Methods

Database construction

Databases of protein sequence and secondary structure annotations were assembled as follows. Sequences and secondary structures of all proteins in the PDB (updated 05/31/2017) were determined using both DSSP (1) and PROSS (2). DSSP determines secondary structure through hydrogen bonding patterns while PROSS identifies secondary structure through backbone dihedral angles. The resulting sets of secondary structure annotations and their corresponding amino acid sequences were updated to include chain breaks and unobserved regions using a metafile from the PDB (https://cdn.rcsb.org/etl/kabschSander/ss_dis.txt.gz). Six fold-switching proteins were identified using DSSP that were not identified with PROSS; 27 were identified through PROSS but not found using DSSP; the remaining 64 were identified using both libraries. All sequences in the database were used except those with continuous stretches of ≥4 amino acids annotated as X (0.05% of the database).

Classification of fold switches

Proteins with different oligomeric states were assigned to class A, B, or C as follows. Those reported to hetero-oligomerize with another macromolecule (protein, DNA, or RNA) in one conformation, or two different macromolecules in both conformations, were assigned to class A. Those with a ≥30% increase in number of intermolecular hydrogen bonds were assigned to class C. The remainder all showed a

≥30% increase in number of intermolecular hydrophobic contacts, and were assigned to class B. Hydrogen bonds were calculated using the criteria of Kortemme, et al. (3), and hydrophobic contacts were calculated as any carbon atoms within 4.5Å of one another. The oligomeric state of the PDB's "biological assembly" was assumed to be its biologically relevant conformation; if more than one oligomeric state was provided, the literature was searched for mention of oligomeric state changes between the fold-switch pair. The rest of the fold-switching proteins (those that did not change oligomeric state) were assigned to class D.

Functional, organismal, and date-solved distributions

Protein functions (**Fig. 1**) were drawn from the HEADER of their PDB files. Similarly, the organism from which these proteins originated was taken from SOURCE ORGANISM_SCIENTIFIC in the PDB files. For comparison, we used the same method to determine the distribution of all source organisms from a non-redundant subset of the PDB (<75% sequence identity, <10 Å, R < 0.3, non-X-ray structures included; (4)). These distributions were compared in **Fig. S3**. To eliminate selection bias, proteins with both the same function and trigger were excluded. This led to a decrease in the number of viral fusion proteins (8 -> 2), β-pores (7->2), serpins (4->1), and MacA (2->1) included in the organismal distribution. All fibril-forming proteins were counted because they were formed by proteins with disparate functions. Release dates were used to measure the publication age of a protein structure. For each fold switch pair, the newer of the two dates was used in the distribution of solved-by dates (**Fig. S4**). This is because the newer of the two structures was required to demonstrate that a fold switch occurred.

Identification of independent folding units

Calculations of independent folding units (IFUs) were performed with SEED (5). Because this method is computationally intensive, the protein regions corresponding to fold-switching segments were given as reference for the IFU search. SEED then searched for IFUs by calculating the qualifying ratio (QR, the measure by which SEED determines independent folding cooperativity) of the reference sequence. The QR of the reference sequence was calculated. If it exceeded 0.78, SEED searched for a maximum QR by extending the reference sequence up to 15 residues both N-terminally and C-terminally (for a total extension of up to 30 residues) in increments of 1. If the QR of the reference <0.78, SEED searched for a segment containing the fold-switching region with a QR ≥0.78. To do this, it extended the C-terminus of the reference sequence up to 50 residues in increments of 1; if no segment with QR ≥0.78 was found, it extended the N-terminus in increments of 1 followed by C-terminal extension of to 50 residues in increments of 1; this continued until either a segment with a QR≥0.78 was found or both the N- and C-termini were extended by 50 residues. If a segment with QR ≥0.78 was found, SEED searched for a maximum QR by extending the segment up to 15 residues both N-terminally and

C-terminally (for a total extension of up to 30 residues). The maximized regions with QRs ≥0.78 were reported (**Table S2**).

To test the statistical significance of the QR values corresponding to the IFUs above, we calculated the maximal QR of a segment of specified length containing the randomly-selected reference sequence. Specifically, we randomly selected 96 reference segments from the non-redundant PDB (<75% sequence identity, <10 Å, R < 0.3, non-X-ray structures included; (4)) whose lengths corresponded one-to-one with the reference segments in set of 96 fold switchers. Using SEED, we then calculated the QR of every segment containing the random reference segments whose length equaled the SEED-calculated IFU corresponding to the fold-switch reference segment. The IFU with the maximum QR was chosen as the best IFU containing the randomly-selected reference segment. For example, if a reference sequence had 40 residues (residues 30-69 of a 100-residue protein), and its corresponding IFU had 50 residues, SEED would calculate the QRs of residues 30-79, 29-78, 28-77…20-69 and select the IFU with the highest QR. We chose this approach to preserve IFU length because QRs tend to increase with sequence length. The same 96-protein simulation was repeated 10 times, and the QRs from all 10 simulations were used to make the distribution in **Fig. 2**. Using the Kolmogorov-Smirnov test, we found that the randomly-generated distribution and fold switch distribution differed significantly: $p<10^{-34}$. Occasionally, simulations on a given protein failed due to incompatibilities with size or significant chain breaks. We counted these failed simulations as correct predictions, giving our random simulations the benefit of the doubt. Had we counted the failed simulations as incorrect predictions, $p<10^{-37}$.

Whole-PDB fold switcher predictions

SEED calculations were performed on all protein regions from a non-redundant subset of the PDB (<75% sequence identity, ≤3.0Å, R < 0.3, non-X-ray structures included (4)) whose secondary structure annotations from SPIDER2 (6) differed substantially from their experimentally determined secondary structure annotations from PROSS (2), where substantial differences were the same as those defined in *Identification...* (**Methods**, main text). This yielded 11281/35060 unique protein structures, or 32% of the PDB.

To get the lower bound estimate of possible fold-switching proteins in the PDB, we searched the non-redundant subset of the PDB for homologs of the 96 fold switchers and other proteins with related keywords in their PDB files (e.g. viral fusion proteins or pore-forming toxins). We also performed a text search of all abstracts in the PDB to identify key terms associated with fold switchers (“structural rearrangement”, “refold”, “prion”, “amyloid”, ”pre-fusion”, ”post-fusion”, ”prefusion”, ”postfusion”, and “dramatic”). We then read the relevant abstracts containing these key terms to identify fold switchers with literature support. We read full papers when further clarification was needed. Combining these approaches, we found a total of 92 additional fold switchers **(Table S3A)**. These 92 combined with the 96 tabulated switchers constituted 0.5% of the non-redundant PDB.

SI figures.

A

Dynamic protein (not fold switch)

Adenylate kinase (4akeA apo, 1akeA +Ap5A)

```
EEEEE  TT  TTTTHHHHHHH  EEEEEHHHHH
EEEEE  TT  HHHHHHHHHHHT   EEEHHHHH

HHHHH     HHHHHHHHHHHT     HHHHHHHH
HHHHHH       HHHHHHHHH     HHHHHHHH

HHHHH TTTTT        EEE         HHHH
HHHHH TTTTT        EEE         HHHH

HHHHHHT EEE EEEEEEE HHHHHHHHTTEEEE
HHHHHHT EEE EEEEEEE HHHHHHHHHTEEEET

  EEETTT     TT  TTT   EEEETT  HHHH
TTEEETTT     TT  TTT   EEEETT  HHHH

HHHHHHHHHHHHHHHHHHHHHHHH  EEEEEEETT
HHHHHHHHH  HHHHHHHHHHHHH  EEEEEEETT

  HHHHHHHHHHHH
  HHHHHHHHHHHH
```

Fold switch

KaiB (5jytA +KaiC, 4ksoA tetramer)

```
EEEEEEEEEETT HHHHHHHHHHHHHHH TTTT
EEEEEEEEEETT HHHHHHHHHHHHHHH TT

HHHHHHT      EEEE        EEEE  TT
      EEEE HHHHHHT   HHHHHHHHHH

 HHHHHHHHHHHHHHHHH
         EEE
```

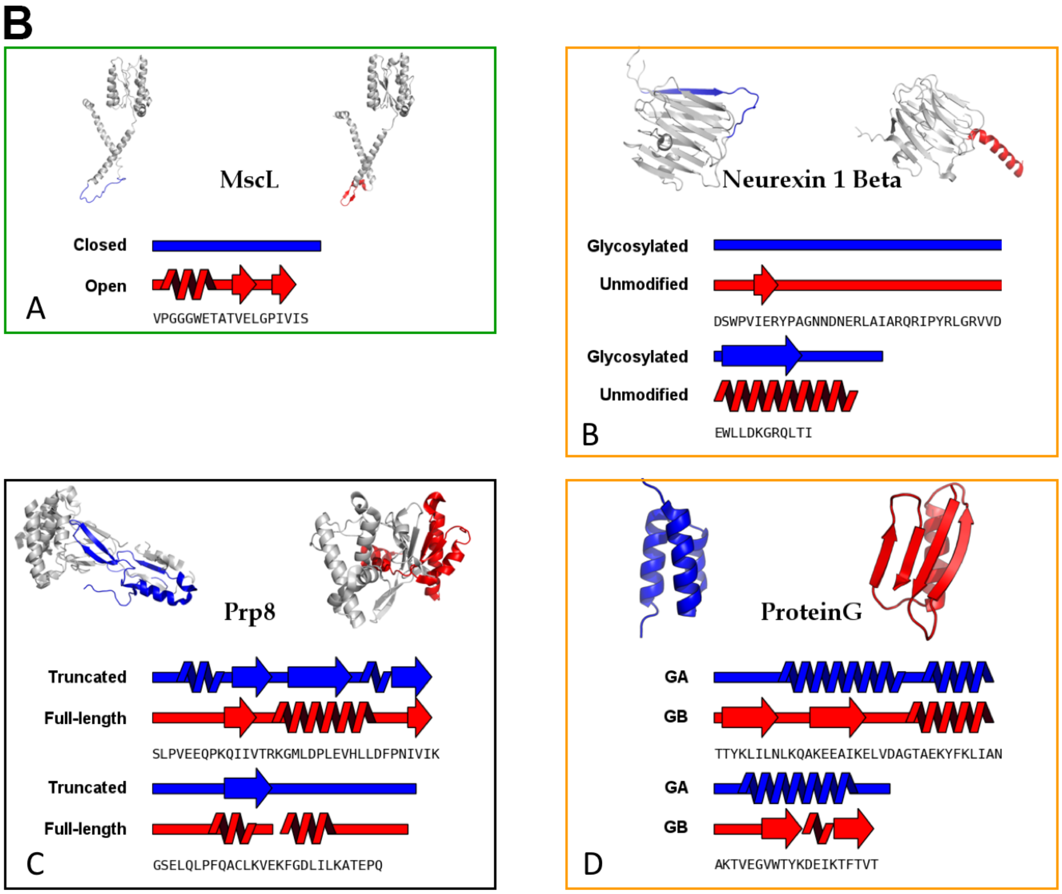
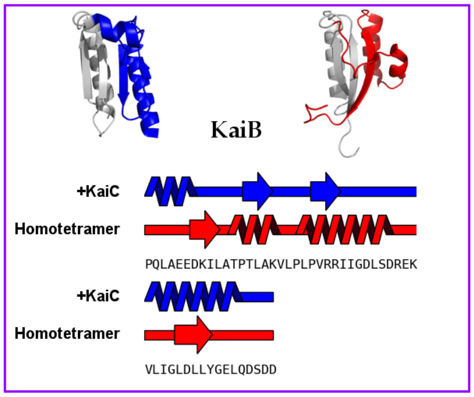
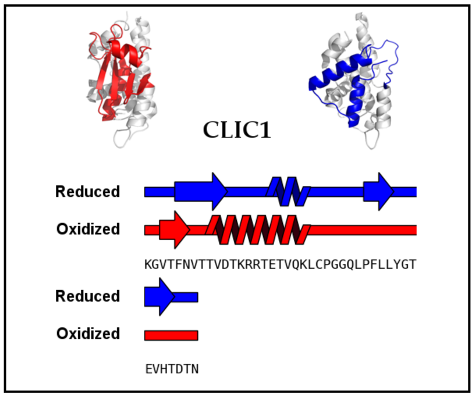


Fig. S1. Examples of proteins not classified as fold switchers. **A.** Secondary structure alignments of a dynamic protein and a fold-switching protein. Although dynamic, the secondary structure annotations of adenylate kinase (left) are essentially identical in both conformations (apo and in complex with Ap5A, di(adenosine-5') pentaphosphate). Therefore, our algorithm does not recognize it as a fold-switching protein. In contrast, the secondary structure alignment of the two KaiB conformations (right) show substantial differences (bold). Thus, our algorithm flagged KaiB as a potential fold-switching protein. **B.** Examples of rejected fold-switch "hits". A. MscL was rejected because it undergoes a transition between coil and regular secondary structure. B. Neurexin 1β was rejected because there is NMR evidence suggesting that the red conformation is an artifact of crystal packing (7), not biologically relevant. C. Prp8 was rejected because the sequence of the blue form was truncated and does not necessarily reflect a biologically relevant variant (8). D. The engineered variants of protein G were rejected both because they do not occur naturally and because they require an amino acid change to trigger the conformational switch (9).

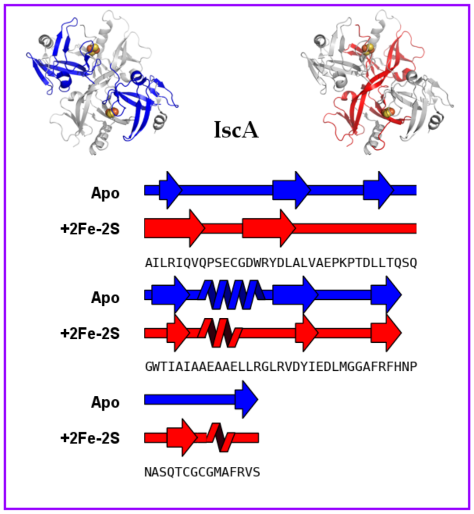
1.


KaiB
+KaiC
Homotetramer
PQLAEEDKILATPTLAKVLPLPVRRIIGDLSDREK
+KaiC
Homotetramer
VLIGLDLLYGELQDSDD

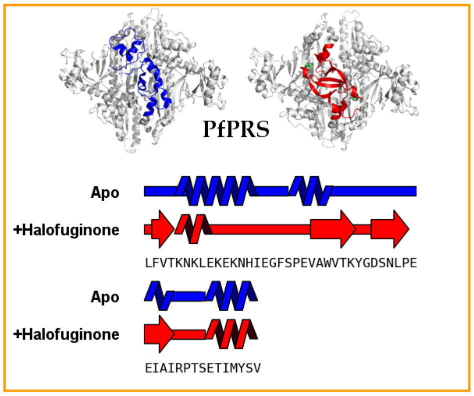
2.


CLIC1
Reduced
Oxidized
KGVTFNVTTVDTKRRTETVQKLCPGGQLPFLLYGT
Reduced
Oxidized
EVHTDTN

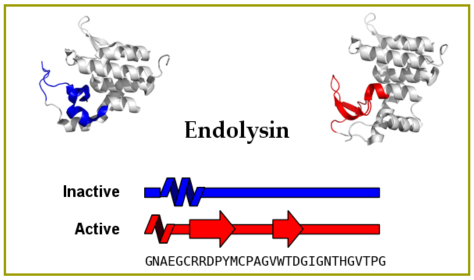
3.


IscA
Apo
+2Fe-2S
AILRIQVQPSECGDWRYDLALVAEPKPTDLLTQSQ
Apo
+2Fe-2S
GWTIAIAAEAAELLRGLRVDYIEDLMGGAFRFHNP
Apo
+2Fe-2S
NASQTCGCGMAFRVS

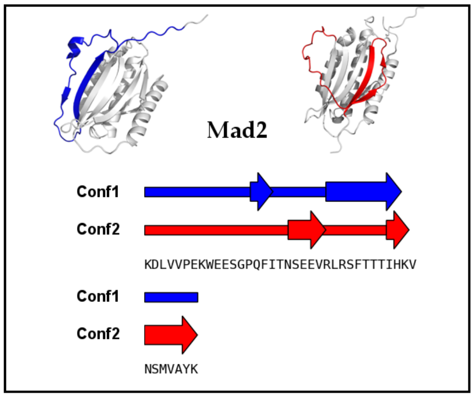
4.


PfPRS
Apo
+Halofuginone
LFVTKNKLEKEKNHIEGFSPEVAWVTKYGDSNLPE
Apo
+Halofuginone
EIAIRPTSETIMYSV

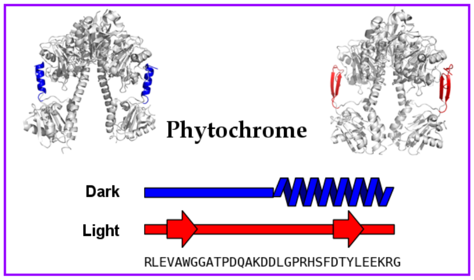
5.


Endolysin
Inactive
Active
GNAEGCRRDPYMCPAGVWTDGIGNTHGVTPG

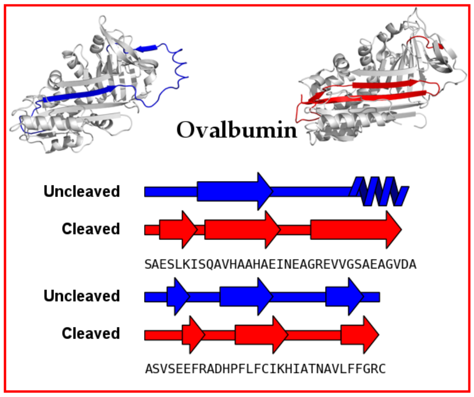
6.


Mad2
Conf1
Conf2
KDLVVPEKWEESGPQFITNSEEVRLRSFTTTIHKV
Conf1
Conf2
NSMVAYK

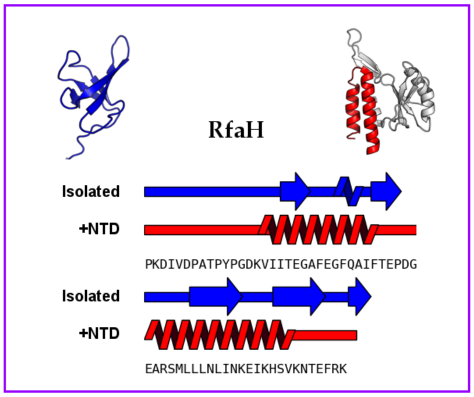
7.


