.0	4.3	true	1/2	rst	No	Metazoa		Neurophysin II	Connecting strands 3 and 4 - inside sandwich facing equivalent disulfide further along strand	Not seen
Cys-loop ligand gated ion channels, ligand-binding domain	2qc1 B 128-142	1	14.1	0.0	end	1*	rst	No	Eukaryota		Nicotinic receptor ligand binding domain-like	Connecting strands 5 and 6 at point where strand 5 crosses from N- to C-terminal sheets	Not seen - in 1i9b H-bonding is not formed between strands 5 and 5 but disulfide is present
Ephrin receptors, ligand binding domain	2bba A 61-184	49	17.5	4.4	end	0	lst	Yes	Metazoa	3nru I 74-192	Galactose-binding domain-like	EM - outside	Galactose oxidase, N-terminal domain (1gof)
Cellulase C, cellulose-binding domain	1gu3 A 33-140	1	12.7	0.0	true	2	rst	Yes	*Cellulomonas*		Galactose-binding domain-like	Connecting strands 2 and 11 - outside	Galactose oxidase, N-terminal domain (1gof)
Fucose binding lectin	1k12 A 50-146	5	14.4	0.8	end	0	rst	Yes	Metazoa		Galactose-binding domain-like	Connecting strands 2' (not detected by HERA) and 9 - outside (near Ca binding site)	Galactose oxidase, N-terminal domain (1gof)
Exo-1,3/1,4-beta-glucanase, C-terminal domain	3f93 A 736-741	5	12.5	0.4	true (A-bent)	1	rst	Yes	Bacteria		Galactose-binding domain-like	Connecting strands 9 and 10 - inside A-bent bulge	Endoglucanase CelJ (1wzx, no bulge)
Hemagglutinin-esterase, hemagglutinin domain	3i26 A 200-246	16	13.7	0.8	true	0	rst	Yes	Viruses		Viral protein domain	Connecting strands 5 and 8 (only counting strands in the two sheets) - outside (central)	Hemagglutinin (1jsd)
Dextranase, N-terminal domain	1ogo X 9-14	1	16.9	0.0	true	1	lst	No	*Penicillium* + *Lipomyces*	1ogm X 9-14[†]	Dextranase, N-terminal domain	Connecting first two (N-terminal) strands (1st not detected by HERA) - outside	Not seen but unlikely to be impossible
Complement c1q globular head: A, B and C chains	2jg9 A 150-168	27	13.9	2.1	true	0/1	rst	Yes	Euteleostomi		TNF-like	Connecting strands 6 and 8 - outside (non-terminal end)	TNF (4tsv)
Tumor necrosis factor ligand superfamily member 4, OX40L	2hew F 98-183	4	12.4	0.2	β-bridge	0	rst	Yes	Mammalia		TNF-like	Connecting C-terminal to fold over between strands 4 and 5	TNF (4tsv)
Thaumatin-like proteins	2pe7 A 56-66	15	15.8	9.9	true (A-bent)	0	rst/lst	Na	Bacteria + eukaryota	1rqw A 56-66	Osmotin, thaumatin-like proteins	Inside A-bent β-ribbon	Not seen - disulfide not in main fold
Pentraxin-like proteins	2w08 A 36-95	108	17.9	7.9	true	2	rst/lcs	Yes	Metazoa	1gyk E 36-95	Concanavalin A-like lectins/glucanases	Connecting strands 3 and 8 - outside of outer layer (central)	Calnexin (1jhn)
Glycosyl hydrolase family 7 #1	1ojk A 18-24	53	13.7	2.2	true	0/1	rst	Yes	Fungi, amoebozoa, parabasalia	1oji A 18-24	Concanavalin A-like lectins/ glucanases	Connecting strands 9 and 10 near β-hairpin - outside of layers	Glycosyl hydrolase family 16 (1gbg)
Glycosyl hydrolase family 7 #2	1ojk A 215-234	55	14.4	4.0	β-bridge	0	rst	Na	Fungi, amoebozoa, parabasalia		Concanavalin A-like lectins/ glucanases	Variant where N-terminal forms β-hairpin to hydrogen bond with C-terminal in parallel rather than antiparallel alignment, disulfide on β-hairpin	Not seen - variant fold

Glycosyl hydrolase family 12, catalytic domain #1	2nlr A 5-31	42	14.8	3.1	true/end	1	rst/lcs	Na	Bacteria, archaea, fungi, plants, stramenophiles		Concanavalin A-like lectins/glucanases	Connecting strands 1 and 4 - outside of outer layer	Not seen - N-terminal extension
Glycosyl hydrolase family 12, catalytic domain #2	2nlr A 64-69	11	12.7	0.8	true	0	rst	Na	Actinobacteria, acidobacteria, bacteroidetes		Concanavalin A-like lectins/glucanases	Connecting strands 7 and 8 (very short β-hairpin)	Not seen - on additional hairpin
Peptidase A4 #1	2ifr A 141-148	4	14.3	3.1	true/end	0	rst	Yes	Pezizomycotina (not eurotiomycetes)		Concanavalin A-like lectins/ glucanases	On extended β-hairpin between strands 10 and 11	Aspergillopepsin II (1y43)
Peptidase A4 #2	2ifr A 194-203	5	11.3	0.2	true	1	rst	Yes	Pezizomycotina (not eurotiomycetes)		Concanavalin A-like lectins/ glucanases	Connecting strands 13 and 14 (C-terminal β-hairpin) - between layers	Aspergillopepsin II (1y43)
Peptidase A4 #3	1y43 B 6-30	2	15.6	3.9	true	1/2*	rst/lst	Yes	Eurotiomycetes	3o0x A 105-137	Concanavalin A-like lectins/glucanases	Connecting strands 1 and 2 - inside cavity (opposite end to β-hairpin)	Scytalidopepsin B (2ifr)
Gustavus	2ihs A160-B160	1	18.2	0.0	β-bridge	0*	rst	Na	Endopterygota		Concanavalin A-like lectins/glucanases	Homodimer interface (extension of strand 10)	na
Galectin CG-1B	3dui A7-B7	1	11.0	0.0	true	2	rst	Yes	*Gallus gallus*	Yes	Concanavalin A-like lectins/ glucanases	Homodimer interface	Galectin CG-1A (1qmj)
Rhamnogalacturonase B, N-terminal domain	1nkg A 164-173	5	14.7	0.5	true	2	rst	Yes	Pezizomycotina		Supersandwich	Connecting strands 12 and 13 - outside (central)	Galactose mutarose (1z45)
InaD with NorpA	1ihj A31-D6	2	12.9	1.1	true	1	rst	Na	*Drosophila*	Yes	PDZ domain-like	Intermolecular	na
Minor coat protein g3p, N-terminal domains	1g3p A 46-53	4	13.3	0.7	true	1/2*	rst	No	Bacterial phages + bacteria		N-terminal domains of the minor coat protein g3p	Connecting strands 3 and 4 - inside barrel (near β-hairpin)	Not seen
Hydrophobins #1	1r2m A 14-26	13	12.9	1.2	true (A-bent)	1/2	rst	No	Dikarya		Hydrophobin II	Connecting strands 1 and 2 - inside barrel (facing Hydrophobin II #2)	Not seen
Hydrophobins #2	1r2m A 53-64	21	14.8	1.9	true/end	1*	rst	No	Dikarya		Hydrophobin II	Connecting strands 3 and 4 - inside barrel (facing Hydrophobin II #1)	Not seen