Phytochrome
Dark
Light
RLEVAWGGATPDQAKDDLGPRHSFDTYLEEKRG

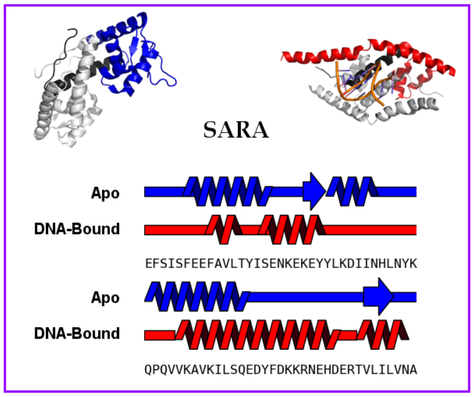
8.


Ovalbumin
Uncleaved
Cleaved
SAESLKISQAVHAAHAEINEAGREVVGSAEAGVDA
Uncleaved
Cleaved
ASVSEEFRADHPFLFCIKHIATNAVLFFGRC

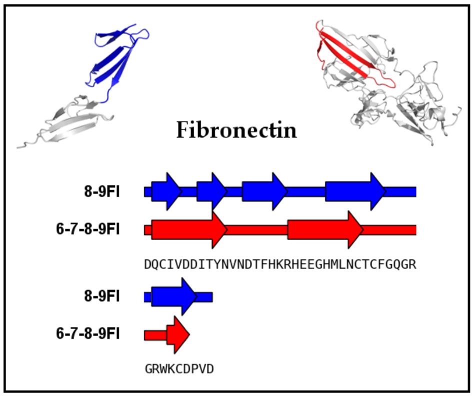
9.



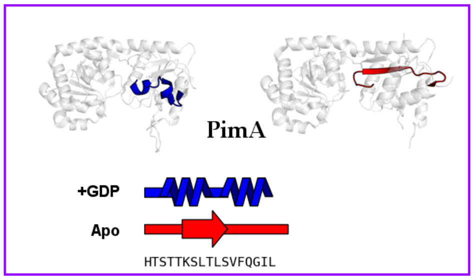
10.



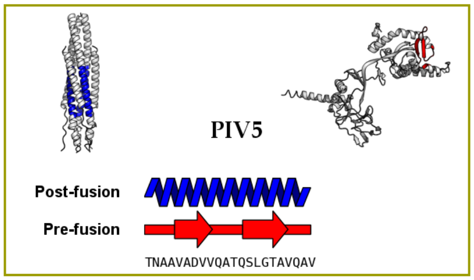
11.



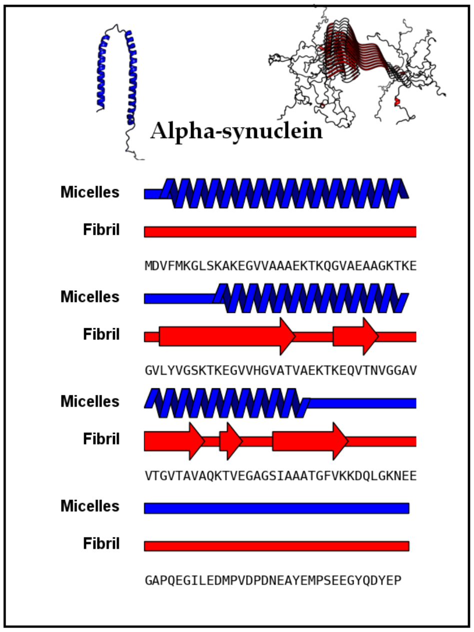
12.



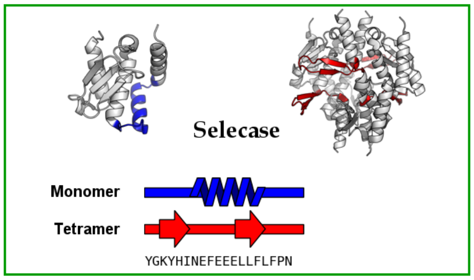
13.



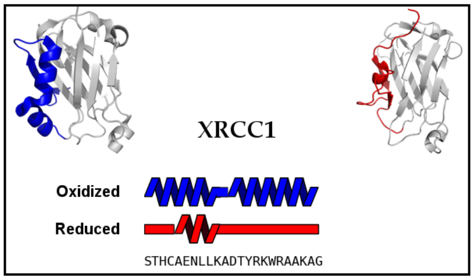
14.



15.



16.



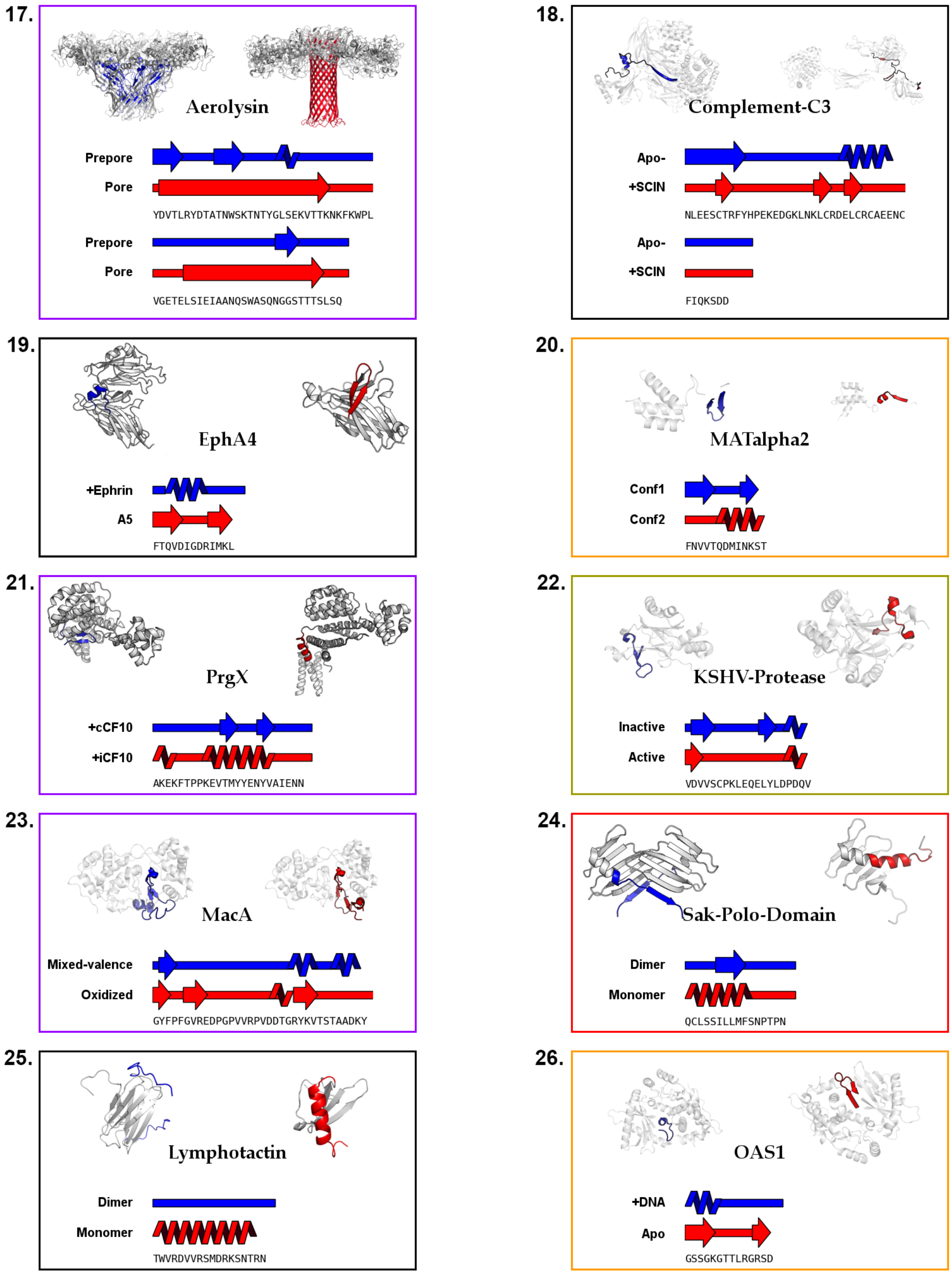


Fig. S2. Representative fold-switching proteins. Two conformations and their corresponding secondary structures are shown in red and blue, respectively. Protein regions that do not switch folds are shown in gray. The common name of the protein is shown between the structures, and the names or triggers of each conformation are shown to the left of the secondary structure diagrams. The colored boxes around each protein represent their organism of origin; black: primates; red: other vertebrates; orange: single-celled eukaryotes; purple: bacteria; green: archaea; olive: virus. Corresponding PDB IDs are as follows: 1. 5jytA, 4ksoA; 2. 1k0nA, 1rk4A; 3. 1x0gA, 1x0gB; 4. 4twaA, 4ydqA; 5. 1xjuA, 1xjtA; 6. 3gmhL, 2vfxL; 7. 4o01A, 4o0pA; 8. 1ovaA, 1jtiA; 9. 2lclA, 2ougD; 10. 2frhA, 2fzpA; 11. 3ejhA, 3m7pA; 12. 4nc9C; 4n9wA; 13. 1svfA, 4wsgC; 14. 2kkwA, 2n0aD; 15. 4qhfA, 4qhhA; 16. 3lqcA; 1xntA. 17. 5jzhA, 5jztG; 18. 2a73B, 2l5nB; 19. 4m4rA, 4w50B; 20. 1mnmC, 1mnmD; 21. 2axzA, 2grmB; 22. 3njqA, 2pbkB; 23. 2cu1C, 2c1vB; 24. 1mbyA, 2n19A; 25. 2n54A, 1j8iA; 26. 4rwnA, 4rwqA. All protein structures were made using PyMOL (10).

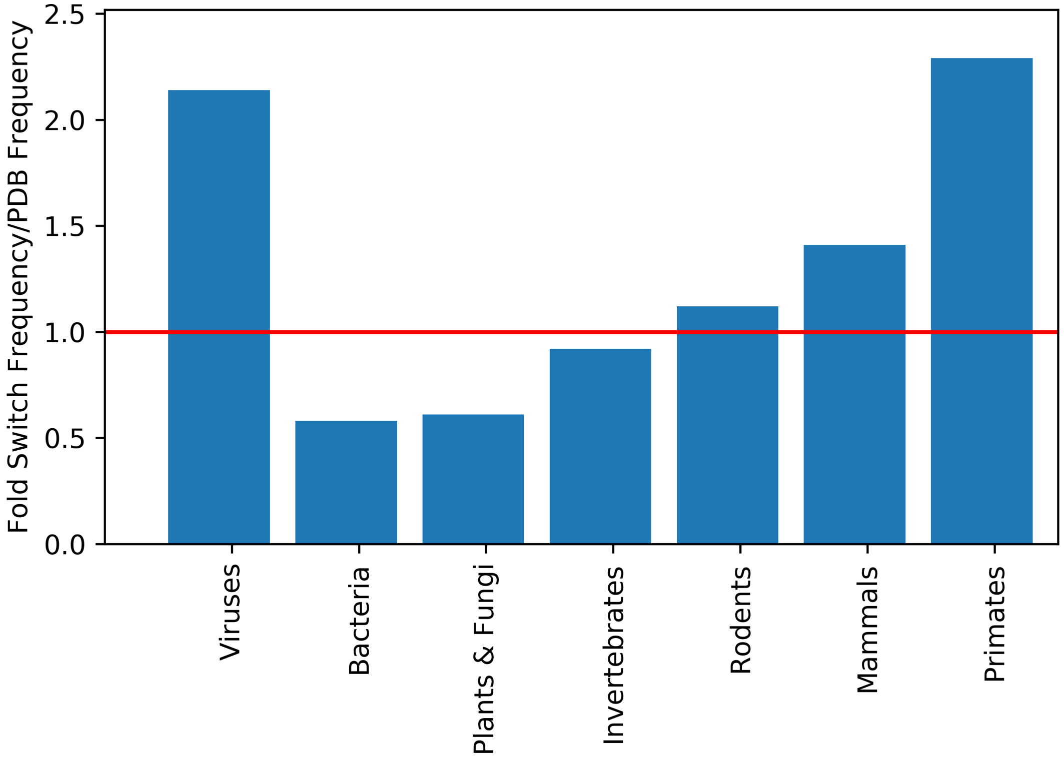


Fig. S3. Organismal distribution of unique fold-switching proteins. Types of life are on the x-axis; y-axis is the fraction of fold switchers from a given type of organism (#fold switchers from organismal type/#fold switchers) normalized by the fraction of non-redundant protein structures (**SI Methods,** *Functional, organismal, and date-solved distributions*) solved from that organism (#protein structures from organismal type/#protein structures). Red line highlights where the fraction of fold switchers from a given type of organism equals the fraction of non-redundant protein structures solved from that organism.

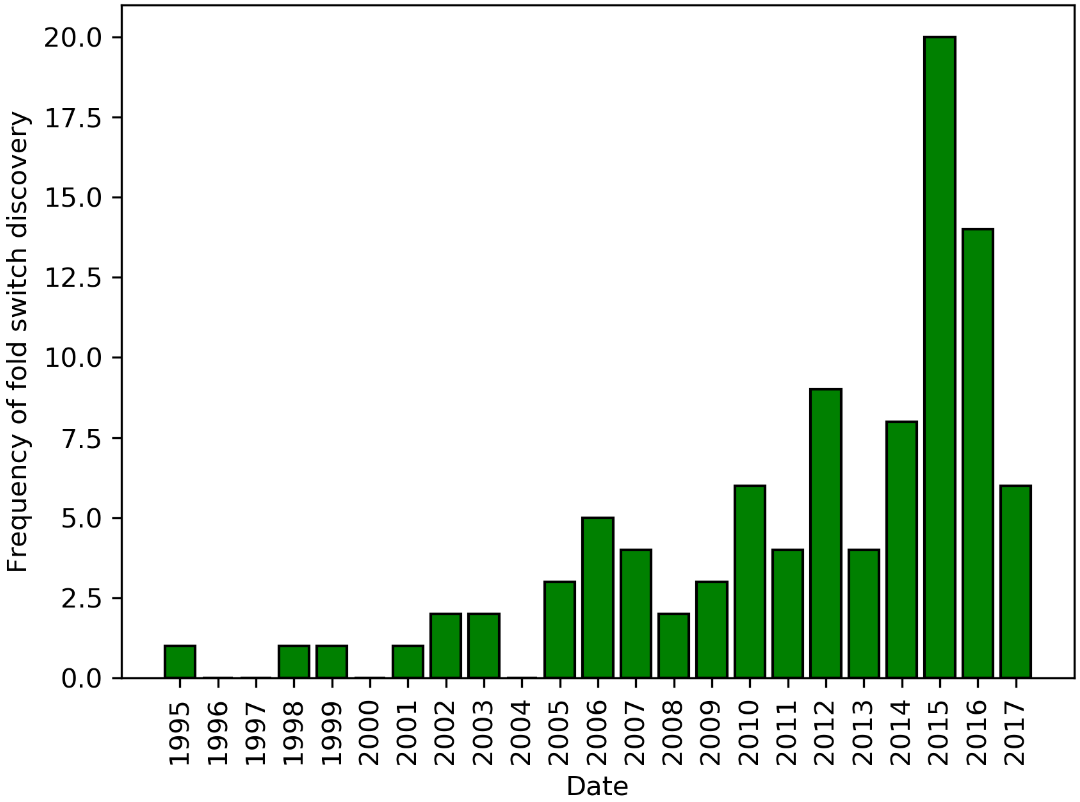
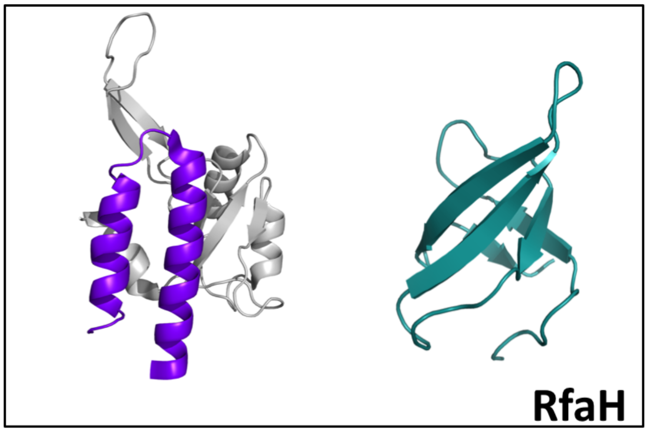
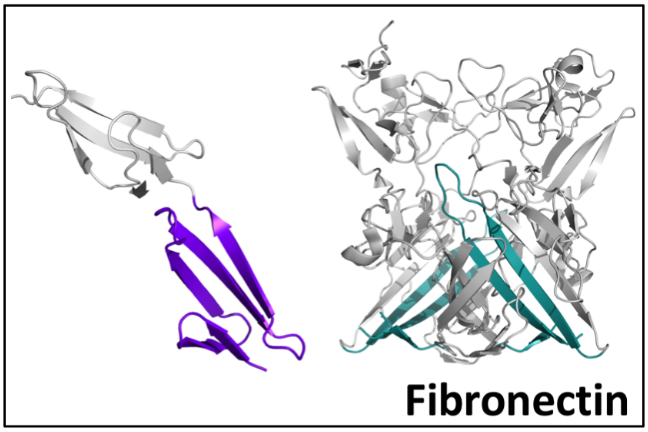
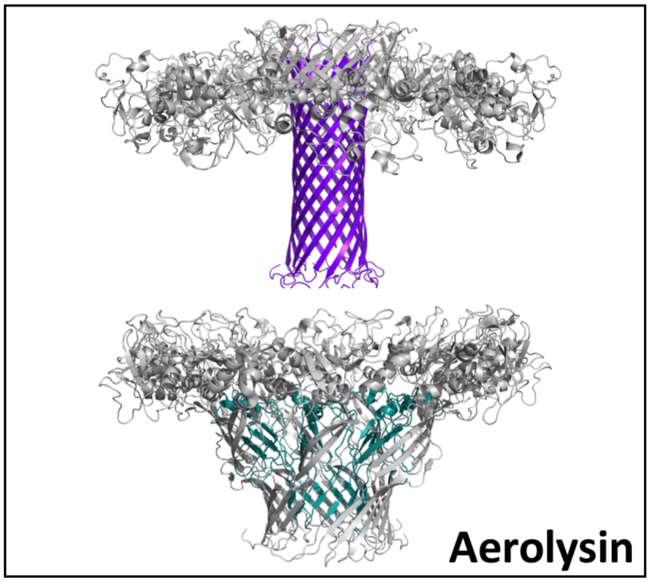


Fig. S4. The frequency of protein fold switchers being solved has increased in recent years. Determined by the release date of the latest solved structure in the fold switch pair.


RfaH


Fibronectin


Aerolysin

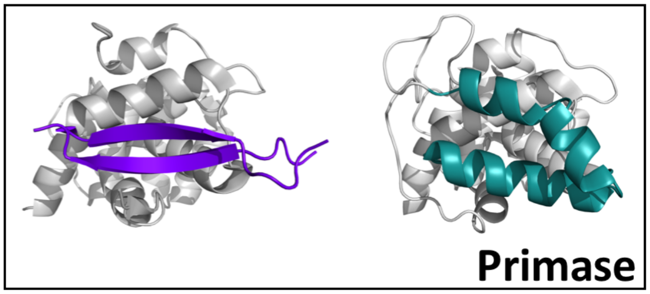

Primase

Fig S5. Fold-switching regions correspond to independent folding units in all four fold-switch classes. Order of classes is the same as **Fig. 1**. Domain boundaries in both the purple and green conformations are drawn from SEED calculations on the purple conformations, and they encompass complete secondary structure segments. Most also comprise domain-like protein subunits in both conformations. The two exceptions are the green conformations of (8)FibronectinI (8FI) and aerolysin. 8FI's boundaries encompass a three-stranded β-sheet, but its structure forms an extensive H-bond network with two structurally-unchanged units of fibronectin (7FI and 9FI), between which it is sandwiched. Accordingly, SEED recognizes 7FI-8FI-9FI as the smallest independently folding cooperative unit in the green conformation. Similarly, the green conformation of aerolysin forms extensive interactions with the β-pore region of the protein, which SEED recognizes as the smallest cooperative unit. Gray structural regions do not differ in secondary structure annotation between the two conformations. Boundaries of independent folding units and PDB ids are as follows: RfaH: 2ougD (115-156), 2lclA (115-165); Fibronectin (516-558): 3m7pA, 3ejhA; Aerolysin (213-281): 5jzt A-G, 5jzh A-G; Primase (317-353): 3l9qB; 4rr2D.

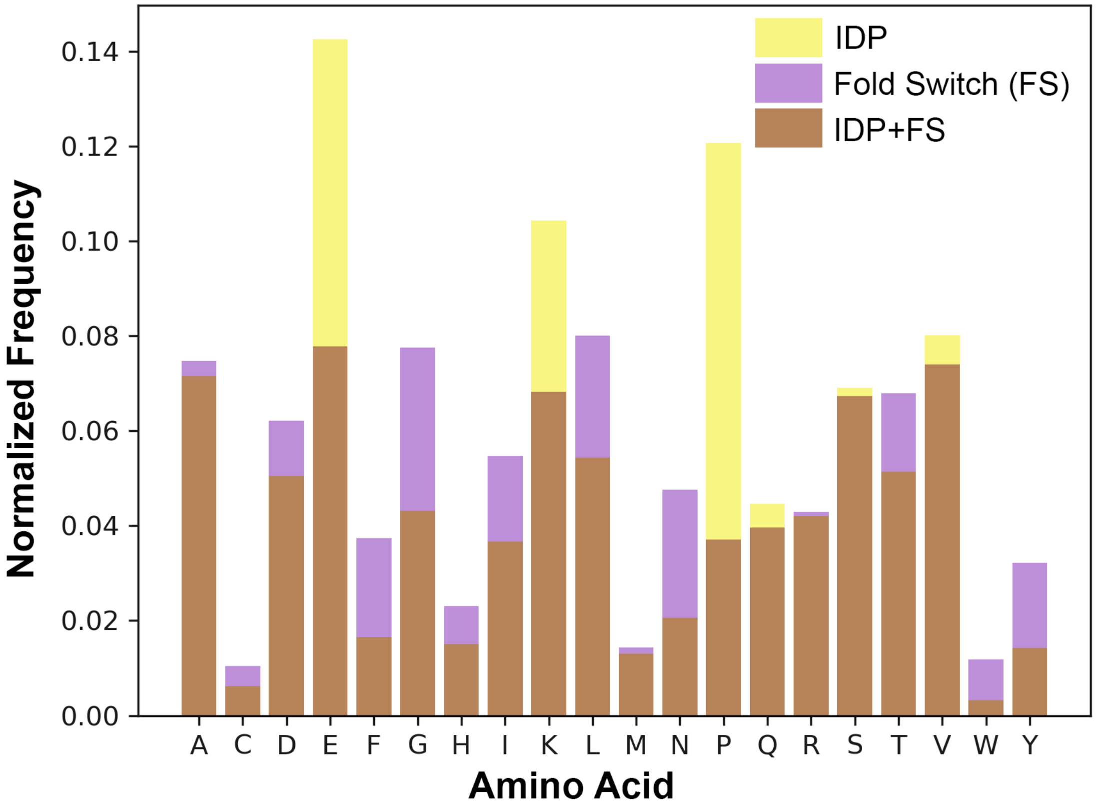


Fig. S6. The amino acid distributions of protein fold switchers and intrinsically disordered proteins (IDPs) have substantial differences. Protein fold switchers tend to have more G, N, and large hydrophobic residues while IDPs have more P, E, and K. One letter abbreviations for amino acids are as follows: A: alanine; C: cysteine; D: aspartate; E: glutamate; F: phenylalanine; G: glycine; H: histidine; I: isoleucine; K: lysine; L: leucine; M: methionine; N: asparagine; P: proline; Q: glutamine; R: arginine; S: serine; T: threonine; V: valine; W: tryptophan; Y: tyrosine. The distribution of amino acids in fold switches/IDPs was calculated from the bold regions in **Table S2**/Reference 11.

Table S1, columns 1-8, Properties of protein fold switches

PDB1+chain	PDB2+chain	Release date PDB1	Release date PDB2	Description	Source organism	Lowest sequence complexity region	Unusual properties
1g2cF	5c6bF	2011	2015	Pre-/Post-fusion RSV F	Viral	2.36	
5ec5P	3zxgB	2016	2012	Lysenin	Bacteria	2.28	Membrane
1uxmK	2namA	2009	2015	SOD1	Human	2.42	Membrane
7ahlE	4yhdG	1998	2015	Hemolysin	Bacteria	2.28	Membrane
5aoeB	5ly6B	2017	2017	Pneumolysin	Bacteria	2.62	Membrane
1ovaA	1jtiB	1992	2001	Ovalbumin (serpin)	Chicken	2.28	
3gmhL	2vfxL	2010	2002	Mad2	Human	2.69	
3m1bF	3lowA	1998	2010	Beta2-microglobulin	Human	2.86	
3j7wB	3j7vG	2014	2014	Bacteriophage T7 capsid	Viral	2.69	
2lqwA	2bzyB	2012	2006	CRKL-SH3 domain	Human	2.63	
5ejbC	1wp8C	2016	2005	Hendra virus	Viral	2.69	
2frhA	1fzpD	2006	2001	SarA	Bacteria	2.63	They grew >250 L of bacteria to get enough protein for one of these structures.
5hmgB	1htmB	1990	1995	Hemagglutinin	Virus	2.58	
4j3oF	2jmrA	2013	2007	FimF	Bacteria	2.91	Membrane
4hddA	2lepA	2013	2012	Rhomboid protease	Bacteria	2.75	
1miqB	1qs8B	2002	1999	Plasmepsin	Protozoan	2.86	
4nc9C	4n9wA	2015	2015	PimA, modulates function at membrane	Bacteria	2.36	

3j97M	1xtgB	2015	2004	SNAP-25 unfolding/refolding by toxin	Human	2.69	
2naoF	1iytA	2016	2002	Amyloid ABeta42	Human	2.36	Fibrillization
5c1vA	5c1vB	2016	2016	Calcineurin: Cis/Trans change in proline leads to a shift in b-sheet register and affects transcription factor binding	Human	2.58	
4zt0C	4cmqB	2015	2014	Cas9 apo and RNA-bound	Bacteria	2.22	
5jytA	2qkeE	2013	2017	KaiB	Bacteria	2.42	
2lejA	2lv1A	2011	2012	Major prion protein	Human	1.45	Fibrillization
3jv6A	1zk9A	2010	2005	RelB Domain swap vs heterodimer	Mouse	2.91	
5k5gA	2kb8A	2016	2009	Islet amyloid peptide/and membrane bound	Human	2.89	Membrane
4uv2D	4q79F	2014	2014	Curli transport lipoprotein CsgG	Bacteria	3.08	Membrane
1wyyB	5wrgA	2005	2017	SARS spike glycoprotein	Viral	2.22	
5fhcJ	1eboE	2016	1999	Ebola viral fusion protein	Viral	2.45	
4wsgC	1svfC	2015	1999	Simian parainfluenza virus 5	Viral	2.36	
1nqdA	1nqjB	2003	2003	Collagen-binding domain, calcium triggered change	Bacteria	2.4	
5fluE	2uy7D	2016	2007	Pillus rod	Bacteria	2.63	

3qy2A	1qb3A	2011	2000	Cks1	Yeast	2.45	
2ougC	2lclA	2012	2007	RfaH	Bacteria	2.42	
1qomB	1nocA	1999	1997	Nitric oxide synthase	Mouse	2.63	
4rr2D	3l9qB	2015	2010	Primase	Human	2.36	
2gedB	1nrjB	2006	2003	GTP-activated switch, monomer<->dimer	Yeast	2.85	
5keqF	1dzlA	2017	2000	HPV, Heparin-induced conformational changes	Virus	2.75	Cryo-EM, special cells
4y0mJ	4xwsD	2015	2015	OxyR	Bacteria	2.12	
2c1uC	2c1vB	2006	2006	Oxidized vs mixed-valence forms of MacA (from Paracoccus)	Bacteria	2.58	
4zrbC	4zrbH	2015	2015	Apo and CoA-bound	Bacteria	2.86	
4qhfA	4qhhA	2014	2014	Selecase	Archaea	2.12	
4aanA	4aalA	2012	2012	MacA (Geobacter)	Bacteria	2.34	
1x0gA	1x0gD	2006	2006	IscA with 2Fe-2S cluster	Bacteria	2.12	
4ae0A	4ow6B	2012	2014	Diphtheria toxin	Bacteria	2.86	
1mnmC	1mnmD	1998	1998	MATalpha2	Yeast	3.25	
2nntA	2mwfA	2015	2006	WW2	Human	2.58	
4jphB	5hk5H	2013	2016	Gremlin2	Mouse	2.52	
3ifaA	5et5A	2010	2016	Fructose 1,6 bisphosphatase	Human	3.08	
3k2sA	2a01C	2010	2006	Apolipoprotein bound to lipids and apo	Human	2.36	
4fu4C	4g0dZ	2013	2013	Variable collagenase 3 pro-domain peptide	Human	2.63	

1h38D	1qlnA	2002	2000	T7 RNA Polymerase	Viral	2.29	
5b3zA	5bmyA	2016	2016	Folding intermediate	Human	2.63	
1xjtA	1xjuB	2005	2005	P1 Phage endolysin	Viral	2.69	
3hdeA	3hdfA	2009	2009	R21 endolysin, membrane bound/unbound	Viral	3.08	
4b3oB	3meeA	2013	2010	p66 reverse transcriptase complexed with NNRTI and DNA/RNA	Viral	2.75	
4twaA	4ydqB	2014	2015	Prolyl-tRNA synthase Liganded/not liganded	Protozoan (eukaryotic)	2.63	
3ejhA	3m7pA	2009	2010	Fibronectin (zinc binding leads to homodimer formation and inactivation)	Human	2.52	
1k0nA	1rk4B	2001	2003	CLIC1	Human	2.28	
1xntA	3lqcA	2010	1999	XRCC1	Human	2.58	
2axzA	2grmB	2005	2005	PrgX	Bacterial	2.63	
4gqcC	4gqcB	2012	2012	Peroxiredoxin Q; Conformational change required for activity	Archaea	3.09	
4o0pA	4o01D	2014	2014	Bacteriophytochrome; light-activated changes	Bacteria	3.39	
4dxtA	4dxrA	2012	2012	SUN	Human	2.58	
4rwnA	4rwqB	2015	2015	Apo-/DNA-bound OAS1	Pig	2.42	

2hdmA	2n54B	2001	2008	Lymphotactin monomer/dimer	Human	2.46	Inclusion bodies, needed NMR to notice conformational change
3vo9B	3vpaD	2012	2012	Apo- and GDP- bound FtsZ	Bacterial	2.92	
2p3vA	2p3vD	2007	2007	*Thermotoga maritima* IMPase TM1415	Bacteria	2.86	
3zwgN	4tsyD	2015	2015	FraC	Anemone	2.19	Membrane
3ewsB	3g0hA	2009	2009	Human Deadbox	Human	3.08	
3tp2A	5lj3M	2012	2016	Splicing Factor Cwc2; in isolation/as part of spliceosome	Yeast	2.69	Cryo-EM, fermented in yeast, complicated assembly protocol
3njqA	2pbkB	2017	2007	KSHV protease	Viral	2.69	
4rmbA	4rmbB	2015	2015	Keratin 4 binding domain of Srr-1	bacteria	2.86	
2ce7C	3kdsG	2006	2009	FtsH Active site conformational switch	bacteria	2.86	
4phqA	2wcdX	2014	2009	ClyA	Bacteria	2.05	Membrane
3t1pA	1kctA	2011	1996	Serpin	Human	2.46	
3j9cA	3q8fA	2015	2012	Anthrax pore/prepore	Bacterial	2.13	Membrane
2nxqB	1jfkA	2007	2001	Calcium-binding domain of calcium binding protein 1	Entamoeba histolytica	2.63	
5l35D	5l35G	2017	2017	Sf6 capsid has conformational flexibility in subunits	Virus	2.92	Cryo-EM

5i2mA	5i2sA	2006	2007	Vesicular stomatitis virus glycoprotein G, pre-/post-fusion	Virus	2.46	
5f3kA	5f5rB	2016	2016	heat shock protein 75 kDa, mitochondrial; ATP-binding leads to helix -> domain-swapped β-strand change	Human	2.86	
4qdsA	2qqjA	2015	2007	neuropilin-2; splicing leads to change in C-terminus	Human	2.69	
5jzhA	5jztG	2016	2016	Aerolysin pore/prepore	Bacteria	2.52	Membrane
4pyiA	4pyjA	2014	2014	Domain-swapped ComT	Human	3.08	
5ineA	3mkoA	2016	2010	Pre-/post-fusion arenavirus	Viral	3.08	
1mbyA	4yypA	2002	2015	Domain-swapped dimer/bound to peptide	Mouse	2.52	
2a73B	3l5nB	2005	2005	Cleaved/uncleaved forms of human complement C3	Human	2.46	
2k0qA	2lelA	2011	2008	CopK, conformational change upon copper binding	Bacterial	2.86	
3uyiA	3v0tA	2012	2012	Perakine reductase changes upon binding NADPH	Plant	2.86	

4a5wB	3t5oA	2012	2012	Complement C6 in different complexes	Human	2.69	
1rkpA	2h44A	2004	2006	PDE5A1 Bound to different ligands	Human	2.42	
1ceeB	2k42A	1999	2008	WASP in different conformations	Human	3.08	
3o44A	1xezA	2011	2005	Cytolysin	Bacteria	2.92	Membrane
1repC	2z9oB	2000	2007	RepE54 Apo-/bound to DNA	Bacteria	2.75	
3kuyA	5c3iF	2009	2015	Histone H3.2, in different complexes	Human	2.92	
2n0aD	2kkwA	2004	2015	Alpha synuclein (micelles)	Human	3.02	Fibrillization
4m4rA	4w50B	2013	2014	Eph4A ectodomain	Human	1.89	

Table S1, columns 9-13, More properties of protein fold switches

PDB1+ chain	PDB2+ chain	Expression organism	Protein type	Trigger	Fold-switching region length (#AAs)	Class
1g2cF	5c6bF	Human	viral fusion protein	pH change	25	B
5ec5P	3zxgB	E. coli	beta-toxin	Membrane insertion	54	C
1uxmK	2namA	Yeast	Oxidoreductase	Membrane insertion	153	C
7ahlE	4yhdG	E. coli	beta-toxin	Membrane insertion	53	C
5aoeB	5ly6B	E. coli	beta-toxin	Membrane insertion	128	D
1ovaA	1jtiB	E. coli	hydrolase	Cleavage	52	D
3gmhL	2vfxL	E. coli	Cell cycle	Binding	38	D
3m1bF	3lowA	E. coli	Immune system	Change in Redox Potential	21	C
3j7wB	3j7vG	E. coli	Viral capsid	Viral capsid maturation	22	C
2lqwA	2bzyB	E. coli	Signaling	In equilibrium/possibly phosphorylation	67	C
5ejbC	1wp8C	E. coli	viral fusion protein	pH change	38	B
2frhA	1fzpD	Human	Transcription	Binding	92	A
5hmgB	1htmB	E. coli	viral fusion protein	pH change	113	B
4j3oF	2jmrA	Special expression	Cell adhesion	Membrane insertion	25	C
4hddA	2lepA	E. coli	Hydrolase	Temperature	25	B
1miqB	1qs8B	E. coli	Hydrolase	Cleavage	17	D
4nc9C	4n9wA	E. coli	Transferase	Binding	30	D
3j97M	1xtgB	E. coli	Other	Binding	40	A