Pertussis toxin / Subtilase cytotoxin #1	1bcp B 120-134	29	14.7	4.4	true	0/1/2*	rst/lcs	Yes	Proteobacteria	3dwq C 31-45	OB-fold	Connecting strands 3 and 4 - non-helix side (central)	Cholera toxin (3chb)
Pertussis toxin / Subtilase cytotoxin #2	3dwa A 103-109	15	14.9	1.4	true	0	rst	Yes	Proteobacteria		OB-fold	Forming a β-hairpin in the C-terminus, following strand 6	Enterotoxin A (1dyq)
Kunitz (STI) inhibitors	2qn4 B 140-144	15	14.2	2.3	true/end	0/1	rst/lst	Yes	Viridiplantae		β-trefoil	Connecting strands 11 and 12 - outside (near β-hairpin)	Two-chain trypsin inhibitor (1r8o)
α-L-Arabinofuranosidase B, C-terminal domain	1wd3 A 401-439	4	9.6	1.0	end	2	rst	Yes	Fungi and Bacteria		β-trefoil	Connecting strands 5 and 8 - outside	Two-chain trypsin inhibitor (1r8o)
α-Lytic protease	1ssx A 137-159	48	13.0	0.8	true	1	rst	Yes	Fungi and Bacteria	2ull A 137-159	Trypsin-like serine proteases	Connecting strands 1 and 2 of 2nd β-barrel - outside (central)	Achromobacter protease (1arb)
Plasminogen, catalytic domain	1ddj B 737-765	2	14.0	11.2	end	0	rst	Yes	Eukaryota		Trypsin-like serine proteases	Connecting strands 4 and 5 of 2nd β-barrel - outside (central) - holding together one side of a very large bulge in the barrel wall	Viral capsid protein (1wyk)
Eukaryotic trypsin-like serine proteases	1s5s A 136-201	936	16.5	3.8	true/end	1/2*	rst	Yes	Vertebrates + some other metazoa	1y3w A 136-203	Trypsin-like serine proteases	Connecting strands 1 and 4 of 2nd β-barrel (binding β-hairpin to neighbouring strand) - inside (near β-hairpin)	Thrombin (1h8d)

Serine protease NS3	3p8n A 47-52	2	14.8	0.5	true	1*	rst	Yes	Flaviviridae		Trypsin-like serine proteases	Connecting strands 2 and 3 of 1st β-barrel - inside (near β-hairpin)	Trypsin (1s5s)
Prophenoloxidase activating factor-I	2olg A 184-191	3	17.2	3.2	true	0	rst	No	Arthropoda		Trypsin-like serine proteases	On β-hairpin in elaboration between strands 4 and 5 of 1st β-barrel (near Ca binding site)	Not seen
Hypothetical protein ORF137	2x4j A 55-133	1	12.9	0.0	true	2	rst	No	Pyrobaculum spherical virus		Two interconnected β-barrel-like motifs	Connecting strands 4 and 11	Not seen - no other structures with this fold
β-Secretase / Xylanase inhibitor TAXI-I	1w50 A 269-319	230	14.0	2.1	true	0	rst	No	Eukaryota	3ivi C 330-380	Acid proteases	On 3 strand sheet, on end of C-terminal domain looping around binding site	Usually a tail-abddh on other side of sheet (1pso)
Dual adaptor of phosphotyrosine and 3-phosphoinositides, DAPP1/PHISH	1fao A 227-240	2	15.2	1.0	true	1*	rst	Yes	Euteleostomi		PH domain-like barrel	Connecting strands 6 and 7 - outside barrel (opposite end to β-hairpin)	Rac-α serine/threonine kinase (1unr)
Rac-α serine/threonine kinase	1unr A 60-77	3	17.1	2.9	end	1*	rst	Yes	Metazoa	1unq A 60-77	PH domain-like barrel	Connecting strands 5 and 6 - inside barrel (opposite end to β-hairpin)	Dual adaptor of phosphotyrosine and 3-phosphoinositides, DAPP1/PHISH (1fao)
β-Lactoglobulin	1beb A 106-119	38	14.2	2.9	true	2	rst	Yes	Mammalia	2q2m A 106-119†	Lipocalins	Connecting strands 7 and 8 - outside barrel (middle)	Thrombin inhibitor (1avg)
Thrombin inhibitor	1avg I 69-84	1	12.9	0.0	true	2	rst	Yes	Hemiptera		Lipocalins	Connecting strands 4 and 5 - inside barrel (middle)	β-Lactoglobulin (1beb)
Membrane bound lysozyme inhibitor of C-type lysozyme, MliC	3f6z B 51-124	2	15.5	0.8	true	2	rst	Yes	Bacteria		Streptavidin-like	Connecting strands 1 and 8 - inside	D-aminopeptidase, middle domain (1ei5)
Glucosidase II β subunit-like proteins	3aih A 110-123	2	14.8	0.9	end	1	rst	Yes	Eukaryota		Mannose 6-phosphate receptor domain	Connecting strands 1 and 2 (equivalent to strands 3 and 4 in MIR-receptor) - outside (N-terminal end)	MIR-receptor (1gp3)
Influenza neuraminidase #1	1f8e A 124-129	207	15.0	4.7	true	0	rst	Yes	Influenza A, B		6-bladed β-propeller	Inner edge of blade 1 - anticlockwise side - on β-hairpin sticking out top of propeller	Salmonella sialidase (3sil)
Influenza neuraminidase #2	1f8e A 232-237	205	14.6	2.0	true/end	1	rst	Yes	Influenza A, B		6-bladed β-propeller	Inner edge of blade 3 - anticlockwise side - on β-hairpin sticking out top of propeller	Salmonella sialidase (3sil)
Influenza neuraminidase #3	1f8e A 280-289	203	14.1	1.7	true	1	rst	Yes	Influenza A, B		6-bladed β-propeller	Inner edge of blade 4 - anticlockwise side - central	Salmonella sialidase (3sil)
Influenza neuraminidase #4	1f8e A 421-447	199	14.8	3.1	true	1/2	rst	Yes	Influenza A, B		6-bladed β-propeller	Middle of blade 6 - clockwise side	Salmonella sialidase (3sil)
Influenza neuraminidase #5	2htq A 278-291	3	16.9	1.6	end	0	rst	Yes	Influenza A, B		6-bladed β-propeller	Inner edge of blade 4 - anticlockwise side - central (usually within a β-bulge)	Salmonella sialidase (3sil)
Hemagglutinin-neuraminidase #1	2zb6 A 287-300	3	21.1	9.8	true	1	rst/lcs	Yes	Paramyxoviridae		6-bladed β-propeller	Inner edge of blade 2 - anticlockwise side - central	Salmonella sialidase (3sil)
Hemagglutinin-neuraminidase #2	1z4v A 365-375	6	12.4	0.4	end	1	rst	Yes	Paramyxoviridae		6-bladed β-propeller	Outer edge of blade 3 - anticlockwise side - near β-hairpin	Salmonella sialidase (3sil)

Hemagglutinin-neuraminidase #3	1e8t A 531-542	38	14.7	2.2	true	1	rst	Yes	Paramyxoviridae		6-bladed β-propeller	Inner edge of blade 6 - anticlockwise side - on β-hairpin sticking out top of β-propeller	Salmonella sialidase (3sil)
Nidogen	1npe A 1094-1105	1	12.2	0.0	true	2	rst	Yes	Metazoa		6-bladed β-propeller	Middle of blade 5 - clockwise side	Low density lipoprotein receptor (1ijq)
Arylsulfate sulfotransferase	3elq A 418-424	4	17.4	3.0	true	0	rst	Na	Proteobacteria		6-bladed β-propeller	On β-hairpin of elaboration between blades 4 and 5	Not seen - not in main fold
Amine dehydrogenases, heavy chain	2bbk H 168-183	84	15.0	1.7	true/end / β-bridge	1	rst	Yes	Protebacteria + Actinobacteria	1mda J 167-183	7-bladed β-propeller	Inner edge of blade 4 - anticlockwise side - bottom (blade 4 is twisted so disulfide is on protein surface)	Quinohemoprotein amine dehydrogenase (1pby B)
Integrin α, N-terminal domain #1	3ije A 59-67	1	13.3	0.0	true	1	rst	Yes	Metazoa		7-bladed β-propeller	Outer edge of blade 1 - anticlockwise side - central	Quinohemoprotein amine dehydrogenase (1pby B)
Integrin α, N-terminal domain #2	3ije A 108-128	23	13.9	2.7	true	2	rst	Yes	Metazoa		7-bladed β-propeller	Middle of blade 2 - clockwise side - central	Quinohemoprotein amine dehydrogenase (1pby B)
Sema domain	1q47 A 132-141	21	13.9	1.6	true/end	1/2	rst	Yes	Eukaryota + viruses	1q47 B 132-141	7-bladed β-propeller	Middle of blade 2 - clockwise side - central	Quinohemoprotein amine dehydrogenase (1pby B)
Fimbrial biogenesis protein type IV, PilY1	3hx6 A 1136-1141	2	15.1	1.3	true	1	rst	No	*Pseudomonas*		7-bladed β-propeller	Outer edge of blade 7 (3 strand blade) - clockwise side - central	Not seen - propeller split between blades 1 and 7
Dipeptidyl aminopeptidase-like proteins, N-terminal domain	1xfd A 411-418	4	15.6	0.1	true	1	rst	Yes	Metazoa		8-bladed β-propeller	Outer edge of blade 5 - anticlockwise side - near β-hairpin	Dipeptidyl peptidase IV/CD26, N-terminal domain (3c45)
Methanol dehydrogenase, heavy chain [f]	2d0v I 386-415	1	11.1	0.0	end	1*	rst	Yes	Bacteria		8-bladed β-propeller	Inner edge of blade 6 - anticlockwise side - blade 6 is double length with bulges in strands 1 and 3 between normal blade length and extension region, disulfide is just below bulge in extended region	Ethanol dehydrogenase (1flg)
CucA	2xlg A 65-267	12	11.8	0.7	true	1	rst	Na	Chroococcales		Double-stranded beta-helix	Connecting C-terminus to strand 2	Not seen - C-terminal strand different in related proteins
PR domain zinc finger protein 4, SET domain	3db5 B 49-124	1	12.4	0.0	true	1	rst	Yes	Euteleostomi	3db5 A 49-124	β-clip	Connecting strands 2 and 6 (strands that circle fold) - outside	Histone H3 K4-specific methyltransferase SET7/9 catalytic domain (1muf)
Translationally-controlled tumor protein	3ebm B172-C172	1	13.7	0.0	true	0*	rst	Na	Chordata + plants	Yes	Mss4-like	Homodimer interface	na
Cyanovirin-N	1lom A 8-22	30	12.7	0.6	true	1	rst	No	Bacteria, plants, fungi, viruses	1l5b B 8-22[†]	Cyanovirin-N	Connecting N-terminal end of strand 1 to strand 2	Not seen but unlikely to be impossible - just a 3 strand b-sheet
NGO0477 protein	3kxa A 8-17	4	16.0	0.6	end	1	rst	No	*Neisseria gonorrhoeae*		*4 strand β-sheet constructed from the terminii of all four strands*	On β-hairpin, facing disulfide from another chain	Not seen
Oligosaccharyl transferase stt3 subunit related protein	2zai A 638-658	8	14.7	0.9	true/end /β-bridge	0/1*	rst	No	*Pyrococcus*		*barrel, flattened; n=8, S=10; meander; strands 5 and 6 twisted*	Connecting strands 5 and 6 before twist	Not seen - in related structure 3aag strands 5 and 6 are fully inverted

E3 ubiquitin-protein ligase, UBR1	3nin A163-B163	1	13.6	0.0	β-bridge	2	rst	Na	*Saccharomyces*	3nil B163-F163	*Based on a β-ribbon; bears some resemblance to a single turn of a double stranded β-helix*	Homidimer interface	Not seen - no other structures with this fold
Zona pellucida domain containing proteins	3nk4 A 229-295	6	14.8	3.3	true	2	rst	No	Metazoa		*Complex all-β domain; subdomain fold based on greek key*	CF - inside (central)	Not seen - no other structures with this fold
Alphavirus envelope glycoprotein, E2 #1	3n40 P 155-169	4	15.0	0.2	end	0	rst	No	Alphaviruses		*Complex all-β domain; subdomain fold based on greek key*	DE - inside (non-β-hairpin end)	Not seen - no other structures with this fold
Alphavirus envelope glycoprotein, E2 #2	3n40 P 265-289	4	14.8	1.2	true	0	rst	No	Alphaviruses		*Complex all-β domain; subdomain: sandwich, 6 strands in 2 sheets*	Connecting strands 3 and 6 in subdomain	Not seen - no other structures with this fold
Flavivirus polyprotein propeptide	3c5x C 34-68	2	15.2	0.7	true/end	2	rst	No	Flaviviruses		*sandwich; 7 strands in 2 antiparallel sheets, top sheet 1275, bottom sheet 346*	Connecting strands 4 and 6 - inside	Not seen - no other structures with this fold
Sodium-potassium pump, β subunit	2zxe B 215-278	2	12.3	0.7	true	1	rst	Yes	Metazoa		*sandwich; 7 strands in 2 sheets; greek-key (Immunoglobulin-like but having many significant elaborations*	CF - inside (central)	Immunoglobulin (1bww)
Exo-1,3/1,4-beta-glucanase, N-terminal domain	3f93 A 303-308	5	16.0	0.7	end	0	rcs	Na	Proteobacteria		TIM β/α-barrel	Forming short β-hairpin in helix between strands 7 and 8	Not in main fold
Alcohol dehydrogenase 1	2hcy A277-B277	2	14.0	2.2	true	0	rst	Yes	*Saccharomyces*	Yes	NAD(P)-binding Rossmann-fold domains	Homodimer interface	Alcohol dehydrogenase (1hdx)
NGR ectodomain-like proteins #1 (Toll-like receptor 2)	2z80 A 30-36	6	18.0	6.3	true	0/1	rst/lcs	Yes	Euteleostomi		Leucine-rich repeat, LRR (right-handed β-α superhelix)	Connecting N-terminal strand to first loop of superhelix forming β-hairpin	Splicesomal U2A' protein (1a9n)