2naoF	1iytA	E. coli	Protein binding	Fibrillization	28	C
5c1vA	5c1vB	E. coli	Hydrolase	Binding	31	D
4zt0C	4cmqB	E. coli	Hydrolase	Binding	168	D
5jytA	2qkeE	E. coli	Signaling	Binding	52	A
2lejA	2lv1A	E. coli	Membrane protein	Fibrillization	15	C
3jv6A	1zk9A	E. coli	Transcription	Binding	29	A
5k5gA	2kb8A	E. coli	Hormone	Membrane insertion	28	C
4uv2D	4q79F	E. coli	Transport	Membrane insertion	29	C
1wyyB	5wrgA	Spodoptera frugiperda	viral fusion protein	Receptor binding	58	B
5fhcJ	1eboE	Human	viral fusion protein	Receptor binding	56	B
4wsgC	1svfC	Trichoplusia ni	viral fusion protein	pH change	39	B
1nqdA	1nqjB	E. coli	Hydrolase	Calcium binding	33	D
5fluE	2uy7D	E. coli	Structural protein	oligomerization	30	C
3qy2A	1qb3A	E. coli	Cell cycle	oligomerization	15	C
2ougC	2lclA	E. coli	Transcription	Binding	66	A
1qomB	1nocA	E. coli	Oxidoreductase	Binding	22	C
4rr2D	3l9qB	E. coli	Transferase	Binding	46	A
2gedB	1nrjB	E. coli	Transport	Binding	29	B
5keqF	1dzlA	Simian virus 40	Viral capsid	oligomerization	34	C
4y0mJ	4xwsD	E. coli	DNA binding protein	Change in Redox Potential	69	D
2c1uC	2c1vB	E. coli	Oxidoreductase	Change in Redox Potential	39	D
4zrbC	4zrbH	E. coli	hydrolase	Binding	20	D
4qhfA	4qhhA	E. coli	Hydrolase	oligomerization	19	C
4aanA	4aalA	E. coli	Oxidoreductase	Change in Redox Potential	34	D

1x0gA	1x0gD	E. coli	Metal binding protein	Binding	98	C
4ae0A	4ow6B	Purchased	Toxin	pH change	12	D
1mnmC	1mnmD	E. coli	Transcription	Binding	13	C
2nntA	2mwfA	E. coli	Transcription	Fibrillization	37	C
4jphB	5hk5H	E. coli	Cytokine	Binding	22	A
3ifaA	5et5A	E. coli	hydrolase	Binding AMP	12	D
3k2sA	2a01C	E. coli	lipid binding protein	Binding	39	D
4fu4C	4g0dZ	E. coli	Prodomain	Binding	23	D
1h38D	1qlnA	E. coli	Transferase	Binding	16	D
5b3zA	5bmyA	E. coli	isomerase	Folding intermediate	29	D
1xjtA	1xjuB	E. coli	Hydrolase	Membrane insertion	31	D
3hdeA	3hdfA	E. coli	Hydrolase	Membrane insertion	12	D
4b3oB	3meeA	E. coli	Hydrolase	Different binding partners	40	D
4twaA	4ydqB	E. coli	Ligase	Binding	38	D
3ejhA	3m7pA	Yeast	Cell adhesion	Zn binding	23	A
1k0nA	1rk4B	E. coli	Transport	Change in Redox Potential	42	B
1xntA	3lqcA	E. coli	DNA repair	Change in Redox Potential	23	D
2axzA	2grmB	E. coli	Transcription	Binding	22	D
4gqcC	4gqcB	E. coli	Oxidoreductase	Change in Redox Potential	23	D
4o0pA	4o01D	E. coli	Transferase	Light	13	D
4dxtA	4dxrA	E. coli	Structural protein	Binding	30	C
4rwnA	4rwqB	E. coli	Transferase	Binding	14	A

2hdmA	2n54B	E. coli	Chemokine	Cleavage	22	B
3vo9B	3vpaD	E. coli	Cell cycle	Binding	14	D
2p3vA	2p3vD	E. coli	Hydrolase	Binding	16	D
3zwgN	4tsyD	E. coli	Alpha-toxin	Membrane insertion	15	B
3ewsB	3g0hA	E. coli	Hydrolase	pH change	12	D
3tp2A	5lj3M	Fermented yeast	Splicing	Binding	42	A
3njqA	2pbkB	E. coli	Hydrolase	Binding	25	D
4rmbA	4rmbB	E. coli	Cell adhesion	oligomerization	17	C
2ce7C	3kdsG	E. coli	Cell division protein	Binding	17	D
4phqA	2wcdX	E. coli	Alpha-toxin	Change in Redox Potential	17	B
3t1pA	1kctA	E. coli	Hydrolase	Cleavage	84	D
3j9cA	3q8fA	E. coli	Beta-toxin	Membrane insertion	22	C
2nxqB	1jfkA	E. coli	Metal binding protein	Binding	87	B
5l35D	5l35G	S. flexneri	viral capsid	Different geometries	28	D
5i2mA	5i2sA	None given	Viral fusion protein	pH change	40	B
5f3kA	5f5rB	E. coli	Chaperone	Binding	25	B
4qdsA	2qqjA	E. coli	Cell adhesion	Binding or splicing (?)	20	B
5jzhA	5jztG	E. coli	Beta-toxin	pH change	64	C
4pyiA	4pyjA	E. coli	Transferase	Oligomerization	14	C
5ineA	3mkoA	Drosophila	viral fusion protein	pH change	48	B
1mbyA	4yypA	E. coli	Transferase	Binding	16	C
2a73B	3l5nB	Human serum	Immune system	Membrane insertion	37	D

2k0qA	2lelA	E. coli	Metal binding protein	Binding	23	D
3uyiA	3v0tA	E. coli	Oxidoreductase	Binding	15	B
4a5wB	3t5oA	Human serum	Immune system	Binding	16	A
1rkpA	2h44A	E. coli	Hydrolase	Binding	22	D
1ceeB	2k42A	E. coli	Signaling	Binding	24	A
3o44A	1xezA	E. coli	Beta-toxin	membrane insertion	18	C
1repC	2z9oB	E. coli	Replication	Binding	17	D
3kuyA	5c3iF	E. coli	Structural protein	Binding	16	C
2n0aD	2kkwA	E. coli	Lipid binding protein	Fibrillization	138	C
4m4rA	4w50B	Trichoplusia ni	Transferase	Binding	14	D

Table S2, columns 1-3. Coincidence of fold-switching regions with cooperative folding units in proteins

PDB1[1]	PDB1 prediction[2]	QR 1[3]
1g2cF	FYD**PLVFPSDEFDASISQVNEKINQSLAFIRKSDELLHN**VNAG	1.11
5ec5P	GMSAKAAEGYEQIEVDVVAVWKEGYVYENRGSTSVDQKITITKGM**KNVNSETRTVTATHSIGSTISTGDAFEIGSVEVSYSHSHQKSQVSMTQTEVYSS**KVIEHTITIPPTSKFTRWQLNADVGGAGIEYMYLIDEVTPIGGTQSIPQVITSRAKIIVGRQIILGKTEIRIKHAERKEYMTVVSRKSWPAATLGHSKLFKFVLYEDWGGFRIKTLNTMYSGYEYAYSSDQGGIYFDQGTDNPKQRWAINKSLPLRHGDVVTFMNKYFTRSGLCYDDGPATNVYCLDKREDKWILEVVG	1.04
1uxmK	**ATK**V**VCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGF**HVH**EFGDNTAGCTSAGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVVHEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ**	1.03
7ahlE	MADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMHKKVFYSFIDDKNHNKKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQIS**DYYPRNSIDTKEYMSTLTYGFNGNVTGDDTGKIGGLIGANVSIGHTLKYVQPD**FKTILESPTDKKVGWKVIFNNMVNQNAGPYDRDSWNPVYGNQLFMKTANGSMKAAENFLDPNKASSLLSSGFSPDFATVITMDRKASKQQTNIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDRSSERYKIDWEKEEMTNLEHHHHHH	0.90
5aoeB	AHHHHHHSSGLVPRGSHMANKAVNDFILAMNYDKKKLLTHQGESIENRFIKEGNQLPDEFVVIERKKRSLSTNTSDISVTATNDSRLYPGALLVVDETLLENNPTLLAVDRAPMTYSIDLPGLASSDSFLQVEDPSNSSVRGAVNDLLAKWHQDYGQVNNVPARMQYEKITA**HSMEQLKVKFGS**A**FEK**TG**NSLDIDFN**S**VHSGEKQIQIVNFKQIYYTVSVDAVKNPGDVFQDTVTVEDLKQRGISAERPLVYISSVAYGRQVYLKLETTSKSDEV**E**AAFEA**L**I**K**GVKVAPQT**E**WKQILD**NTEVKAVILGGDPSSGARVVTGKVDMVEDLIQEGSRFTADHPGLPISYTTSFLRDNVVATFQNSTDYVETKVTAYRNGDLLLDHSGAYVAQYYITWDELSYDHQGKEVLTPKAWDRNGQDLTAHFTTSIPLKGNVRNLSVKIRECTGLAWEWWRTVYEKTDLPLVRKRTISIWGTTLYPQVEDKVEND	0.86
1ovaA	XGSIGAASMEFCFDVFKELKVHHANENIFYCPIAIMSALAMVYLGAKDSTRTQINKVVRFDKLPGFGDSIEAQCGTSVNVHSSLRDILNQITKPNDVYSFSLASRLYAEERYPILPEYLQCVKELYRGGLEPINFQTAADQARELINSWVESQTNGIIRNVLQPSSVDSQTAMVLVNAIVFKGLWEKAFKDEDTQAMPFRVTEQESKPVQMMYQIGLFRVASMASEKMKILEPFASGTMSMLVLLPDEVSGLEQLESIINFEKLTEWTSSNVMEERKIKVYLPRMKMEEKYNLTSVLMAMGITDVFSSSAN**LSGISSAESLKISQAVHAAHAEINEAGREVVGSAEAGVDAASVSEEFRADHP**FLFCIKHIATNAVLFFGRCVSP	
3gmhL	MRGSHHHHHHGSITLRGSAEIVAEFFSFGINSILYQRGIYPSETFTRVQKYGLTLLVTTDLELIKYLNNVVEQLKDWLYKCSVQKLVVVISNIESGEVLERWQFDIECDKTAKDDSAPREKSQKAIQDEIRSVIRQITATVTFLPLLEVSCSFDLLIYTDKDLVVP**EKWEESGPQFITNSEEVRLRSFTTTIHKVNSMVAYKIP**VND	0.87
3m1bF	IQRTPKIQVYSRHPAENGKSNFLNCYVSGFHPSDIEVDLL**KNGERIEKVEHSDLSFSKDWS**FYLLYYTEFTPTEKDEYACRVNHVTLSQPKIVKWDRDM	0.89
3j7wB	MASMTGGQQMGTNQGKGVVAA**GDKLALFLKVFGGEVLTAFART**SVTTSRHMVRSISSGKSAQFPVLGRTQAAYLAPGENLDDKRKDIKHTEKVITIDGLLTADVLIYDIEDAMNHYDVRSEYTSQLGESLAMAADGAVLAEIAGLCNVESKYNENIEGLGTATVIETTQNKAALTDQVALGKEIIAALTKARAALTKNYVPAADRVFYCDPDSYSAILAALMPNAANYAALIDPEKGSIRNVMGFEVVEVPHLTAGGAGTAREGTTGQKHVFPANKGEGNVKVAKDNVIGLFMHRSAVGTVKLRDLALERARRANFQADQIIAKYAMGHGGLRPEAAGAVVFKVE	2.68

2lqwA	MSSARFDSSDRSAWYMGPVSRQEAQTRLQGQRHGMFLVRDSSTCPGDYVLSVSENSRVSHYIINSLPNRRFKIGDQEFDHLPALLEFYKIHYLDTTTLIEPAPRYPSPPMGSVSAPNLPTAEDNLEYVRTLYDFPGNDAEDLPFKKGEILVIIEKPEEQWWSARNKDGRVGMIPVPYVEKLVRSSPHGKHGNRNSNSYGIPEPAHAXAQPQTTTPLPAVSGSPGAAITPLPSTQNG**PVFAKAIQKRVPCAYDKTALALEVGDIVKVTRMNINGQWEGEVNGRKGLFPFTHVKIFDPQNPDENE**	1.07
5ejbC	GILHYEKLSKIGLVKGITRKYKIKSNPLTKDIVIKMKPNVSNVSKCTGTVMENYKSRLTGILSPIKGAIELYNNNTHDLVGDVKLAGVVMAGIAIGIATAAQITAGVALYEAMK**NADNINKLKSSIESTNEAVVKLQETAEKTVYVLTALQDY**INTNLVPSIDQISCKQTELALDLALSKYLSDLLFVFGPNLQDPVSNSMTIQAISQAFGGNYETLLRTLGYATEDFDDLLESDSITGQIVYVDLSSYYIIVRVYFPILTEIQQAYVQELLPVSFNNDNSEWISIVPNFVLIRNTLISNIEVKYCLITKKSVICNQDYATPMTASVRECLTGSTDKCPRELVVSSHVPRFALSGGVLFANCISVTCQCQTTGRAISQSGEQTLLMIDNTTCTTVVLGNIIISLGKYLGSINYNSESIAVGPPVYTDKVDISSQISSMNQSLQQSKDYIKEAQKILDNVEDKIEEILSKIYHIENEIARIKKLIGEAPGGIEGRHHHHHHH	0.83
2frhA	GSHMAITKINDCFELLSMVTYADKLKSLIKK**EFSISFEEFAVLTYISENKEKEYYLKDIINHLNYKQPQVVKAVKILSQEDYFDKKRNEHDERTVLILVNAQQRKKIESLLSRVNKRITEANN**EIEL	1.04
5hmgB	GLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINGKLNRVIEK**TNEKFHQIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTGSEMNKLFEKTRRQLRENAEEMGNGCFKIYHKCDNACIESIRNGTYDHDVYRDEALNNRF**QIKG	
4j3oF	**ADSTITIRGYVRDNGCSVAAESTNF**TVDLMENAAKQFNNIGATTPVVPFRILLSPCGNAVSAVKVGFTGVADSHNANLLALENTVSAASGLGIQLLNEQQNQIPLNAPSSALSWTTLTPGKPNTLNFYARLMATQVPVTAGHINATATFTLEYQ	4.72
4hddA	LXITSFANPRVAQAFV**DYMATQGVILTIQQHNQSDVWLADE**SQAERVRAELARFLENPADPRYLAASWQAGHTENLYFQ	1.14
1miqB	TEHLTLAFKIERPYDKVLKTISKKNLKNYIKETFNFFKSGYMKQNY**LGSENDVIELDDVANIMFYGEGEVGDNHQKFM**LIFDTGSANLWVPSKKCNSSGCSIKNLYDSSKSKSYEKDGTKVDITYGSGTVKGFFSKDLVTLGHLSMPYKFIEVTDTDDLEPIYSSVEFDGILGLGWKDLSIGSIDPIVVELKNQNKIDNALFTFYLPVHDVHAGYLTIGGIEEKFYEGNITYEKLNHDLYWQIDLDVHFGKQTMEKANVIVDSGTTTITAPSEFLNKFFANLNVIKVPFLPFYVTTCDNKEMPTLEFKSANNTYTLEPEYYMNPILEVDDTLCMITMLPVDIDSNTFILGDPFMRKYFTVFDYDKESVGFAIAKN	
4nc9C	GSGAMRIGMVCPYSFDVPGGVQSHVLQLAEVLRDAGHEVSVLAPASPHVKLPDYVVSGGKAVPIPYNGSVARLRFGPATHRKVKKWIAEGDFDVLHIHEPNAPSLSMLALQAAEGPIVATF**HTSTTKSLTLSVFQGILRPYHEKIIGRIAV**SDLARRWQMEALGSDAVEIPNGVDVASFADAPLLDGYPREGRTVLFLGRYDEPRKGMAVLLAALPKLVARFPDVEILIVGRGDEDELREQAGDLAGHLRFLGQVDDATKASAMRSADVYCAPHLGGESFGIVLVEAMAAGTAVVASDLDAFRRVLADGDAGRLVPVDDADGMAAALIGILEDDQLRAGYVARASERVHRYDWSVVSAQIMRVYETVSGAGIKVQVSGAANRDETAGESV	0.85
3j97M	MRNELEEMQRRADQLADESLESTRRMLQLVEESKDAGIRTLVMLDEQGEQLDRVEEGMNHINQDMKEAEKNLKDLGKFCGLCVCPCNKLKSSDAYKKAWGNNQDGVVASQPARVVDEREQMAISGGFIRRVTNDARENEMDENLEQVSGIIGNLRH**MALDMGNEIDTQNRQIDRIMEKADSNKTRIDEANQRATKM**LG	1.02
2naoF	DAEFRHDSGYEVHH**QKLVFFAEDVGSNKGAIIGLMVGGVVIA**	1.54

5c1vA	SEPKAIDPKLSTTDRVVKAVPFPPSHRLTAKEVFDNDGKPRVDILKAHLMKEGRLEESVALRIITEGASILRQEKNLLDIDAPVTVCGDIHGQFFDLMKLFEVGGSPANTRYLFLGDYVDRGYFSIECVLYLWALKILYPKTLFLLRGNHECRHLTEYFTFKQECKIKYSERVYDACMDAFDCLPLAALMNQQFLCVHGGLSPEINTLDDIRKLDRFKEPPAYGPMCDILWSDPLEDFGNEKTQEHFTHNTVRGCSYFYSYPAVCEFLQHNNLLSILRAHEAQDAGYRMYRKSQTTGFPSLIT**IFSAPNYLDVYNNKAAVLKYENNVMNIRQFN**CSPHPSWAPNFD	0.86
4zt0C	SMDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEXAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHXIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASXIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKXDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIE**KILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDN**EENEDILEDIVLTLTLFEDREXIEERLKTYAHLFDDKVXKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFXQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVXGRHKPENIVIEXARENQTTQKGQKNSRERXKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDXYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKXKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRXNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIXNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSXPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIXERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRXLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	0.89
5jytA	MAPLRKTAVLKLYVAGNTPNSVRALKTLANILEKEFKGVYALKVIDVLKN**PQLAEEDKILATPTLAKVLPPPVRRIIGDLSNREKVLIALRLLAEEIGDYKDD**DDK	0.97
2lejA	GAMDPGQGGGTHSQWNKPSKPKTNMKHM**AGAAAAGAVVGGLGG**YMLGSAMSRPIIHFGSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVIEQMCITQYERESQAYYQRGSS	0.88
3jv6A	TSELRICRINKESGPCTGGEELYLLCDKVQKEDISVVF**STASWEGRADFSQADVHRQIAIVFKTPPY**EDLEISEPVTVNVFLQRLTDGVCSEPLPFTYLPR	0.90
5k5gA	KCNTATCAT**QRLANFLVHSSNNFGAILSSTNVGSNTY**	2.81
4uv2D	SLTAPPKEAARPTLMPRAQSYKDLTHLPAPTGKIFVSVYNIQDETGQFKPYPASNFSTAVPQSATAMLVTALKDSRWFIPLERQGLQNLLNERKIIRAAQENGTVAINNRIPLQSLTAPPKEAARPTLMPRAQSYKDLTHLPAPTGKIFVSVYNIQDETGQFKPYPASNFSTAVPQSATAXLVTALKDSRWFIPLERQGLQNLLNERKIIRAAQENGTVAINNRIPLQSLTAANIXVEGSIIGYESNVKSGGVGARYFGIGADTQYQLDQIAVNLRVVNVSTGEILSSVNTSKTILSYEVQAGVFRFIDYVGYTSNEPVXLCLXSAIETGVIFLINDGID**RGLWDLQNKAERQNDILVKYRHSVPPES**	0.91