NGR ectodomain-like proteins #2	1ozn A 31-43	57	13.8	4.7	end	0/1	rst	Yes	Eukaryota	3g3a A 6-15	Leucine-rich repeat, LRR (right-handed β-α superhelix)	Connecting N-terminal strand to first loop of superhelix forming β-hairpin - disulfide further from β-hairpin than in NGR ectodomain-like proteins #1	Splicesomal U2A' protein (1a9n)
Hemagglutinin-esterase, esterase domain	3i26 A 88-136	14	12.0	0.5	true	0	rst	Yes	Orthomyxoviridae + torovirinae		Flavodoxin-like	Connecting loose β-strand from esterase domain to first strand of hemagglutinin domain	Hemagglutinin (1jsd)
Vancomycin A-type resistance protein, VanA	1e4e A 52-64	2	17.4	4.5	end	0*	rst	No	Bacilli		PreATP-grasp domain	Binds very short strand 4 to strand 3	Strand 4 of β-sheet is not seen when disulfide is not present
Heparan sulfate sulfotransferases	1t8t A 351-363	10	7.4	0.5	end	1	rst	Na	Metazoa		P-loop containing nucleoside triphosphate hydrolases	Holding together 3 strand b-sheet lying beside main fold (outside)	Not seen - disulfide not in main fold
Alkaline serine protease, Apa1	1y9z B 260-269	3	21.0	0.7	end	0	lcs	Na	Proteobacteria		Subtilisin-like	On additional α/β/β domain, holding together β-hairpin on the side of this domain	Not seen - disulfide not in main fold
Tripeptidyl-peptidase 1 precursor	3edy A 522-537	2	16.0	1.8	true	0	rst	Yes	Euteleostomi		Subtilisin-like	On β-hairpin C-terminal to main fold - outside	Serine-carboxyl proteinase (1t1e)
Thioredoxin	2ifq A73-B73	2	14.2	0.3	β-bridge	0	rst	Na	Metazoa	Yes	Thioredoxin fold	Homodimer interface	Na
Bacterioferritin comigratory protein, BCP	2cx4 G 49-54	6	14.1	1.6	true/end	0	rst	Na	Bacteria + archaea	Yes	Thioredoxin fold (glutatione peroxidase-like)	On additional β-ribbon β-hairpin	Not seen - disulfide not in main fold

												between strands 1 and 2	
Archaeal peroxiredoxin	2cv4 A 207-213	25	19.0	14.9	true	0	rst/lcs	Yes	Archaea	3a2x A 207-213	Thioredoxin fold (glutatione peroxidase-like)	On β-hairpin between strands 3 and 4 of C-terminal β-sheet sub-domain	Peroxiredoxin 6 (2v3g)
Glutathione peroxidase	3dwv A 47-95	2	10.4	0.6	end	0	rst	Yes	Euglenozoa	Yes	Thioredoxin fold (glutatione peroxidase-like)	On short β-ribbon section connecting the loop between strand 1 and helix 1 to the loop between helix 2 and strand 2'	Chordate glutathione peroxidases (1gp1)
Putative DsbA oxidoreductase SCO1869	3gl5 A 227-232	1	12.7	0.0	true	0	rst	Na	Bacteria		Thioredoxin fold	On C-terminal β-hairpin - detached from main domain	Na
Hypothetical protein	2oka A41-B41	2	14.2	0.0	true	2	rst	Yes	*Pseudomonas aeruginosa*		Thioredoxin fold	Homodimer interface	Hypothetical protein Atu0228 (2fa8)
Maf protein	1ex2 A 74-79	4	14.9	0.5	true	1	rst	Yes	*Bacillus*		Anticodon-binding domain-like (ITPase-like)	On β-hairpin at the opposite end of strand 3 to the main fold structure (outside)	Maf homologue Tb11.01.5890 (2amh)
RNase HII	1i3a A 24-57	1	16.0	0.0	end	0	rst	Yes	*Archaeoglobus fulgidus*	1i39 A 24-57	Ribonuclease H-like motif	Connecting strands 2 and 3 at loop end	RNase HII of Methanococcus jannaschii (1eke)
Nucleoside triphosphate diphosphohydrolase #1	3agr A 234-244	2	16.2	0.6	true	0	rst	Yes	Alveolata	Yes	Ribonuclease H-like motif	On β-hairpin immediately before 2nd Ribonucleause H-like unit	Ectonucleoside triphosphate diphosphohydrolase 2 (3cj1)
Nucleoside triphosphate diphosphohydrolase #2	3agr A 35-64	2	14.6	0.8	end	0	rst	Yes	Alveolata		Ribonuclease H-like motif	Connecting strands 1 and 2 of N-terminal domain - N-terminal end	Mlc protein (1z6r)
Pancreatic ribonuclease [g]	2op2 A 43-85	1	12.8	0.0	true	1	rst	Yes	Engineered		RNase A-like	Connecting strands 1 and 2 - outside (N-terminal end)	Ribonuclease A (1dy5)
CinA	3kbq A 122-137	2	15.3	1.4	true	1*	rst	Yes	*Thermoplasma*		Molybdenum cofactor biosynthesis proteins	Connecting strands 4 and 5 of main fold	Gephyrin, domain 5 (2fu3)
Caspase-2	1pyo B289-D289	1	11.5	0.0	true	2	rst	Na	Euteleostomi, not fish		Caspase-like	Homodimer interface	Na
Histidine acid phosphatases, eukaryotic	1qwo A 48-391	13	14.6	5.4	true/end	0/1	rst/lcs	Yes	Eukaryota	2gfi A 62-385	Phosphoglycerate mutase-like	Connecting strands 1 and 5	Phytase of Klebsiella (2wnh)
Histidine acid phosphatases	2wnh A 370-379	10	13.7	0.6	end	1*	rst	Yes	Eukaryota and proteobacteria		Phosphoglycerate mutase-like	Holding together C-terminal β-hairpin (β-hairpin forms β-bridge to strand 6)	Phytase of Aspergillus (1qwo)
TP53-induced glycolysis and apoptosis regulator B, TIGAR B	3e9c A201-B201	2	20.4	1.5	true	2	rst	Na	*Danio rerio*		Phosphoglycerate mutase-like	Homodimer interface	na
SAM-dependent tRNA/rRNA cytosine-C5 methylase	1ixk A 46-81	1	14.7	0.0	true	1	rst	Yes	*Thermococcus*		S-adenosyl-L-methionine-dependent methyltransferases (N-terminal ferriredoxin-like domain)	Connecting strands 1 and 3 of N-terminal ferriredoxin-like domain	NOL1R (2b9e)
Fringe proteins	2j0a A 139-202	2	12.4	1.0	true	2	rst	Yes	Metazoa		Nucleotide-diphospho-sugar transferases	Connecting strands 4 and 6 of N-terminal sheet	Spore coat polysaccharide biosynthesis protein (1qg8)
Mycobacterial antigens	1f0n A 87-92	5	12.8	1.1	true	0	rst	Yes	*Mycobacterium*		α/β Hydrolases	On β-hairpin formed in loop between strand 4 and the following helix	Carboxylesterase (1jkm)
Serine carboxypeptidase-like proteins #1	1ivy A 60-334	12	12.8	1.1	β-bridge	0	rst	Na	Eukaryota		α/β Hydrolases	Connecting loop between strand 3 and following helix to elaborate loop between strands 5 and 6	Not seen - loop between 5 and 6 much shorter in most families