1wyyB	GSAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAE**VQIDRLITGGRGGSDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKY**E	0.90
5fhcJ	EAIVNAQPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIY**IEGLMHNQDGLICGLRQLANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGG**	0.87
4wsgC	LDLAALMQIGVIPTNVRQLMYYTEASSAFIVVKLMPTIDSPISGCNITSISSYNATVTKLLQPIGENLETIRNQLIPTRRRFAGVVIGLAALGVATAAQVTAAVALVKANENAAAILNLKNAIQ**KTNAAVADVVQATQSLGTAVQAVQDHINSVVSPAITAAN**CKAQDAIIGSILNLYLTELTTIFHNQITNPALSPITIQALRILLGSTLPTVVEKSFNTQISAAELLSSGLLTGQIVGLDLTYMQMVIKIELPTLTVQPATQIIDLATISAFINNQEVMAQLPTRVMVTGSLIQAYPASQCTITPNTVYCRYNDAQVLSDDTMACLQGNLTRCTFSPVVGSFLTRFVLFDGIVYANCRSMLCKCMQPAAVILQPSSSPVTVIDMYKCVSLQLDNLRFTITQLANVTYNSTIKLESSQILPIDPLDISQNLAAVNKSLSDALQHLAQSDTYLSAIEDKIEEILSKIYHIENEIARIKKLIGEAPGGIEGRHHHHHH	0.78
1nqdA	**GSPGIPGNEKLKEKENNDSSDKATVIPNFNTTM**QGSLLGDDSRDYYSFEVKEEGEVNIELDKKDEFGVTWTLHPESNINDRITYGQVDGNKVSNKVKLRPGKYYLLVYKYSGSGNYELRVNK	
5fluE	**APTIPQGQGKVTFNGTVVDAPCSISQKSAD**QSIDFGQLSKSFLEAGGVSKPMDLDIELVNCDITAFKGGNGAKKGTVKLAFTGPIVNGHSDELDTNGGTGTAIVVQGAGKNVVFDGSEGDANTLKDGENVLHYTAVVKKSSAVGAAVTEGAFSAVANFNLTYQ	4.26
3qy2A	MYHHYHAFQGRKLTDQERARVLEFQDSIHYSPRYSDDNYEYRHVMLPKAMLKVIPSDYFNSEVGTLRILTEDEWRGLGITQSLGWEHYECHAAEPHILLFKR**PLNYEAELRAATAAA**	
2ougC	MQSWYLLYCKRGQLQRAQEHLERQAVNCLAPMITLEKIVRGKRTAVSEPLFPNYLFVEFDPEVIHTTTINATRGVSHFVRFGASPAIVPSAVIHQLSVYK**PKDIVDPATPYPGDKVIITEGAFEGFQAIFTEPDGEARSMLLLNLINKEIKHSVKNTEFRK**L	0.96
1qomB	SLDKLHVTSTRPQYVRIKNWGSGEILHDTLHHKATSDFTCKSKSCLGSIMNPKSLTRGPRDKPTPLEELLPHAIEFINQYYGSFKEAKIEEHLARLEAVTKEIETTGTYQLTLDELIFATKMAWRNAPRCIGRIQWSNLQVFDARNCSTAQEMFQHICRHILYATNNGNIRSAITVFPQRSDGKHDFRLWNSQLIRYAGYQMPDGTIRGDAATLEFTQLCIDLGWKPRYGRFDVLPLVLQADGQDPEVFEIPPDLVLEVTMEHPKYEWFQELGLKWYALPAVANMLLEVGGLEFPACPFNGWYMGTEIGVRDFCDTQRYNILEEVGRRMGLETHTLASLWKDRAVTEINVAVLHSFQKQNVTIMDHHTASESFMKHMQNEYRARGGCPADWIW**LVPPVSGSITPVFHQEMLNYVL**SPFYYYQIEPWKTHIWQNEHHHHHH	0.96
4rr2D 4rr2D cont’d	MEFSGRKWRKLRLAGDQRNASYPHCLQFYLQPPSENISLIEFENLAIDRVKLLKSVENLGVSYVKGTEQYQSKLESELRKLKFSYRENLEDEYEPRRRDHISHFILRLAYCQSEELRRWFIQQEMDLLRFRFSILPKDKIQDFLKDSQLQFEAISDEEKTLREQEIVASSPSLSGLKLGFESIYKIPFADALDLFRGRKVYLEDGFAYVPLKDIVAIILNEFRAKLSKALALTARSLPAVQSDERLQPLLNHLSHSYTGQDYSTQGNVGKISLDQIDLLSTKSFPPCMRQLHKALRENHHLRHGGRMQYGLFLKG**IGLTLEQALQFWKQEFIKGKMDPDKFDKGYSYNIRHSFGKEGKRTD**YTPFSCLKIILSNPPSQGDYHGCPFRHSDPELLKQKLQSYKISPGGISQILDLVKGTHYQVACQKYFEMIHNVDDCGFSLNHPNQFFCESQRILNGGKDIKKEPIQPETPQPKPSVQKTKDASSALASLNSSLEMDMEGLEDYFSEDS	0.84

2gedB	HMDGFKFANLEASVVAFEGSINKRKISQWREWIDEKLYQPSIIIAGPQNSGKTSLLTLLTKTSLLTLLTTDSVRPT**ADYDGSGVTLVDFPGHVKLR**YKLSDYLKTRAKFVKGLIFMVDSTVDPKKLTTTAEFLVDILSITESSCENGIDILIACNKSELFTARPPSKIKDALESEIQKVIERRK	0.90
5keqF	LWLPSEATVYLPPVPVSKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQYRVFRIHLPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASAYAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTNVAVNPGDCPPLELINTVIQDGDMVDTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSEPYGDSLFFYLRREQMFVRHLFNRAGAVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFNKPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRHGEEYDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPP**PGGTLEDTYRFVTSQAIACQKHTPPAPKEDPLKK**YTFWEVNLKEKFSADLDQFPLGRKFLLQAGLKAKPKFTLGKR	
2c1uC	ETEAIDNGALREEAKGVFEAIPEKMTAIKQTEDNPEGVPLTAEKIELGKVLFFDPRMSSSGLISCQTCHNVGLGGVDGLPTSIGHGWQKGPRNAPTMLNAIFNAAQFWDGRAADLAEQAKGPVQAGVEMSNTPDQVVKTINSMPEYVEAFKAAFPEEADPVTFDNFAAAIEQFEATLITPNSAFDRFLAGDDAAMTDQEKRGLQAFMETGCTACHYGVNFG**GQDYHPFGLIAKPGAEVLPAGDTGRFEVTRTTDDEYVFR**AAPLRNVALTAPYFHSGVVWELAEAVKIMSSAQIGTELTDQQAEDITAFLGTLTGEQPVIDHPILPVRTGTTPLPTPM	0.83
4zrbC	SNAM**KDFHFDAISAFENYEIEKMR**DGHVVVTTKVVNSSLNYYGNAHGGYLFTLCDQISGLVVISLGLDGVTLQSSINYLKAGKLDDVLTIKGECVHQGRTTCVMDVDITNQEGRNVCKATFTMFVTGQRSEERRVRI	0.86
4qhfA	MKDRKILNEILSNTINELNLNDKKANIKIKIKPLKRKIASISLTNKTIYINKNILPYLSDEEIRFILAHELLHLK**YGKYHINEFEEELLFLFPN**KEAILINLINKLHQKK	
4aanA	WSHPQFEKGAETAVPNSEDVMKRAQGLFKPIPAKPPVMKDNPASPSRVELGRMLFFDPRLSASHLISCNTCHNVGLGGTDILETSIGHGWQKGPRNSPTVLNAVYNIAQFWDGRAEDLAAQAKGPVQASVEMNNKPENLVATLKSIPGYPPLFRKAFPGQGDPVTFDNVAKAIEVFEATLVTPDAPFDKYLKGNRKAISSTAEQGLALFLDKGCAACHSGVNMGG**TGYFPFGVREDPGPVVRPVDDTGRYKVTSTAADKY**VFRSPSLRNVAITMPYFHSGKVWKLKDAVKIMGSAQLGISITDADADKIVTFLNTLTGAQPKVMHPVLPPNSDDTPRPVSN	0.84
1x0gA	MVELTPAAIQELERLQTHGVRRGQAAILRIQVQPSECGDWRY**DLALVAEPKPTDLLTQSQGWTIAIAAEAAELLRGLRVDYIEDLMGGAFRFHNPNASQTCGCGMAFRVSRS**	
4ae0A	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQTALEHP**ELSELKTVTGTN**PVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS	0.87
1mnmC	QLTQKNKSADGLV**FNVVTQDMINKST**KPYRGHRFTKENVRILESWFAKNIENPYLDTKGLENLMKNTSLSRIQIKNWVSNRRRKEKT	3.80

2nntA	GSM**GATAVSEWTEYKTADGKTYYYNNRTLESDWEKPQELK**	0.93
4jphB	MRKNRPAGAIPSPYKDGSSNNSERW**HHQIKEVLASSQEALVVTERKY**LKSDWCKTQPLRQTVSEEGCRSRTILNRFCYGQCNSFYIPRHVKKEEDSFQSCAFCKPQRVTSVIVELECPGLDPPFRIKKIQKVKHCRCMSVNLSDSDKQ	1.29
3ifaA	TDRSPFET**DMLTLTRYVMEK**GRQAKGTGELTQLLNSMLTAIKAISSAVRKAGLAHLYGIAGSVNVTGDQVKKLDVLSNSLVINMLQSSYSTCVLVSEENKDAIITAKEKRGKYVVCFDPLDGSSNIDCLASIGTIFAIYRKTSEDEPSEKDALQCGRNIVAAGYALYGSATLVALSTGQGVDLFMLDPALGEFVLVEKDVKIKKKGKIYSLNEGYAKYFDAATTEYVQKKKFPEDGSAPYGARYVGSMVADVHRTLVYGGIFLYPANQKSPKGKLRLLYECNPVAYIIEQAGGLATTGTQPVLDVKPEAIHQRVPLILGSPEDVQEYLTCVQKNQAGS	0.95
3k2sA	DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQ**EKLSPLGEEMRDRARAHVDALRTHLAPYSDELRQRLAAR**LEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQ	0.91
4fu4C	GGD**EDDLSEEDLQFAERYLRSYYHPT**	9.26
1h38D	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLKAGEVADNAAAKPLITTLLPKMIARINDWFEEVKAKRGKRPTAFQFLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAIGRAIEDEARFGRIRDLEAKHFKKNVEEQLNKRVGHVYKKAFMQVVEADMLSKGLLGGEAWSSWHKEDSIHVGVRCIEMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRAGALAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIWFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYYWLKIHGANCAGVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNCSLPLAFDGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYGIVAKKVNEILQADAIN**GTDNEVVTVTDENTGE**ISEKVKLGTKALAGQWLAYGVTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGYMAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCAVHWVTPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFA	0.87
5b3zA	MEKLPPG**WEKRMSRSSGRVYYFNHITNASQWERPSG**KIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTN	
1xjtA	MKGKTAAGGGAICAIAVXITIVXGNGNVRTNQAGLELI**GNAEGCRRDPYCPAGVWTDGIGNTHGVTPG**VRKTDQQIAADWEKNILIAERCINQHFRGKDXPDNAFSAXTSAAFNXGCNSLRTYYSKARGXRVETSIHKWAQKGEWVNXCNHLPDFVNSNGVPLRGLKIRREKERQLCLTGLVNEHHHHHH	0.89

3hdeA	MPPSLRKAVAAAIGGGAIAIASVLITGPSGNDGLEGVSYIPYKDIVGVWTVCHGHTGKDIMLGKTYTKAECKALLNKDLATVARQINPYIKVDIPETMRGALYSFVYNVGAGNFRTSTLLRKINQGDIKGACD**QLRRWTYAGGKQ**WKGLMTRREIEREICLWGQQ	0.84
4b3oB	GPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDGTKWRKLVDFRELNKKTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWET**WWTEYWQATWVPEWEFVNTPPLVKLWYQLEKEPIVGAETF**	
4twaA	GAMAITSKKIENFSDWYTQVIVKSELIEYYDISGCYILRPAAYYIWECVQAFFNKEIKKLNVENSYFPLFVTKNKLEKEK**NHIEGFSPEVAWVTKYGDSNLPEEIAIRPTSETIMYSV**FPKWIRSYRDLPLKLNQWNTVVRWEFKQPTPFIRTREFLWQEGHTAHKNEEEAVKLVFDILDLYRRWYEEYLAVPIIKGIKSEGEKFGGANFTSTAEAFISENGRAIQAATSHYLGTNFAKMFKIEFEDENEVKQYVHQTSWGCTTRSIGIMIMTHGDDKGLVLPPNVSKYKVVIVPIFYKTTDENAIHSYCKDIEKILKNAQINCVYDDRASYSPGYKFNHWELRGIPIRIEVGPKDLQNNSCVIVRRDNNEKCNVKKESVLLETQQMLVDIHKNLFLKAKKKLDDSIVQVTSFSEVMNALNKKKMVLAPWCEDIATEEEIKKETQRLSLNQTNSETTLSGAMKPLCIPLDQPPMPPNMKCFWSGKPAKRWCLFGRSY	0.97
3ejhA	**DQCIVDDITYNVNDTFHKRHEEG**HMLNCTCFGQGRGRWKCDPVDQCQDSETGTFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQPLQTYPSS	1.07
1k0nA	MAEEQPQVELFVKAGSDGAKIGNCPFSQRLFMVLWL**KGVTFNVTTVDTKRRTETVQKLCPGGELPFLLYGTEVHTDTN**KIEEFLEAVLCPPRYPKLAALNPESNTAGLDIFAKFSAYIKNSNPALNDNLEKGLLKALKVLDNYLTSPLPEGVDETSAEDGVSQRKFLDGNELTLADCNLLPKLHIVQVVCKKYRGFTIPEAFRGVHRYLSNAYAREEFASTCPDDEEIELAYEQVAKALK	
1xntA	MPEIRLRHVVSCSSQD**STHCAENLLKADTYRKWRAAKAG**EKTISVVLQLEKEEQIHSVDIGNDGSAFVEVLVGSSAGGAGEQDYEVLLVTSSFMSPSESRSGSNPNRVRMFGPDKLVRAAAEKRWDRVKIVCSQPYSKDSPFGLSFVRFHSPPDKDEAEAPSQKVTVTKLGQFRVKEEEESAN	0.79
2axzA	XFKIGSVLKQIRQELNYHQIDLYSGIXSKSVYIKVEADSRPISVEELSKFSERLGVNFFEILNRAGXNTKSVNETGKEKLLISKIFTNPDLFDKNFQRIEPKRLTSLQYFSIYLGYISIAHHYNIEVPTFNKTITSDLKHLYDKRTTFFGIDYEIVSNLLNVLPYEEVSSIIKPXYPIVDSFGKDYDLTIQTVLKNALTISIXNRNLKEAQYYINQFEHLKTIKNISINGYYDLEINYLKQIYQFLTDKNIDSYLNAVNIINIFKIIGKEDIHRSLVEELTKISAKEKFTPP**KEVTYYENYVAIENNPIPEIK**EQS	0.89
4gqcC	MKGLVELGEKAPDFTLPNQDFEPVNLYEVLKRGRPAVLIFFPAAF**SPVCTKELCTFRDKMAQLEKANA**EVLAISVDSPWCLKKFKDENRLAFNLLSDYNREVIKLYNVYHEDLKGLKMVAKRAVFIVKPDGTVAYKWVTDNPLNEPDYDEVVREANKIAGELVA	0.80
4o0pA	MASMTGGQQMGRGSMSRDPLPFFPPLYLGGPEITTENCEREPIHIPGSIQPHGALLTADGHSGEVLQMSLNAATFLGQEPTVLRGQTLAALLPEQWPALQAALPPGCPDALQYRATLDWPAAGHLSLTVHRVGELLILEFEPTEAWDSTGPHALRNAMFALESAPNLRALAEVATQTVRELTGFDRVMLYKFAPDATGEVIAEARREGLHAFLGHRFPASDIPAQARALYTRHL	