Serine carboxypeptidase-like proteins #2	1ivy A 213-218	2	13.6	0.0	true	0	rst	Na	Metazoa		α/β Hydrolases	On short β-hairpin in loop between strands 5 and 6	Not seen - loop between 5 and 6 much shorter in most families

Tricorn interacting factor F1	1mtz A 5-22	12	12.4	3.3	end	1	rst	Yes	*Thermoplasma*		α/β Hydrolases	Connecting N-terminal end of strand 1 to strand 2	Carboxylesterase (1jkm)
Triacylglycerol lipase B	1tcb A 22-64	15	12.6	0.5	true	1	rst	Na	Fungi + Bacteria		α/β Hydrolases	Connecting strands 2 and 4 near end (strand 1 missing in this protein)	Not seen - not in main fold
Pancreatic lipase, N-terminal domain	1bu8 A 4-10	14	13.5	0.8	true/end	0	rst	Yes	Metazoa		α/β Hydrolases	On β-hairpin in N-terminal loop	Carboxylesterase (1jkm)
Arylsulfatase B	1fsu A 181-192	1	11.9	0.0	true	0	rst	No	Euteleostomi		Alkaline phosphatase-like	On β-hairpin of elaboration between strands 4 and 5	Not seen - elaboration does not involve hairpin if disulfide is not present
Nucleotide pyrophosphatase/phosphodiesterase	2gso A 314-337	8	13.5	0.5	true	2	rst	Yes	Proteobacteria		Alkaline phosphatase-like	Connecting strands 3 and 4 of side domain (central)	Phosphonoacetate hydrolase (1ei6)
Chitin oligosaccharide binding protein	1zu0 A 179-194	2	13.4	0.2	true	2	rst	Yes	α + γ proteobacteria		Periplasmic binding protein-like II	Connecting strands 3 and 4 of N-terminal sheet (2341576) - just before loop	Oligo-peptide binding protein of bacillus subtilis (1xoc)
Oligo-peptide binding protein	1jet A 271-417	32	13.3	0.7	true	0*	rst	Yes	γ proteobacteria		Periplasmic binding protein-like II	Connecting strands 1 and 4 of C-terminal sheet (1'51423) - disulfide in core of protein immediately adjacent to peptide binding site	Oligo-peptide binding protein of bacillus subtilis (1xoc)
Ectoine/hydroxyectoine ABC transporter solute-binding protein	2q88 A 97-201	2	11.7	0.2	true	0	rst	Yes	α proteobacteria		Periplasmic binding protein-like II	Connecting the two strands that run the length of the fold - just below C-terminal helix	Lysine-,arginine-,ornithine-binding (LAO) protein (1lst)
DNA repair helicase XPD	2vl7 A 244-297	1	10.5	0.0	true	1*	rst	No	Sulfolobaceae		*β-α(2)-β-α(2)-β(2); 2 layers; antiparallel β-sheet, order: 1432*	Connecting strands 2 and 3 - outside	Not seen - no other structures with this fold
Acid-sensing ion channel	2qts A 94-195	8	14.0	2.4	end	2	rst	No	Eukaryota		*core: 3 layers, β/β/α; central antiparallel beta-sheet of 5 strands, order 32145; top mixed beta-sheet of 6 strands, order 213654*	Connecting strands 1 and 3 of central sheet, α layer side	No other protein with this fold yet known - likely to be possible to exist without disulf
Acidianus filamentous virus 1 protein 102	2wb6 A 56-62	1	12.8	0.0	true	1	rst	No	Acidianus filamentous virus 1		*intertwined dimer of α-β(2)-α-β(2) motifs; 3 layers, β/α/β; antiparallel β-sheets, order: 1243*	Connecting strands 3 and 4 of mid-sequence sheet - inside	Not seen - no other structures with this fold
Fungal ribonucleases/Ribotoxin	1i0x A 6-103	32	9.8	2.2	end	1*	rst	Na	Dikarya	5hoh A 6-103	Microbial ribonucleases (Fungal ribonucleases/Ribotoxin)	On small section of β structure connecting terminii	Not seen - this part of fold not seen in other families
Destabilase	2dqa A 27-36	3	14.8	0.2	true	1	rst	Yes	Metazoa		Lysozyme-like	Connecting strands 1 and 2	Lysozyme (Gallus gallus) (2bpu)
Cathepsin Z #1	1deu A 112-118	4	17.0	4.2	end	0	rst	Na	Eukaryota		Cysteine proteinases	On additional β-hairpin	Not seen - disulfide not in main fold
Cathepsin Z #2	1deu A 153-235	4	13.3	0.9	true	2	rst	Yes	Eukaryota		Cysteine proteinases	Connecting strands 2 and 6 - non-helix side	Papain (1khq)
Serine-repeat antigen protein	2wbf X 567-572	3	16.5	1.1	true	1	lst	Yes	Aconoidasida		Cysteine proteinases	On N-terminal β-hairpin	(Pro)cathepsin X (1deu)
Adenain bound to pVIc cofactor	1nln A104-B309	2	13.2	0.1	true	0	rst	Na	Adenoviridae		Cysteine proteinases	Intermolecular	na
Amyloid β a4 protein, copper binding domain	2fk1 A 133-187	22	11.7	1.2	end	1/2	rst	Yes	Metazoa		Dodecin subunit-like	Connecting strands 1 and 3 (N to C terminii) - non-helix side	Dodecin (2cc6)
Membrane lipoprotein, rcsF	2y1b A 109-124	1	13.9	0.0	true	1	rst	Yes	γ proteobacteria		Dodecin subunit-like	Connecting strands 3 and 4 - helix side (central)	Dodecin (2cc6)

L-aminoacid oxidase	2iid A 331-412	22	17.9	3.0	true	1*	rst	Yes	Mammals, reptiles, birds		FAD-linked reductases, C-terminal domain	Connecting strands 1 and 5 - non-terminal end, helix side	Polyamine oxidase (1b5q)
Cystatin/Cathelicidin	1roa A 99-119	28	13.1	1.9	true/end	1	rst	Yes	Metazoa		Cystatin-like	Connecting strands 4 and 5 - non-β-hairpin end, helix side	Monellin (2o9u)
Copper amine oxidase, N-terminal region, 2nd domain	1ksi A 137-158	12	12.4	0.8	true	0	rst	Yes	Plants and bacteria		Cystatin-like	Connecting strands 1 and 2 - helix side (central)	Copper amine oxidase of Escherichia coli (1d6z)
Class I MHC homologs	2o5n A 101-108	3	13.5	1.1	true/end	1/2	rst	Yes	Muromegalovirus		MHC antigen-recognition domain	Connecting strands 5 and 6 - helix side (near β-hairpin)	Class I MHC homolog Mic-b (human) (1je6)
Class I MHC homolog Mic-a	1hyr C 36-41	2	20.0	3.0	true	1	rst	Yes	Primates		MHC antigen-recognition domain	Connecting strands 3 and 4 - helix side	Class I MHC homolog Mic-b (human) (1je6)
Class II MHC homolog HLA-DM, β chain, N-terminal domain	2bc4 B 25-35	3	13.2	0.9	true	2	rst	Yes	Mammals		MHC antigen-recognition domain	Connecting strands 2 and 3 - non-helix side	Class II MHC β chain, N-terminal domain HLA-DR1 (human) (1pyw)
Retinoic acid early transcript, RAE-1	1jfm A 7-26	5	13.2	1.1	true	2	rst	Yes	Murinae		MHC antigen-recognition domain	Connecting strands 1 and 2 - helix side (central)	Class I MHC homolog Mic-b (human) (1je6)
Yellow fluorescent protein [g]	1h6r A 149-202	4	12.6	0.9	true	1	rst	Yes	Engineered	Yes	GFP-like	Outside of β-barrel - connecting strands 7 and 10	Wild type (1ema)
Green fluorescent protein [g]	1jc1 A 147-204	4	12.9	1.2	end	1	rst	Yes	Engineered	Yes	GFP-like	Outside of β-barrel - connecting strands 7 and 10	Wild type (1ema)
FKBP13	1u79 A 106-111	3	16.7	2.3	true (A-bent)	0	rst	Yes	Tracheophyta	Yes	FKBP-like	Inside A-bent β-ribbon	FKBP25 (1pbk)
FK506-binding domain of FKBP35	3ihz A105-B105	1	14.2	0.0	β-bridge	0*	lst	Yes	*Plasmodium*		FKBP-like	Homidimer interface (connecting β-ribbons with A-bent bulges)	FKBP13 (1u79)
tRNA(Ile)-lysidine synthase (TilS), substrate-binding domain	2e89 A 253-312	4	22.1	6.0	true/end	1/2	rst/lcs	Yes	Aquificaceae		MesJ substrate recognition domain-like	Connecting strands 1 and 4 - outside	TilS substrate-binding domain in E. coli (1ni5)
Fluoroacetyl-CoA thioesterase, FIK	3kv7 A73-B73	22	17.0	1.3	true	2	lst	Yes	*Streptomyces cattleya*		Thioesterase/thiol ester dehydrase-isomerase	Homodimer interface	Hypothetical protein TTHA0967 (2cwz)
Integrin β, tail domain	3ije B 663-687	1	15.2	0.0	end	0	rst	No	Metazoa		Integrin β tail domain	Connecting strands 3 and 4 - helix side (opposide end to β-hairpin)	Not seen but unlikely to be impossible
Killer toxin KP6, α-subunit	1kp6 A 16-74	1	13.1	0.0	true	1	rst	Yes	*Ustilago maydis* P6 virus		Ferredoxin-like	Connecting strands 1 and 4 - helix side (central)	D-ribose-5-phosphate isomerase, lid domain (1o8b)
IS200 transposase, orfA	2ec2 A 49-67	6	16.3	1.4	end	0	rst	Yes	Crenarchaeota	2f5g A 49-67	Ferredoxin-like	Connecting strands 2 and 3 of main fold - inside (opposite end to β-hairpin)	ISHP608 transposase (2vjv)
Insulinase family protease, yhjJ	3go9 A 328-337	1	15.0	0.0	true	1	rst	Yes	Enterobacteriales		LuxS/MPP-like metallohydrolase	Connecting strands 3 and 4 - helix side	Cytochrome bc1 core subunit 1 (1kb9)
3-phosphoshikimate 1-carboxyvinyltransferase	3roi A 186-216	2	12.8	0.2	true	1/2*	rst	Yes	γ proteobacteria		IF3-like	Non-terminal sub-domain, 3rd subunit, connecting strands 2 and 4 - outside	EPSP synthase (1rf6)
Type IV pilin, PilX	2ope A 125-144	6	13.8	0.9	true/end	1	rst	Na	Neisseria		MK0786-like	Connecting additional C-terminal strand to strand 4	Not seen - no structures found with additional strand
Coronavirus spike proteins	2ghw A 366-419	12	13.7	2.1	true/end /β-bridge	0/1	rst	No	Coronavirinae	2ghw C 366-419	SARS receptor-binding domain-like	Connecting strands 2 and 4	Not seen - no other proteins with this fold

Putative nitroreductase TM1586	1vkw A 162-201	1	11.5	0.0	true	2	rst	Yes	*Thermotoga*		FMN-dependent nitroreductase-like	Connecting strands 2 and 4 of second domain - non-helix side	NADH oxidase (1nox)
Peptide chain release factor subunit 1	3agk A 97-108	1	67.0	0.0	true	1	rcs	Yes	Thermoprotei		N-terminal domain of eukaryotic peptide chain release factor subunit 1, ERF1	Connecting strands 2 and 3 - same side as helicies 2 and 3	N-terminal domain of eukaryotic peptide chain release factor subunit 1, ERF1 (1dt9)
Angiotensin converting enzymes	1o86 A 152-158	22	15.8	3.9	true/end	0	rst	Na	Metazoa	2c6n A 128-136†	Zincin-like	On β-hairpin elaboration in middle of helix 3	Not seen - disulfide not in main fold
Clostridium neurotoxins	1epw A 436-445	9	13.4	1.2	true/end	1	rst	Na	*Clostridium*	Yes	Zincin-like	On β-hairpin in inter-domain region	Not seen - disulfide not in main fold
β-Lactamase-inhibitor protein-I, BLIP-I #1	3gmu B 30-42	26	13.9	2.0	true	1	rst	Yes	*Streptomyces*	2b5r C 1030-1042	BLIP-like	AB - helix side	Unnamed protein (3elg)
β-Lactamase-inhibitor protein-I, BLIP-I #2	3gmu B 109-131	29	14.5	1.8	true/end	1	rst	Yes	*Streptomyces*		BLIP-like	EF - helix side	Unnamed protein (3elg)
Hypothetical protein BT3689	2hqy B 92-294	1	13.8	0.0	true	1*	rst	Yes	*Bacteroides*		Acyl-CoA N-acyltransferases (Nat)	Connecting strands 5 and 7 of N-terminal domain	Aminoglycoside 6'-N-acetyltransferase (1n71)
Putative acetyltransferase Ta0821	3c26 A 173-182	1	15.5	0.0	true	0	rst	Na	*Thermoplasma acidophilum*		Acyl-CoA N-acyltransferases (Nat)	On β-hairpin of elaboration between strands 2 and 3 of C-terminal β-sheet	Not seen - not in main fold
Gelsolin/Macrophage capping protein, Cap G/Adseverin	1kcq A 188-201	6	14.6	1.0	true	1*	rst	Yes	Euteleostomi	Yes	Gelsolin-like	CD - facing helix	Severin (1svy)
PR-1-like proteins	1rc9 A 129-145	29	17.5	2.7	true	1	rst	Yes	Eukaryota		PR-1-like	Connecting strands 2 and 3	Golgi-associated PR-1 protein (1smb)
Methionine aminopeptidase 2	1qzy A 228-448	16	15.0	1.5	true	1*	rst	Yes	Eukaryota	1b6a A 228-448	Creatinase/aminopeptidase	Connecting loop between strands 1 and 2 to β-haripin between strands 5 and 6	Methionine aminopeptidase (2ggc)
Bifunctional DNA primase/polymerase, N-terminal domain	1ro2 A 185-206	4	15.1	0.6	true	0	rst	Yes	Sulfolobaceae		Prim-pol domain	On β-hairpin connecting the C-terminal strands of the two sheets	DNA primase (1g71)
NAD:arginine ADP-ribosyltransferase, ART	1gxy A 121-173	8	13.0	0.4	end	0	rst	Yes	Euteleostomi		ADP-ribosylation	Connecting strands 1 and 2 of main section of sheet - outer suface	Diphtheria toxin, N-terminal domain (1f0l)