	LRLTADTRAAAVPLDPVLNPQTNAPTPLGGAVLRATSPMHMQYLRNMGVGSSLSVSVVVGGQLWGLIACHHQTPYVLPPDLRTTLEYLGRLLSLQVQVKEAADVAAFRQSLREHHARVALAAAHSLSPHDTLSDPALDLLGLMRAGGLILRFEGRWQTLGEVPPAPAVDALLAWLETQPGALVQTDALGQLWPAGADLAPSAAGLLAISVGEGWSECLVWLRPELRLEVAWGGATPDQAKDDLGP**RHSFDTYLEEKRG**YAEPWHPGEIEEAQDLRDTLTGALEHHHHHH	
4dxtA	GPGVTEEQVHHIVKQALQRYSEDRIGLADYALESGGAS**VISTRCSETYETKTALLSLFGIPLWYHSQS**PRVILQPDVHPGNCWAFQGPQGFAVVRLSARIRPTAVTLEHVPKALSPNSTISSAPKDFAIFGFDEDLQQEGTLLGKFTYDQDGEPIQTFHFQAPTMATYQVVELRILTNWGHPEYTCIYRFRVHGEPAH	0.79
4rwnA	MELRHTPARDLDKFIEDHLLPNTCFRTQVKEAIDIVCRFLKERCFQGTADPVRVSKVVKG**GSSGKGTTLRGRSD**ADLVVFLTKLTSFEDQLRRRGEFIQEIRRQLEACQREQKFKVTFEVQSPRRENPRALSFVLSSPQLQQEVEFDVLPAFDALGQWTPGYKPNPEIYVQLIKECKSRGKEGEFSTCFTELQRDFLRNRPTKLKSLIRLVKHWYQTCKKTHGNKLPPQYALELLTVYAWEQGSRKTDFSTAQGFQTVLELVLKHQKLCIFWEAYYDFTNPVVGRCMLQQLKKPRPVILDPADPTGNVGGGDTHSWQRLAQEARVWLGYPCCKNLDGSLVGAWTMLQKIGSHHHHHH	0.80
2hdmA	GSEVSDKRTCVSLTTQRLPCSRIKTYTITEGSLRAVIFITKRGLKVCA**DPQATWVRDVVRSMDRKSNTRN**NMIQTKPTGTQQSTNTAVTLTG	0.81
3vo9B	GHMATLKVIGVGGGGNNAVNRXIDHGMNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQGADXVFVTSGXGGGTGTGAAPVVAKIAKEXGALTVGVVTRPFSFEGRKRQTQAAAGVEAXKAAVDTLIVIPNDRLLDIVDKSTPXXEAFKEADNVLRQ**GVQGISDLIAVSGE**VNLDFADVKTIXSNQGSALXGIGVSSGENRAVEAAKKAISSPLLETSIVGAQGVLXNITGGESLSLFEAQEAADIVQDAADEDVNXIFGTVINPELQDEIVVTVIATGFD	0.81
2p3vA	MDRLDFSIKLLRKVGHLLMIHWG**RVDNVEKKTGFKDIVT**EIDREAQRMIVDEIRKFFPDENIMAEEGIFEKGDRLWIIDPIDGTINFVHGLPNFSISLAYVENGEVKLGVVHAPALNETLYAEEGSGAFFNGERIRVSENASLEECVGSTGSYVDFTGKFIERMEKRTRRIRILGSAALNAAYVGAGRVDFFVTWRINPWDIAAGLIIVKEAGGMVTDFSGKEANAFSKNFIFSNGLIHDEVVKVVNEVVEEIGGK	0.93
3zwgN	**SADVAGAVIDGAGLG**FDVLKTVLEALGNVKRKIAVGIDNESGKTWTAMNTYFRSGTSDIVLPHKVAHGKALLYNGQKNRGPVATGVVGVIAYSMSDGNTLAVLFSVPYDYNWYSNWWNVRVYKGQKRADQRMYEELYYHRSPFRGDNGWHSRGLGYGLKSRGFMNSSGHAILEIHVTKA	
3ewsB	MHHHHHHSSGVDLGTENLYFQSMEDRAAQSLLNKLIRSNLVDN**TNQVEVLQRDPN**SPLYSVKSFEELRLKPQLLQGVYAMGFNRPSKIQENALPLMLAEPPQNLIAQSQSGTGKTAAFVLAMLSQVEPANKYPQCLCLSPTYELALQTGKVIEQMGKFYPELKLAYAVRGNKLERGQKISEQIVIGTPGTVLDWCSKLKFIDPKKIKVFVLDEADVMIATQGHQDQSIRIQRMLPRNCQMLLFSATFEDSVWKFAQKVVPDPNVIKLKREEETLDTIKQYYVLCSSRDEKFQALCNLYGAITIAQAMIFCHTRKTASWLAAELSKEGHQVALLSGEMMVEQRAAVIERFREGKEKVLVTTNVCARGIDVEQVSVVINFDLPVDKDGNPDNETYLHRIGRTGRFGKRGLAVNMVDSKHSMNILNRIQEHFNKKIERLDTDDLDEIE	9.25
3tp2A	GAMT**SWRDKSAKVQVKESELPSSIPAQTGLTFNIWYNKWSQGFAGNT**RFVSPFALQPQLHSGKTRGDNDGQLFFCLFFAKGMCCLGPKCEYLHHIPDEEDIGKLALRTEVLDCFGREKFADYREDMGGIGSFRKKNKTLYVGGIDGALNSKHLKPAQIESRIRFVFSRLGDIDRIRYVESKNCGFVKFKYQANAEFAKEAMSNQTLLLPSDKEWDDRREGTGLLVKWAN	0.91

3njqA	GLYVGG**FVDVVSCPKLEQELYLDPDQVTDYL**PVTEPLPITIEHLPETEVGWTLGLFQVSHGIFCTGAITSPAFLELASRLADTSHVARAPVKNLPKEPLLEILHTWLPGLSLSSIHPRELSQTPSGPVFQHVSLCALGRRRGTVAVYGHDAEWVVSRFSSVSKSERAHILQHVSSCRLEDLSTPNFVSPLETL	
4rmbA	RIMKLDDERQTETYITQINPEGKEMYFASGLGNLYTIIGSDGTSGSPVNLLNAEVKILKTNSKNLTDSMDQNYDSPEFEDVTSQYSYTNDGSKITIDWKTNSISSTTSYVVLVKIPKQSGVLYSTVSDINQTYGSKYSYGHTN**ISGDSDANAEIKLLSLE**HHHHHH	9.60
2ce7C	MATMYKPSGNKRVTFKDVGGAEEAIEELKEVVEFLKDPSKFNRIGARMPKGILLVGPPGTGKTLLARAVAGEANVPFFHISGSDFVELFVGVGAARVRDLFAQAKAHAPCIVFIDEIDAVGRHRGAGLGGGHDEREQTLNQLLVEMDGFDSKEGIIVMAATNRPDILDPALLRPGRFDKKIVVDPPDMLGRKKILEIHTRNKPLAEDVNLEIIAKRTPGFVGADLENLVNEAALLAAREGRDKITMKDFEEAIDRVIAGPARKSLLISPAEKRIIAYHEAGHAVVSTVVPNGEPVHRISII**PRGYKALGYTLHLPEED**KYLVSRNELLDKLTALLGGRAAEEVVFGDVTSGAANDIERATEIARNMVCQLGMSEELGPLAWGKEEQEVFLGKEITRLRNYSEEVASKIDEEVKKIVTNCYERAKEIIRKYRKQLDNIVEILLEKETIEGDELRRILSEEFEKVVEAAALEHHHHHH	0.88
4phqA	CDKTVEVVKNAIETADGALDLYNKYLDQVIPWQTFDETIKELSRFKQEYSQAASVLVGDIKTLLMDSQDKYFEATQTVYEWAGVATQLLAAYILLFDEYNEKKASAQKDILIKVLDDGITKLNEAQKSLLVSSQSFNNASGKLLALDSQLTNDFSEKSSYFQSQVDKIRKEAY**AGAAAGVVAGPFGLIIS**YSIAAGVVEGKLIPELKNKLKSVQNFFTTLSNTVKQANKDIDAAKLKLTTEIAAIGEIKTETETTRFYCDYDDLMLSLLKEAAKKMINTANEYQKRHGKKTLFEVPEV	0.92
3t1pA	NKITPNLAEFAFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGNGLFLSEGLKLVDKFLEDVKKLYHSEAFTVNFGDTEEAKKQINDYVEKGTQGKIVDLVKELDRDTVFALVNYIFFKGKWERPFEVKDTEEEDFHVDQVTTVKVPMMKRLGMFNIQHSKKLSSWVLLMKYLGNATAIFFLPDEGKLQHLENELTHDIITKFLENED**RRSASLHLPKLCITGTYDLKSVLGQLGITKVFSNGADLSGVTEEAPLKLSKAVHKAVLCIDEKGTEAAGAMFLEAIPRSIPPEV**KFNKPFVFLMIEQNTKSPLFMGKVVNPTQK	
3j9cA	TVPDRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSTASDPYSDFEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSQTRTISKNTSTSRTHTSEVHGNAEVHASFFDIGGSVSAGFSNSNSSTVAI**DHSLSLAGERTWAETMGLNTAD**TARLNANIRYVNTGTAPIYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPSKNLAPIALNAQDDFSSTPITMNYNQFLELEKTKQLRLDTDQVYGNIATYNFENGRVRVDTGSNWSEVLPQIQETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNLQYQGKDITEFDFNFDQQTSQNIKNQLAELNATNIYTVLDKIKLNAKMNILIRDKRFHYDRNNIAVGADESVVKEAHREVINSSTEGLLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNISSLRQDGKTFIDFKKYNDKLPLYISNPNYKVNVYAVTKENTIINPSENGDTSTNGIKKILIFSKKGYEIG	1.00
2nxqB	MAEALFKEIDVNGDGAVSYEEVKAFVSKKRAIKNEQLLQLIFKSID**ADGNGEIDQNEFAKFYGSIQGQDLSDDKIGLKVLYKLMDVDGDGKLTKEEVTSFFKKHGIEKVAEQVMKADANGDGYITLEEFLEFS**L	1.84
5l35D	PNNLDSNVSQIVLKKFLPGFMSDLVLAKTVDRQLLAGEINSSTGDSVSFKRPHQFSSLRTPTGDISGQNKNNLISGKATGRVGNYITVAVEYQQLEEAIKLNQLEEILAPVRQRIVTDLETELAHFMMNNGALSLGSPNTPITKWSDVAQTASFLKDLGVNEGENYAVMDPWSAQRLA**DAQTGLHASDQLVRTAWENAQIPTNFGG**IRALMSNGLASRTQGAFGGTLTVKTQPTVTYNAVKDSYQFTVTLTGATASVTGFLKAGDQVKFTNTYWLQQQTKQALYNGATPISFTATVTADANSDSGGDVTVTL	0.90

	SGVPIYDTTNPQYNSVSRQVEAGDAVSVVGTASQTMKPNLFYNKFFCGLGSIPLPKLHSIDSAVATYEGFSIRVHKYADGDANVQKMRFDLLPAYVCFNPHMGGQFFGNP	
5i2mA	KFTIVFPHNQKGNWKNVPSNYHYCPSSSDLNWHNDLIGTALQVKMPKSHKAIQADGWMCHASKWVTTCDFRWYGPKYITHSIRSFTPSVEQCKESIEQTKQGTWLNPGFPPQSCGYATVTDAEAVIVQVTPHHVLVDEYTGEWVDSQFINGKCSNYICPTVHNSTTWHSDYKVKGLCDSNLISMDITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVRLPSGVWFEMADKDLFAAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAGLPISPVDLSYLAPKNPGTGPAFTIINGTLKYFETRYIRVDIAAPILSRMVGMISGTTTERELWDDWAPYEDVEIGPNGVLRTSSGYKF**PLYMIGHGMLDSDLHLSSKAQVFEHPHIQDAASQLPDDES**	
5f3kA	AGHSTQTAEDKEEPLHSIISSTESVQ**GSTSKHEFQAETKKLLDIVARSLYS**EKEVFIRELISNASDALEKLRHKLVSDGQALPEMEIHLQTNAEKGTITIQDTGIGMTQEELVSNLGTIARSGSKAFLDALQNQAEASSKIIGQFGVGFYSAFMVADRVEVYSRSAAPGSLGYQWLSDGSGVFEIAEASGVRTGTKIIIHLKSDCKEFSSEARVRDVVTKYSNFVSFPLYLNGRRMNT	1.06
4qdsA	GSHMFQCNVPLGMESGRIANEQISASSTYSDGRWTPQQSRLHGDDNGWTPNLDSNKEYLQVDLRFLTMLTAIATQGAISRETQNGYYVKSYKLEVSTNGEDWMVYRHGKNHKVFQANNDATEVVLNKLHAPLLTRFVRIRPQTWHSGIALRLELFGCRVTDA**PCSNMLGMLSGLIADSQISA**SSTQE	
5jzhA	AEPVYPDQLRLFSLGQGVCGDKYRPVNREEAQSVKSNIVGMMGQWQISGLANGWVIMGPGYNGEIKPGTASNTWCYPTNPVTGEIPTLSALDIPDGDEVDVQWRLVHDSANFIKPTSYLAHYLGYAWVGGNHSQYVGEDMDVTRDGDGWVIRGNNDGGCDGYRCGDKTAIKVSNFAYNLDPDSFKHGDVTQSDRQLVKTVVGWAVNDSDTPQSG**YDVTLRGDTATNWSKTNTYGLSEKVTTKNFKWPLVGETELSIEIAANQSWASQNGGSTTTSLS**QSVRPTVPARSKIPVKIELYKADISYPYEFKADVSYDLTLSGFLRWGGNAWYTHPDNRPNWNHTFVIGPYKDKASSIRYQWDKRYIPGEVKWWDWNWTIQQNGLSTMQNNLARVLRPVRAGITGDFSAESQFAGNIEIGAPVPLAA	0.81
4pyiA	MGDTKEQRILNHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGKIVDAVIQEHQPSVLLELGAYCGYSAVRMARLLSPGARLITIEINPDCAAITQRMVDFAGVKDKVTLVVGASQDIIPQLKKKYDVDTLDMVFLDHWKDRYLPDTLLLEECGLLRKGTVLLADNVICPGAPDFLAHVRGSSCFECTHYQSFLEYRE**VVDGLEKAIYKGPG**SEAGP	0.86
5ineA	MGQIVTMFEALPHIIDEVINIVIIVLIIITSIKAVYNFATCGILALVSFLFLAGRSCGMYGLNGPDIYKGVYQFKSVEFDMSHLNLTMPNACSANNSHHYISMGSSGLELTFTNDSILNHNFCNLTSAFNKKTFDHTLMSIVSSLHLSIRGNSNHKAVSCDFNNGITIQYNLSFSDPQSAISQCRTFRGRVLDMFRTAFGGKYMRSGWGWAGSDGKTTWCSQTSYQYLIIQNRTWENHCRYAGPFGMSRILFAQEKTKFLTRRLAGTFTWTLSDSSGVENPGGYCLTKWMILAAELKCFGNTAVAKCNVNHDEEFCDMLRLIDYNKAALSKFKQDVESALHVFKTTVNSLISDQLLMRN**HLRDLMGVPYCNYSKFWYLEHAKTGETSVPKCWLVTNGSYLNETHFSD**QIEQEADNMITEMLRKDYIKRQGSTPLALMD	0.85
1mbyA	GSAQLLKSVFVKNVGWATQLTSGAVWVQFNDGSQLVMQAGVSSISYTSPDGQTTRYGENEKLPEYIKQKL**QLLSSILLMFSNPTPV**FQ	0.80
2a73B	VQLTEKRMDKVGKYPKELRKCCEDGMRENPMRFSCQRRTRFISLGEACKKVFLDCCNYITELRRQHARASHLGLARSNLDEDIIAEENIVSRSEFPESWLWNVEDLKEPPKNGISTKLMNIFLKDSITTWEILAVSMSDKKGICVADPFEVTVMQDFFIDLRLPYSVVRNEQVEIRAVLYNYRQNQELKVRVELLHNPAFCSLATTKRRHQQTVTIPPKSSLSVPYVIVPLKTGLQEVEVKAAVYHHFISDGVRKSLKVVPEGIRMNKTVAVRTLDPERLGREGVQKEDIPPADLSDQVPDTESETRILLQG	

	TPVAQMTEDAVDAERLKHLIVTPSGCGEQNMIGMTPTVIAVHYLDETEQWEKFGLEKRQGALELIKKGYTQQLAFRQPSSAFAAFVKRAPSTWLTAYVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDGVFQEDAPVIHQEMIGGLRNNNEKDMALTAFVLISLQEAKDICEEQVNSLPGSITKAGDFLEANYMNLQRSYTVAIAGYALAQMGRLKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSYALLALLQLKDFDFVPPVVRWLNEQRYYGGGYGSTQATFMVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITHRIHWESASLLRSEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFDLKVTIKPAPETEKRPQDAKNTMILEICTRYRGDQDATMSILDISMMTGFAPDTDDLKQLANGVDRYISKYELDKAFSDRNTLIIYLDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYY**NLEESCTRFYHPEKEDGKLNKLCRDELCRCAEENCFI**QKSDDKVTLEERLDKACEPGVDYVYKTRLVKVQLSNDFDEYIMAIEQTIKSGSDEVQVGQQRTFISPIKCREALKLEEKKHYLMWGLSSDFWGEKPNLSYIIGKDTWVEHWPEEDECQDEENQKQCQDLGAFTESMVVFGCPN	
2k0qA	VDMSNVVKTYDLQDGSKVHVFKDGKMGMENKFGKSMNMPEGKVMETRD**GTKIIMKGNEIFRLDEALRKGHS**EGG	0.88
3uyiA	HMPRVKLGTQGLEVSKLGFGCMGLSGDYNDALPEEQGIAVIKEAFNCGITFFDTSDIYGENGSNEELLGKALKQLPREKIQVGTKFGIHEIGFSGVKAKGTPDYVRSCCEASLKRLDVDYIDLFYIHRIDTTVPIEITMGELXKLVEEGKIKYVGLSEASPDTIRRAHAVHPVTALQIEYSLWTRDIEDEIVPLCRQLGIGIVPY**SPIGRGLFAGKAIKE**SLPENSVLTSHPRFVGENLEKNKQIYYRIEALSQKHGCTPVQLALAWVLHQGEDVVPIPGTTKIKNLHNNVGALKVKLTKEDLKEISDAVPLDEVAGESIHEVIAVTNWKFANTPPL	0.80
4a5wB	CFCDHYAWTQWTSCSKTCNSGTQSRHRQIVVDKYYQENFCEQICSKQETRECNWQRCPINCLLGDFGPWSDCDPCIEKQSKVRSVLRPSQFGGQPCTAPLVAFQPCIPSKLCKIEEADCKNKFRCDSGRCIARKLECNGENDCGDNSDERDCGRTKAVCTRKYNPIPSVQLMGNGFHFLAGEPRGEVLDNSFTGGICKTVKSSRTSNPYRVPANLENVGFEVQ**TAEDDLKTDFYKDLTS**LGHNENQQGSFSSQGGSSFSVPIFYSSKRSENINHNSAFKQAIQASHKKDSSFIRIHKVMKVLNFTTKAKDLHLSDVFLKALNHLPLEYNSALYSRIFDDFGTHYFTSGSLGGVYDLLYQFSSEELKNSGLTEEEAKHCVRIETKKRVLFAKKTKVEHRCTTNKLSEKHEGSFIQGAEKSISLIRGGRSEYGAALAWEKGSSGLEEKTFSEWLESVKENPAVIDFELAPIVDLVRNIPCAVTKRNNLRKALQEYAAKFDPCQCAPCPNNGRPTLSGTECLCVCQSGTYGENCEKQSPDYKSNAVDGQWGCWSSWSTCDATYKRSRTRECNNPAPQRGGKRCEGEKRQEEDCTFSIMENNGQPCINDDEEMKEVDLPEIEADSGCPQPVPPENGFIRNEKQLYLVGEDVEISCLTGFETVGYQYFRCLPDGTWRQGDVECQRTECIKPVVQEVLTITPFQRLYRIGESIELTCPKGFVVAGPSRYTCQGNSWTPPISNSLTCEKDTLTKLKGHCQLGQKQSGSECICMSPEEDCSHHSEDLCVFDTDSNDYFTSPACKFLAEKCLNNQQLHFLHIGSCQDGRQLEWGLERTRLSSNSTKKESCGYDTCYDWEKCSASTSKCVCLLPPQCFKGGNQLYCVKMGSSTSEKTLNICEVGTIRCANRKMEILHPGKCLA	
1rkpA	EETRELQSLAAAVVPSAQTLKITDFSFSDFELSDLETALCTIRMFTDLNLVQNFQMKHEVLCRWILSVKKNYRKNVAYHNWRHAFNTAQCMFAALKAGKIQNKLTDLEILALLIAALSHDL**DHRGVNNSYIQRSEHPLAQLYC**HSIMEHHHFDQCLMILNSPGNQILSGLSIEEYKTTLKIIKQAILATDLALYIKRRGEFFELIRKNQFNLEDPHQKELFLAMLMTACDLSAITKPWPIQQRLAELVATEFFDQGDRERKELNIEPTDLMNREKKNKIPSMQVGFIDAICLQLYEALTHVSEDCFPLLDGCRKNRQKWQALAEQQ	0.824
1ceeB	KKKISKAD**IGAPSGFKHVSHVGWDPQNGFDVN**NLDPDLRSLFSRAGISEAQLTDAETSK	1.39
3o44A	SGFASPAPANSETNTLPHVAFYISVNRAISDEECTFNNSWLWKNEKGSRPFCKDANISLIYRVNLERSLQYGIVGSATPDAKIVRISLDDDSTGAGIHLNDQLGYRQFGASYTTLDAYFREWSTDAIAQDYRFVFNASNNKAQILKTFPVDNINEK	1.10