Poly(ADP-ribose) polymerase 14, PARP14	3se2 A1537-C1537	2	13.0	0.4	β-bridge	2	rst	Yes	Euteleostomi		ADP-ribosylation	Homidimer interface	Poly(ADP-ribose) polymerase (2rd6)
NK cell receptor subfamily A / Tetranectin	3c8j A 145-150	18	16.0	4.8	true	1/2	rst	Yes	Metazoa		C-type lectin-like	Connecting first two N-terminal β strands	Galactose-specific C-type lectin (1jzn)
C-type lectin (venom and eggshell)	1jzn A 38-133	10	13.1	0.6	true/end	1	rst	Yes	Euteleostomi		C-type lectin-like	Connecting short β strand between the two helicies to C-terminal strand	CD44, hyaluronan binding domain (2jcq)
CD44, hyaluronan binding domain	2jcq A 32-134	5	12.2	1.1	end	0	rst	Na	Euteleostomi		C-type lectin-like	Connecting first N-terminal β strand to short β section on C-terminal extension	Not seen - disulfide not in usual fold
Collagen IV, noncollagenous (NC1) domain	1t60 A 64-70	96	13.4	1.1	true (A-bent)	2	rst	Na	Metazoa	Yes	C-type lectin-like	On β-hairpin protruding from domain	Not seen - disulfide not in usual fold
Hemolytic lectin CEL-III, C-terminal domain	1vcl A 298-380	6	12.5	0.4	true	1	rst	No	Cnidaria + echinodermata		Hemolytic lectin CEL-III, C-terminal domain	Connecting first and fifth β strands	Not seen - no other proteins with this fold
Envelope glycoprotein gp120 #1	2nxy A 126-196	17	19.3	4.0	true/end	0	rst	No	Primate lentiviruses	Yes	gp120 core	Inside β-barrel-like region	Not seen
Envelope glycoprotein gp120 #2	2nxy A 218-247	31	15.6	2.0	end/ β-bridge	0/1	rst	No	Primate lentiviruses		gp120 core	On β-ribbon	Not seen
Envelope glycoprotein gp120 #3	2nxy A 296-331	32	14.0	3.0	true/end β-bridge	1/2	rst	No	Primate lentiviruses		gp120 core	On β-ribbon	Not seen

Envelope glycoprotein gp120 #4	2nxy A 385-418	32	22.4	11.0	true	1/2*	rst/lst/lcs	No	Primate lentiviruses	Yes	gp120 core	Inside β-barrel-like region	Not seen
Envelope glycoprotein gp120 #5	2ny2 A 123-431	2	15.1	0.1	end	1*	rst	Yes	Engineered		gp120 core	Connecting β-ribbon containing #2 to adjacent β-hairpin	Wild type (2nxy)
Major envelope glycoprotein gp64 #1	3duz A 177-184	1	14.0	0.0	true	0	rst	No	Baculoviridae		*2 intertwined domains; all-β and α+β*	On β-hairpin facing middle sheet (disulfide facing #3)	No other proteins structures with this fold
Major envelope glycoprotein gp64 #2	3duz A 228-246	1	14.6	0.0	true	2	rst	No	Baculoviridae + Orthomyxoviridae		*2 intertwined domains; all-β and α+β*	On end 4 strand sheet - facing middle 4 strand sheet	No other proteins structures with this fold
Major envelope glycoprotein gp64 #3	3duz A 262-267	1	17.2	0.0	true	1	rst	No	Baculoviridae + Thogotovirus		*2 intertwined domains; all-β and α+β*	On middle 4 strand sheet (disulfide facing #1)	No other proteins structures with this fold
CagD	3cwx A172-B172	1	10.4	0.0	true	0*	rst	Na	*Helicobacter pylori*		*α-β(5)-α(2); meander β-sheet (closely related to UBC-like, additional β strand means different alignment between α and β layers)*	Homodimer interface	Na
YfeY	2qzb A 137-150	2	14.2	0.7	true	1*	rst	No	Bacteria		*antiparallel β-sheet of 8 strands, order 12348765; second layer has α and β components*	Connecting strands 5 and 6 - inside	Not seen - no other structures with this fold
Perforin-1	3nsj A 22-75	1	13.6	0.0	true	1	rst	Na	Euteleostomi		*Complex α+β fold*	Connecting N-terminus to β-ribbon in N-terminal region	Not seen - no other structures with this part of fold
Hypothetical protein PH0156	2p62 A203-B203	1	17.9	0.0	β-bridge	0*	rst	Na	*Pyrococcus horikoshii*		PH0156-like	Short β-bridge forming homodimer interface	Na
Class D β-lactamase	1k55 A 44-51	74	15.7	2.5	true/end	1/2*	rst	Yes	γ proteobacteria		β-lactamase/transpeptidase-like	Connecting strands 2 and 3	Class D β-lactamase of Escherichia coli (1m6k)
Glycoprotein B	2gum B 116-573	13	16.6	5.4	end	0	rst	Na	Herpesviridae		Viral glycoprotein ectodomain-like	Connecting N-terminus to linker region between helical region and C-terminal β domain	Not seen - in variable part of fold
HasR protein	3ddr A 715-750	6	13.5	3.5	true/ β-bridge	0	rst	Na	Proteobacteria		Transmembrane β-barrels (Porins)	On elaboration between strands 17 and 18	Not seen - not in main fold
Ferric hydroxamate uptake receptor, FhuA	2fcp A 701-707	8	12.5	0.6	true	0/1	rst	Yes	Proteobacteria		Transmembrane β-barrels (Porins)	Connecting strands 21 and 22 of β-barrel (bulged strand and C-terminal strand) - outside (near β-hairpin)	Ferric enterobactin receptor FepA (1fep)
Flavivirus envelope glycoprotein	1ok8 A 92-116	18	13.4	1.5	true	0/1	rst	Yes	Flaviviridae		Viral glycoprotein, central and dimerisation domains	On 3 strand sheet at opposite end of molecule to terminii	Fusion glycoprotein E1 (2ala)
Flavivirus envelope glycoprotein	1ok8 A 185-285	18	14.1	1.5	true/end	1	rst	Yes	Flaviviridae		Viral glycoprotein, central and dimerisation domains	Connecting strands 7 and 8 of β domain containing terminii	Fusion glycoprotein E1 (2ala)
Flavivirus polyprotein propeptide	3c5x C 34-68	2	15.2	0.7	true/end	2	rst	No	Flaviviruses		*sandwich; 7 strands in 2 antiparallel sheets, top sheet 1275, bottom sheet 346*	Connecting strands 4 and 6 - inside	Not seen - no other structures with this fold
Thrombomodulin, fifth EGF-like domain	1dx5 I 399-407	5	14.7	1.0	true	0	rst	Yes	Metazoa		Knottins	On β-hairpin of central domain	Epidermal growth factor, EGF (1nql)
Bowman-Birk inhibitors	2g81 I 24-32	30	13.7	1.7	true/end	0/1/2	rst	Yes	Eudicotyldons, not liliopsida	2r33 A 24-32	Knottins	One disulfide across β-hairpin of each BBI domain in fold	C-terminal BBI domain in barley (2fj8)
PMP inhibitors	1gl1 I 17-28	7	15.1	2.0	true/end	2	rst	No	Metazoa		PMP inhibitors	Connecting strands 2 and 3 (central)	Not seen
Ly-6 antigen / uPA receptor -like proteins	2fd6 U 6-12	15	14.0	2.3	true/end	1	rst	Yes	Euteleostomi		Snake toxin-like	Connecting strands 1 and 2	Type II activin receptor (1nys)

Defensin HNP-3	1dfn A 3-31	49	14.9	3.0	true/end	1/2	rst	Yes	Metazoa	1zmp A 3-31	Defensin-like	Connecting terminii	Beta-defensin, BD (1ijv)
Crambin-like proteins	1jxt A 4-32	22	12.3	1.0	end	0/1	rst	No	Eukaryota + bacteria	2v9b B 4-32	Crambin-like	Connecting two β strands in core of structure	Not seen
Plant proteinase inhibitors	1pju A 37-73	11	14.7	1.9	end	0	rst	Yes	Tracheophyta		Plant proteinase inhibitors	Connecting central 2 strands of 4 strand β-sheet (some proteins have strand on either side of central pair missing)	Multidomain proteinase inhibitor (1ce3 - disulfide at opposite end of β-ladder)
Noggin	1m4u A 207-215	1	11.8	0.0	true	0	rst	Yes	Eukaryota		Cystine-knot cytokines	At β-hairpin end of second β-ribbon	Neurotrophin 3 (1bnd)
2-deoxy-scyllo-inosose synthase, btrC	2gru A 7-12	4	12.8	1.1	true	0	rst	Yes	Bacillales		Dehydroquinate synthase-like	Connecting strands 1 and 2 - outside (near β-hairpin)	Dehydroquinate synthase, DHQS (1sg6)
Outer membrane protein, ORF4	3dzm A 159-177	3	16.1	2.1	end	0	rst	No	*Thermus*		Transmembrane β-barrels (OMPA-like)	Connecting strands 7 and 8 - outside (sheet side of juction between sheet-like and barrel-like regions)	Not seen - no good matches to this region of structure
Fusion glycoprotein F0 #1	3rrr F 358-367	9	17.7	6.3	true/end	0/1	rst	No	Mononegavirales		Head and neck region of the ectodomain of NDV fusion glycoprotein	Connecting strands 5 and 6 of 6 strand sheet in centre of molecule (containing N-terminal strand)	Not seen - no structures found with this fold but no disulfide
Fusion glycoprotein F0 #2	3rrr B 322-333	3	23.2	12.4	true/end	0/1	rst	Yes	Pneumovirinae		Head and neck region of the ectodomain of NDV fusion glycoprotein	Connecting strands 2 and 3 of 6 strand sheet in centre of molecule (containing N-terminal strand)	F0 in Paramyxovirinae (2b9b)
Fusion glycoprotein F0 #3	3rrr B 313-343	6	45.9	47.4	true/end	1/2*	rst/rcs	Yes	Pneumovirinae		Head and neck region of the ectodomain of NDV fusion glycoprotein	On long β-hairpin between strands 2 and 3 of 6 strand sheet in centre of molecule	F0 in Paramyxovirinae (2b9b)
Fibronectin I-like domain	3cal A 92-104	46	12.9	1.5	true	1/2	rst	No	Metazoa	3m7p A 543-555†	FnI-like domain	Connecting strands 2 and 3 of 3 strand sheet	Not seen
β-microseminoprotein, N-terminal domain	3ix0 A 2-50	4	15.7	1.9	true	2	rst	No	Metazoa		FnI-like domain	Connecting strands -1 and 5 (in other structures N-terminus is loose, ie. strand -1 doesn't exist)	Not seen - strand 1' loose when not disulfide bonded to strand 5
β-microseminoprotein, N-terminal domain	3ix0 A 40-49	4	12.3	0.9	true	2	rst	No	Metazoa		FnI-like domain	Connecting strands 4 and 5	Not seen - disulfide found in all structures
Crossveinless 2 and BMP-binding endothelial regulator protein	3bk3 C 29-38	2	14.8	0.4	true	0	rst	No	Metazoa		FnI-like domain	Connecting strands 2 and 3 - central = 184.5	Not seen - disulfide found in all structures
Thyroglobulin type-1 domain	1icf I 227-234	6	13.3	1.4	true	0/1	rst	No	Eukaryota + bacteria	2dsq H 192-203	Thyroglobulin type-1 domain	On central β-hairpin	Not seen
Carboxypeptidase inhibitor #1	1dtd B 11-34	5	12.4	1.3	true	0/1	rst	No	*Hirudo medicinalis*		Carboxypeptidase inhibitor	Connecting strands 1 and 3 - non-helix side - immediately adjacent to #2	Not seen - no other proteins in with this fold
Carboxypeptidase inhibitor #2	1dtd B 18-62	4	13.7	0.3	true	1	rst	No	*Hirudo medicinalis*		Carboxypeptidase inhibitor	Connecting strands 2 and 4 - non-helix side - immediately adjacent to #1	Not seen - no other proteins in with this fold
Carboxypeptidase inhibitor #3	1dtd B 22-58	4	12.1	0.2	true	1/2	rst	No	*Hirudo medicinalis*		Carboxypeptidase inhibitor	Connecting strands 2 and 4 - non-helix side	Not seen - no other proteins in with this fold
Resistin	1rgx A 35-88	24	13.1	1.4	true	1	rst	No	Metazoa		Resistin	Connecting N- and C-terminal strands	Not seen - no other proteins in with this fold

Hemagglutinin	1jsd A4-B137	170	14.7	5.6	true	0/1/2	rst	No	Influenza viruses	Stalk segment of viral fusion proteins	Connecting strands 1 and 4 (strand 1 being N-terminus of viral protein domain) - outside	Not seen - most non-influenza stalk structures show only the helicies
Erythropoietin mimetic peptides	1eba C 6-15	2	18.1	2.4	true/end	0/1	rst	Na	Engineered	Erythropoietin (EPO) mimetic peptides	β-hairpin	Na
Hepcidin	3h0t C 7-23	1	20.6	0.0	end	0	rst	No	Euteleostomi	Antimicrobial β-hairpin	On β-ribbon connecting termini	Not seen
Hepcidin	3h0t C 11-19	1	12.6	0.0	end	0	rst	No	Euteleostomi	Antimicrobial β-hairpin	Outside of A-bent bulge - non-β-hairpin side	Not seen
Complement α chain-like proteins #2	2hr0 B 1484-1489	9	18.5	12.6	true	0	rst/lcs	Na	Metazoa	β-hairpin	On β-hairpin connecting A2M receptor and NTR domains	Na
Engineered peptide	1dn2 E 2-12	2	13.8	0.2	true	0	rst	Na	Engineered	β-hairpin	β-hairpin	Na
BR3 derived peptide	1osg G 23-34	4	16.5	3.9	end	0	rst	Na	Engineered	β-hairpin	β-hairpin	Na
Angiopoietin receptors	2gy5 A 280-287	1	14.4	0.0	true (A-bent)	1	rst	No	Metazoa	β-hairpin	β-hairpin between two series of EGF-type domains	Not seen