	FERKEVSGFELGV**TGGVEVSGDGPKAKLEAR**ASYTQSRWLTYNTQDYRIERNAKNAQAVSFTWNRQQYATAESLLNRSTDALWVNTYPVDVNRISPLSYASFVPKMDVIYKASATETGSTDFIIDSSVNIRPIYNGAYKHYYVVGAHQSYHGFEDTPRRRITKSASFTVDWDHPVFTGGRPVNLQLASFNNRCIQVDAQGRLTANMCDSQQSAQSFIYDQLGRYVSASNTKLCLDGAALDALQPCNQNLTQRWEWRKGTDELTNVYSGESLGHDKQTGELGLYASSNDAVSLRTITAYTDVFNAQESSPILGYTQGKMNQQRVGQDNRLYVRAGAAIDALGSASDLLVGGNGGSLSSVDLSGVKSITATSGDFQYGGQQLVALTFTYQDGRQQTVGSKAYVTNAHEDRFDLPDAAKITQLKIWADDWLVKGVQFDLN	
1repC	MAETAVINHKKRKNSPRIVQSNDLTEAAYSLSRDQKRMLYLFVDQIRKSDGTLQEHDGICEIHVAKYAEIFGLTSAEASKDIRQALKSFAGKEVVFYRPEEDAGDEKGYESFPWFIKPAHSPSRGLYSVHINPYLIPFF**IGLQNRFTQFRLSETKE**ITNPYAMRLYESLCQYRKPDGSGIVSLKIDWIIERYQLPQSYQRMPDFRRRFLQVCVNEINSRTPMRLSYIEKKKGRQTTHIVFSFRDITSMTTG	
3kuyA	ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTV**ALREIRRYQKSTELLI**RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAYLVALFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA	0.97
2n0aD	**MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEP**EA	1.09
4m4rA	APANEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIRTYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPGVMGTCKETFNLYYYESDNDKERFIRENQFVKIDTIAADE**SFTQVDIGDRIMKL**NTEIRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKCPLTVRNLAQFPDTITGADTSSLVEVRGSCVNNSEEKDVPKMYCGADGEWLVPIGNCLCNAGHEERSGECQACKIGYYKALSTDATCAKCPPHSYSVWEGATSCTCDRGFFRADNDAASMPCTRPPSAPLNLISNVNETSVNLEWSSPQNTGGRQDISYNVVCKKCGAGDPSKCRPCGSGVHYTPQQNGLKTTKVSITDLLAHTNYTFEIWAVNGVSKYNPNPDQSVSVTVTTNQAAPSSIALVQAKEVTRYSVALAWLEPDRPNGVILEYEVKYYEKDQNERSYRIVRTAARNTDIKGLNPLTSYVFHVRARTAAGYGDFSEPLEVTTNTVPSRIIGDG	

[1] Pdb ID+chain correspond to column 1 of Table S1

[2] Amino acid sequences of PDB chains from column 1 taken from their FASTA files. Bold represents identified fold-switching region, and yellow highlighting represents regions with independent folding cooperativity identified by SEED. Bold regions that are not highlighted were not recognized as independent cooperative folding units. Those not recognized in both conformations are highlighted in red.

[3] Qualifying ratios calculated by SEED (**Methods**), a measure of independent folding cooperativity. No score is reported for sequences in which SEED could not find a cooperative folding unit.

Table S2, columns 4-6 Coincidence of fold-switching regions with cooperative folding units in proteins

PDB2[4]	PDB2 prediction[5]	QR 2[6]
5c6bF	QNITEEFYQSTCSAVSKGYLSALRTGWYTSVITIELSNIKKIKCNGTDAKIKLIKQELDKYKNAVTELQLLMQSTPATNNQARGSGSGRSLGFLLGVGSAIASGVAVSKVLHLEGEVNKIKSALLSTNKAVVSLSNGVSVLTSKVLDLKNYIDKQLLPIVNKQSCSIPNIETVIEFQQKNNRLLEITREFSVNAGVTTPVSTYMLTNSELLSLINDMPITNDQKKLMSNNVQIVRQQSYSIMSIIKEEVLAYVVQLPLYGVIDTPCWKLHTSPLCTTNTKEGSNICLTRTDRGWYCDNAGSVSFFPQAETCKVQSNRVFCDTMNSLTLPSEVNLCNVDIFNPKYDCKIMTSKTDVSSSVITSLGAIVSCYGKTKCTASNKNRGIIKTFSNGCDYVSNKGVDTVSVGNTLYYVNKQEGKSLYVKGEPIIN**FYDPLVFPSDQFDASISQVNEKINQSLAFIRKSDELLS**AIGGYIPEAPRDGQAYVRKDGEWVLLSTFL	
3zxgB	SAKAAEGYEQIEVDVVAVWKEGYVYENRGSTSVDQKITITKGM**KNVNSETRTVTATHSIGSTISTGDAFEIGSVEVSYSHSHEESQVSMTETEVYES**KVIEHTITIPPTSKFTRWQLNADVGGADIEYMYLIDEVTPIGGTQSIPQVITSRAKIIVGRQIILGKTEIRIKHAERKEYMTVVSRKSWPAATLGHSKLFKFVLYEDWGGFRIKTLNTMYSGYEYAYSSDQGGIYFDQGTDNPKQRWAINKSLPLRHGDVVTFMNKYFTRSGLCYDDGPATNVYCLDKREDKWILEVVGLVPRGSGHHHHHH	0.93
2namA	MGHHHHHH**ATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGF**HVH**EFGDNTAGCTSAGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVVHEKADDLGKGGNEESTKTGNAGSRLACGVIGIAQ**	1.05
4yhdG	MADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMAKKVFYSFIDDKNHNKKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQIS**DYYPRNSIDTKEYMSTLTYGFNGNVTGDDTGKIGGLIGANVSIGHTLKYVQPD**FKTILESPTDKKVGWKVIFNNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAAENFLDPNKASSLLSSGFSPDFATVITMDRKASKQQTNIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDRSSERYKIDWEKEEMTNLEHHHHHH	
5ly6B	MANKAVNDFILAMNYDKKKLLTHQGESIENRFIKEGNQLPDEFVVIERKKRSLSTNTSDISVTATNDSRLYPGALLVVDETLLENNPTLLAVDRAPMTYSIDLPGLASSDSFLQVEDPSNSSVRGAVNDLLAKWHQDYGQVNNVPARMQYEKITA**HSMEQLKVKFGSDFEKAANSLDIDFNAVHSGEKQIQIVNFKQIYYTVSVDAVKNPGDVFQDTVTVEDLKQRGISAERPLVYISSVAYGRQVYLKLETTSKSDEVQAAFEAAILGVKVAPQTQWKQILD**NTEVKAVILGGDPSSGARVVTGKVDMVEDLIQEGSRFTADHPGLPISYTTSFLRDNVVATFQNSTDYVETKVTAYRNGDLLLDHSGAYVAQYYITWDELSYDHQGKEVLTPKAWDRNGQDLTAHFTTSIPLKGNVRNLSVKIRECTGLAWEWWRTVYEKTDLPLVRKRTISIWGTTLYPQVEDKVEND	0.88
1jtiB	GSIGAASMEFCFDVFKELKVHHANENIFYCPIAIMSALAMVYLGAKDSTRTQINKVVRFDKLPGFGDSIEAQCGTSVNVHSSLRDILNQITKPNDVYSFSLASRLYAEERYPILPEYLQCVKELYRGGLEPINFQTAADQARELINSWVESQTNGIIRNVLQPSSVDSQTAMVLVNAIVFKGLWEKTFKDEDTQAMPFRVTEQESKPVQMMYQIGLFRVASMASEKMKILELPFASGTMSMLVLLPDEVSGLEQLESIINFEKLTEWTSSNVMEERKIKVYLPRMKMEEKYNLTSVLMAMGITDVFSSSAN**LSGISSAESLKISQAVHAAHAEINEAGTEVVGSAEAGVDAASVSEEFRADHP**FLFCIKHIATNAVLFFGRCVSP	
2vfxL	GMALQLSREQGITARGSAEIVAEFFSFGINSILYQRGIYPSETFTRVQKYGLTLLVTTDLELIKYLNNVVEQLKDWLYKSSVQKLVVVISNIESGEVLERWQFDIESDKTAKDDSAPREKSQKAIQDEIRSVIRQITATVTFLPLLEVSCSFDLLIYTDKDLVVP**EKWEESGPQFITNSEEVRLRSFTTTIHKVNSMVAYKIP**VND	

3lowA	MIQRTPKIQVYSRHPAENGKSNFLNCYVSGFHPSDIEVDLL**KNGERIEKVEHSDLSFSKDWS**FYLLYYTEFTPTEKDEYACRVNHVTLSQPKIVKWDRDM	9.28
3j7vG	MASMTGGQQMGTNQGKGVVAA**GDKLALFLKVFGGEVLTAFART**SVTTSRHMVRSISSGKSAQFPVLGRTQAAYLAPGENLDDKRKDIKHTEKVITIDGLLTADVLIYDIEDAMNHYDVRSEYTSQLGESLAMAADGAVLAEIAGLCNVESKYNENIEGLGTATVIETTQNKAALTDQVALGKEIIAALTKARAALTKNYVPAADRVFYCDPDSYSAILAALMPNAANYAALIDPEKGSIRNVMGFEVVEVPHLTAGGAGTAREGTTGQKHVFPANKGEGNVKVAKDNVIGLFMHRSAVGTVKLRDLALERARRANFQADQIIAKYAMGHGGLRPEAAGAVVFKVE	0.85
2bzyB	**PVFAKAIQKRVPCAYDKTALALEVGDIVKVTRMNINGQWEGEVNGRKGLFPFTHVKIFDPQNPDENE**	
1wp8C	AMKN**ADNINKLKSSIESTNEAVVKLQETAEKTVYVLTALQDSS**	0.98
1fzpD	AITKINDCFELLSMVTYADKLKSLIKK**EFSISFEEFAVLTYISENKEKEYYLKDIINHLNYKQPQVVKAVKILSQEDYFDKKRNEHDERTVLILVNAQQRKKIESLLSRVNKRITEANN**EIEL	1.08
1htmB	LKSTQAAIDQINGKLNRVIEK**TNEKFHQIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEKTRRQLRENAEEMGNGCFKIYHKCDNACIESIRNGTYDHDVYRDEALNNRF**QIKG	1.05
2jmrA	**ADSTITIRGYVRDNGCSVAAESTNF**TVDLMENAAKQFNNIGATTPVVPFRILLSPCGNAVSAVKVGFTGVADSHNANLLALENTVSAASGLGIQLLNEQQNQIPLNAPSSALSWTTLTPGKPNTLNFYARLMATQVPVTAGHINATATFTLEYQ	4.72
2lepA	LXITSFANPRVAQAFV**DYMATQGVILTIQQHNQSDVWLADE**SQAERVRAELARFLENPADPRYLAASWQAGHTENLYFQ	1.14
1qs8B	**LGSENDVIELDDVANIMFYGEGEVGDNHQKFM**LIFDTGSANLWVPSKKCNSSGCSIKNLYDSSKSKSYEKDGTKVDITYGSGTVKGFFSKDLVTLGHLSMPYKFIEVIDTDDLEPIYSSVEFDGILGLGWKDLSIGSIDPIVVELKNQNKIDNALFTFYLPVHDVHAGYLTIGGIEEKFYEGNITYEKLNHDLYWQIDLDVHFGKQTMEKANVIVDSGTTTITAPSEFLNKFFANLNVIKVPFLPFYVTTCDNKEMPTLEFKSANNTYTLEPEYYMNPILEVDDTLCMITMLPVDIDSNTFILGDPFMRKYFTVFDYDKESVGFAIAKN	
4n9wA	GSGAMRIGMVCPYSFDVPGGVQSHVLQLAEVLRDAGHEVSVLAPASPHVKLPDYVVSGGKAVPIPYNGSVARLRFGPATHRKVKKWIAEGDFDVLHIHEPNAPSLSMLALQAAEGPIVATF**HTSTTKSLTLSVFQGILRPYHEKIIGRIAV**SDLARRWQMEALGSDAVEIPNGVDVASFADAPLLDGYPREGRTVLFLGRYDEPRKGMAVLLAALPKLVARFPDVEILIVGRGDEDELREQAGDLAGHLRFLGQVDDATKASAMRSADVYCAPHLGGESFGIVLVEAMAAGTAVVASDLDAFRRVLADGDAGRLVPVDDADGMAAALIGILEDDQLRAGYVARASERVHRYDWSVVSAQIMRVYETVSGAGIKVQVSGAANRDETAGESV	0.90
1xtgB	MDENLEQVSGIIGNLRH**MALDMGNEIDTQNRQIDRIMEKADSNKTRIDEANQRATKM**LG	1.88
1iytA	DAEFRHDSGYEVHH**QKLVFFAEDVGSNKGAIIGLMVGGVVIA**	1.23
5c1vB	SEPKAIDPKLSTTDRVVKAVPFPPSHRLTAKEVFDNDGKPRVDILKAHLMKEGRLEESVALRIITEGASILRQEKNLLDIDAPVTVCGDIHGQFFDLMKLFEVGGSPANTRYLFLGDYVDRGYFSIECVLYLWALKILYPKTLFLLRGNHECRHLTEYFTFKQECKIKYSERVYDACMDAFDCLPLAALMNQQFLCVHGGLSPEINTLDDIRKLDRFKEPPAYGPMCDILWSDPLEDFGNEKTQEHFTHNTVRGCSYFYSYPAVCEFLQHNNLLSILRAHEAQDAGYRMYRKSQTTGFPSLIT**IFSAPNYLDVYNNKAAVLKYENNVMNIRQFN**CSPHPSWAPNFD	

4cmqB	GAASMDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIE**KILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDN**EENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	0.82
2qkeE	MAPLRKTYVLKLYVAGNTPNSVRALKTLNNILEKEFKGVYALKVIDVLKN**PQLAEEDKILATPTLAKVLPPPVRRIIGDLSNREKVLIGLDLLYEEIGDQAEDD**LGLE	0.87
2lv1A	GAMDPGQGGGTHSQWNKPSKPKTNMKHM**AGAAAAGAVVGGLGG**YMLGSAMSRPIIHFGSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCVNITIKQHTVTTTTKGENFTETDVKMMERVIEQMCITQYERESQAYYQRGSS	9.74
1zk9A	LVPRGSHMNTSELRICRINKESGPCTGGEELYLLCDKVQKEDISVVF**STASWEGRADFSQADVHRQIAIVFKTPPY**EDLEISEPVTVNVFLQRLTDGVCSEPLPFTYLPR	1.21
2kb8A	KCNTATCAT**QRLANFLVHSSNNFGAILSSTNVGSNTY**	1.15
4q79F	MQRLFLLVAVMLLSGCLTAPPKEAARPTLMPRAQSYKDLTHLPAPTGKIFVSVYNIQDETGQFKPYPASNFSTAVPQSATAMLVTALKDSRWFIPLERQGLQNLLNERKIIRAAQENGTVAINNRIPLQSLTAANIMVEGSIIGYESNVKSGGVGARYFGIGADTQYQLDQIAVNLRVVNVSTGEILSSVNTSKTILSYEVQAGVFRFIDYQRLLEGEVGYTSNEPVMLCLMSAIETGVIFLINDGID**RGLWDLQNKAERQNDILVKYRHMSVPPES**	
5wrgA	MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAAYFVGYLKPTTFMLKYDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQ	

5wrgA cont'd	PFQQFGRDVSDFTDSVRDPKTSEILDISPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHTVSLLASTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAE**VQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKY**EQYIKWPWSHPQFEK	
1eboE	MKQIEDKIEEILSKIYHIENEIAR**IKKLIGEADGLIEGLRQLANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGG**TCHILGPDCRIEPHDWTKNITDKIDQIIHDFVDKTLPDQGDNDNWWTGWRQ	0.99
1svfC	TAAVALVKANENAAAILNLKNAIQ**KTNAAVADVVQATQSLGTAVQAVQDHINSVVSPAITAAN**Y	1.07
1nqjB	GGPGNEKLKEKENNDSSD**KATVIPNFNTTM**QGSLLGDDSRDYYSFEVKEEGEVNIELDKKDEFGVTWTLHPESNINDRITYGQVDGNKVSNKVKLRPGKYYLLVYKYSGSGNYELRVNK	0.95
2uy7D	**APTIPQGQGKVTFNNTVVDAPCSISQKSAD**QSIDFGQLSKSFLEAGGVSKPMDLDIELVNCDITAFKGGNGAKKGTVKLAFTGPIVNGHSDELDTNGGTGTAIVVQGAGKNVVFDGSEGDANTLKDGENVLHYTAVVKKSSAVGAAVTEGAFSAVANFNLTYQ	
1qb3A	MYHHYHAFQGRKLTDQERARVLEFQDSIHYSPRYSDDNYEYRHVMLPKAMLKVIPSDYFNSEVGTLRILTEDEWRGLGITQSLGWEHYECHAPEPHILLFKR**PLNYEAELRAATAAA**Q	1.76
2lclA	**GAMGPKDIVDPATPYPGDKVIITEGAFEGFQAIFTEPDGEARSMLLLNLINKEIKHSVKNTEFRK**L	1.08
1nocA	NPKSLTRGPRDKPTPLEELLPHAIEFINQYYGSFKEAKIEEHLARLEAVTKEIETTGTYQLTLDELIFATKMAWRNAPRCIGRIQWSNLQVFDARNCSTAQEMFQHICRHILYATNNGNIRSAITVFPQRSDGKHDFRLWNSQLIRYAGYQMPDGTIRGDAATLEFTQLCIDLGWKPRYGRFDVLPLVLQADGQDPEVFEIPPDLVLEVTMEHPKYEWFQELGLKWYALPAVANMLLEVGGLEFPACPFNGWYMGTEIGVRDFCDTQRYNILEEVGRRMGLETHTLASLWKDRAVTEINVAVLHSFQKQNVTIMDHHTASESFMKHMQNEYRARGGCPADWIW**LVPPVSGSITPVFHQEMLNYVL**SPFYYYQIEPWKTHIWQNEHHHH	1.45
3l9qB	NSSLDQIDLLSTKSFPPCMRQLHKALRENHHLRHGGRMQYGLFLKG**IGLTLEQALQFWKQEFIKGKMDPDKFDKGYSYNIRHSFGKEGKRTD**YTPFSCLKIILSNPPSQGDYHGCPFRHSDPELLKQKLQSYKISPGGISQILDLVKGTHYQVACQKYFEMIHNVDDCGFSLNHPNQFFCESQRILNGGKDIKKE	0.89
1nrjB	GSHMGIKQKSYQPSIIIAGPQNSGKTSLLTLLTTDSVRPT**VVSQEPLSAADYDGSGVTLVDFPGHVKLR**YKLSDYLKTRAKFVKGLIFMVDSTVDPKKLTTTAEFLVDILSITESSCENGIDILIACNKSELFTARPPSKIKDALESEIQKVIERRKKSLNEVERKINEEDYAENTLDVLQSTDGFKFANLEASVVAFEGSINKRKISQWREWIDEKL	0.93
1dzlA	MSLWLPSEATVYLPPVPVSKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQYRVFRIHLPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASAYAANAGVDNRECISMDYKQTQLCLIGCKPPIGEHWGKGSPCTQVAVQPGDCPPLELINTVIQDGDMVDTGFGAMDFTTLQANKSEVPLDICTSICKYPDYIKMVSEPYGDSLFFYLRREQMFVRHLFNRAGTVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMVTSDAQIFNKPYW	0.90

1dzlA cont'd	LQRAQGHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRHGEEYDLQFIFQLCKITLTADVMTYIHSMNSTILEDWNFGLQPP**PGGTLEDTYRFVTSQAIACQKHTPPAPKEDPLKK**YTFWEVNLKEKFSADLDQFPLGRKFLLQLGLKAKPKFTLGKRKATPTTSSTSTTAKRKKRKL	
4xwsD	GAMAQLAAPLKVGAIYTIGPYLFPHLIPQLHRVAPQMPLYIEENFTHILRDKLRTGELDAIIIALPFQEADVLTKPLFDEPFYVLMPADHPW**TAKASIDSELLNDKSLLLLGEGH**D**FRDQVLEACPTVRKGDENKHTTVESSSLETIRHMVASGLGVSVLP**FSAVDSHHYAPGVIEVRPFSAPVPFRTVAIAWRASFPRPRAIEVLADSIRLCSVARPQTQEQPQIA	0.99
4zrbH	SNAM**KDFHFDAISAFENYEIEKMR**DGHVVVTTKVVNSSLNYYGNAHGGYLFTLCDQISGLVVISLGLDGVTLQSSINYLKAGKLDDVLTIKGECVHQGRTTCVMDVDITNQEGRNVCKATFTMFVTGQRSEERRVRI	0.83
4qhhA	MKDRKILNEILSNTINELNLNDKKANIKIKIKPLKRKIASISLTNKTIYINKNILPYLSDEEIRFILAHELLHLK**YGKYHINEFEEELLFLFPN**KEAILINLINKLHQKK	0.97
4aalA	WSHPQFEKGAETAVPNSEDVMKRAQGLFKPIPAKPPVMKDNPASPSRVELGRMLFFDPRLSASHLISCNTCHNVGLGGTDILETSIGHGWQKGPRNSPTVLNAVYNIAQFWDGRAEDLAAQAKGPVQASVEMNNKPENLVATLKSIPGYPPLFRKAFPGQGDPVTFDNVAKAIEVFEATLVTPDAPFDKYLKGNRKAISSTAEQGLALFLDKGCAACHSGVNMGGT**GYFPFGVREDPGPVVRPVDDTGRYKVTSTAADKY**VFRSPSLRNVAITMPYFHSGKVWKLKDAVKIMGSAQLGISITDADADKIVTFLNTLTGAQPKVMHPVLPPNSDDTPRPVSN	
1x0gD	MVELTPAAIQELERLQTHGVRRGQAAILRIQVQPSECGDWRY**DLALVAEPKPTDLLTQSQGWTIAIAAEAAELLRGLRVDYIEDLMGGAFRFHNPNASQTCGCGMAFRVS**RS	1.10
4ow6B	GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKGFYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQTALEHP**ELSELKTVTGTN**PVFAGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVAFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS	9.76
1mnmD	QLTQKNKSADGLV**FNVVTQDMINKST**KPYRGHRFTKENVRILESWFAKNIENPYLDTKGLENLMKNTSLSRIQIKNWVSNRRRKEKT	1.41
2mwfA	**SEWTERKTADGKTYYYNNRTLESTWEKPQELK**	1.00
5hk5H	RKNRPAGAIPSPYKDGSSNNSERW**HHQIKEVLASSQEALVVTERKY**LKSDWCKTQPLRQTVSEEGCRSRTILNRFCYGQCNSFYIPRHVKKEEDSFQSCAFCKPQRVTSVIVELECPGLDPPFRIKKIQKVKHCRCMSVNLSDSDKQ	9.23
5et5A	TDRSPFET**DMLTLTRYVMEK**LTAIKAISSAVRKAGLAHLYGIAGSVNVTGDEVKKLDVLSNSLVINMLQSSYSTCVLVSEENKDAIITAKEKRGKYVVCFDPLDGSSNIDCLASIGTIFAIYRKTSEDEPSEKDALQCGRNIVAAGYALYGSATLVALSTGQGVDLFMLDPALGEFVLVEKDVKIKKKGKIYSLNEGYAKYFDAATTEYVQKKKFPEDGSAPYGARYVGSMVADVHRTLVYGGIFLYPANQKSPKGKLRLLYECNPVAYIIEQAGGLATTGTQPVLDVKPEAIHQRVPLILGSPEDVQEYLTCVQKNQAGS	5.16

2a01C	DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLNLKLLDNWDSVTSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQ**EKLSPLGEEMRDRARAHVDALRTHLAPYSDELRQRLAAR**LEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQ	0.91
4g0dZ	GGD**EDDLSEEDLQFAERYLRSYYHPT**	6.15
1qlnA	MNTINIAKNDFSDIELAAIPFNTLADHYGERLAREQLALEHESYEMGEARFRKMFERQLKAGEVADNAAAKPLITTLLPKMIARINDWFEEVKAKRGKRPTAFQFLQEIKPEAVAYITIKTTLACLTSADNTTVQAVASAIGRAIEDEARFGRIRDLEAKHFKKNVEEQLNKRVGHVYKKAFMQVVEADMLSKGLLGGEAWSSWHKEDSIHVGVRCIEMLIESTGMVSLHRQNAGVVGQDSETIELAPEYAEAIATRAGALAGISPMFQPCVVPPKPWTGITGGGYWANGRRPLALVRTHSKKALMRYEDVYMPEVYKAINIAQNTAWKINKKVLAVANVITKWKHCPVEDIPAIEREELPMKPEDIDMNPEALTAWKRAAAAVYRKDKARKSRRISLEFMLEQANKFANHKAIWFPYNMDWRGRVYAVSMFNPQGNDMTKGLLTLAKGKPIGKEGYYWLKIHGANCAGVDKVPFPERIKFIEENHENIMACAKSPLENTWWAEQDSPFCFLAFCFEYAGVQHHGLSYNCSLPLAFDGSCSGIQHFSAMLRDEVGGRAVNLLPSETVQDIYGIVAKKVNEILQADAIN**GTDNEVVTVTDENTGE**ISEKVKLGTKALAGQWLAYGVTRSVTKRSVMTLAYGSKEFGFRQQVLEDTIQPAIDSGKGLMFTQPNQAAGYMAKLIWESVSVTVVAAVEAMNWLKSAAKLLAAEVKDKKTGEILRKRCAVHWVTPDGFPVWQEYKKPIQTRLNLMFLGQFRLQPTINTNKDSEIDAHKQESGIAPNFVHSQDGSHLRKTVVWAHEKYGIESFALIHDSFGTIPADAANLFKAVRETMVDTYESCDVLADFYDQFADQLHESQLDKMPALPAKGNLNLRDILESDFAFA	8.76
5bmyA	MEKLPPG**WEKRMSRSSGR**GSGMKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTN	0.92
1xjuB	RTNQAGLELI**GNAEGCRRDPYMCPAGVWTDGIGNTHGVTPG**VRKTDQQIAADWEKNILIAERCINQHFRGKDMPDNAFSAMTSAAFNMGCNSLRTYYSKARGMRVETSIHKWAQKGEWVNMCNHLPDFVNSNGVPLRGLKIRREKERQLCLTGLVNEHHHHHH	0.91
3hdfA	MGPSGNDGLEGVSYIPYKDIVGVWTVCHGHTGKDIMLGKTYTKAECKALLNKDLATVARQINPYIKVDIPETMRGALYSFVYNVGAGNFRTSTLLRKINQGDIKGACD**QLRRWTYAGGKQ**WKGLMTRREIEREICLWGQQ	0.89
3meeA	PISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWET**WWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETF**YVDGAANRETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVL	0.81

4ydqB	GAMAITSKKIENFSDWYTQVIVKSELIEYYDISGCYILRPAAYYIWECVQAFFNKEIKKLNVENSYFPLFVTKNKLEKEK**NHIEGFSPEVAWVTKYGDSNLPEEIAIRPTSETIMYSV**FPKWIRSYRDLPLKLNQWNTVVRWEFKQPTPFIRTREFLWQEGHTAHKNEEEAVKLVFDILDLYRRWYEEYLAVPIIKGIKSEGEKFGGANFTSTAEAFISENGRAIQAATSHYLGTNFAKMFKIEFEDENEVKQYVHQTSWGCTTRSIGIMIMTHGDDKGLVLPPNVSKYKVVIVPIFYKTTDENAIHSYCKDIEKILKNAQINCVYDDRASYSPGYKFNHWELRGIPIRIEVGPKDLQNNSCVIVRRDNNEKCNVKKESVLLETQQMLVDIHKNLFLKAKKKLDDSIVQVTSFSEVMNALNKKKMVLAPWCEDIATEEEIKKETQRLSLNQTNSETTLSGAMKPLCIPLDQPPMPPNMKCFWSGKPAKRWCLFGRSY	0.83
3m7pA	QPHPQPPPYGHCVTDSGVVYSVGMQWLKTQGNKQMLCTCLGNGVSCQETAVTQTYGGNSNGEPCVLPFTYNGRTFYSCTTEGRQDGHLWCSTTSNYEQDQKYSFCTDHTVLVQTRGGNSNGALCHFPFLYNNHNYTDCTSEGRRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQWDKQHDMGHMMRCTCVGNGRGEWTCIAYSQLR**DQCIVDDITYNVNDTFHKRHEEG**HMLNCTCFGQGRGRWKCDPVDQCQDSETGTFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQPLQT	0.98
1rk4B	GSMAEEQPQVELFVKAGSDGAKIGNCPFSQRLFMVLWL**KGVTFNVTTVDTKRRTETVQKLCPGGQLPFLLYGTEVHTDTN**KIEEFLEAVLCPPRYPKLAALNPESNTAGLDIFAKFSAYIKNSNPALNDNLEKGLLKALKVLDNYLTSPLPEEVDETSAEDEGVSQRKFLDGNELTLADCNLLPKLHIVQVVCKKYRGFTIPEAFRGVHRYLSNAYAREEFASTCPDDEEIELAYEQVAKALK	0.88
3lqcA	MPEIRLRHVVSCSSQD**STHCAENLLKADTYRKWRAAKAG**EKTISVVLQLEKEEQIHSVDIGNDGSAFVEVLVGSSAGGAGEQDYEVLLVTSSFMSPSESRSGSNPNRVRMFGPDKLVRAAAEKRWDRVKIVCSQPYSKDSPFGLSFVRFHSPPDKDEAEAPSQKVTVTKLGQFRVKEEDESANHHHHHH	0.78
2grmB	MFKIGSVLKQIRQELNYHQIDLYSGIMSKSVYIKVEADSRPISVEELSKFSERLGVNFFEILNRAGMNTKSVNETGKEKLLISKIFTNPDLFDKNFQRIEPKRLTSLQYFSIYLGYISIAHHYNIEVPTFNKTITSDLKHLYDKRTTFFGIDYEIVSNLLNVLPYEEVSSIIKPMYPIVDSFGKDYDLTIQTVLKNALTISIMNRNLKEAQYYINQFEHLKTIKNISINGCYDLEINYLKQIYQFLTDKNIDSYLNAVNIINIFKIIGKEDIHRSLVEELTKISAKEKFTPP**KEVTMYYENYVAIENNPIPEIK**EQS	3.89
4gqcB	MKGLVELGEKAPDFTLPNQDFEPVNLYEVLKRGRPAVLIFFPAAF**SPVCTKELCTFRDKMAQLEKANA**EVLAISVDSPWCLKKFKDENRLAFNLLSDYNREVIKLYNVYHEDLKGLKMVAKRAVFIVKPDGTVAYKWVTDNPLNEPDYDEVVREANKIAGELVA	0.83
4o01D	MASMTGGQQMGRGSMSRDPLPFFPPLYLGGPEITTENCEREPIHIPGSIQPHGALLTADGHSGEVLQMSLNAATFLGQEPTVLRGQTLAALLPEQWPALQAALPPGCPDALQYRATLDWPAAGHLSLTVHRVGELLILEFEPTEAWDSTGPHALRNAMFALESAPNLRALAEVATQTVRELTGFDRVMLYKFAPDATGEVIAEARREGLHAFLGHRFPASDIPAQARALYTRHLLRLTADTRAAAVPLDPVLNPQTNAPTPLGGAVLRATSPMHMQYLRNMGVGSSLSVSVVVGGQLWGLIACHHQTPYVLPPDLRTTLEYLGRLLSLQVQVKEAADVAAFRQSLREHHARVALAAAHSLSPHDTLSDPALDLLGLMRAGGLILRFEGRWQTLGEVPPAPAVDALLAWLETQPGALVQTDALGQLWPAGADLAPSAAGLLAISVGEGWSECLVWLRPELRLEVAWGGATPDQAKDDLGP**RHSFDTYLEEKRG**YAEPWHPGEIEEAQDLRDTLTGALEHHHHHH	0.79
4dxrA	GPGGSGGVTEEQVHHIVKQALQRYSEDRIGLADYALESGGAS**VISTRCSETYETKTALLSLFGIPLWYHSQS**PRVILQPDVHPGNCWAFQGPQGFAVVRLSARIRPTAVTLEHVPKALSPNSTISSAPKDFAIFGFDEDLQQEGTLLGKFTYDQDGEPIQTFHFQAPTMATYQVVELRILTNWGHPEYTCIYRFRVHGEPAH	0.80

4rwqB	MELRHTPARDLDKFIEDHLLPNTCFRTQVKEAIDIVCRFLKERCFQGTADPVRVSKVVKG**GSSGKGTTLRGRSD**ADLVVFLTKLTSFEDQLRRRGEFIQEIRRQLEACQREQKFKVTFEVQSPRRENPRALSFVLSSPQLQQEVEFDVLPAFDALGQWTPGYKPNPEIYVQLIKECKSRGKEGEFSTCFTELQRDFLRNRPTKLKSLIRLVKHWYQTCKKTHGNKLPPQYALELLTVYAWEQGSRKTDFSTAQGFQTVLELVLKHQKLCIFWEAYYDFTNPVVGRCMLQQLKKPRPVILDPADPTGNVGGGDTHSWQRLAQEARVWLGYPCCKNLDGSLVGAWTMLQKIGSHHHHHH	
2n54B	VGSEVSDKRTCVSLTTQRLPVSRIKTYTITEGSLRCVIFITKRGLKVCC**DPQATWVRDVVRSMDRKSNTRN**NMIQTKPTGTQQSTNTAVTLTG	1.57
3vpaD	GHMATLKVIGVGGGGNNAVNRMIDHGMNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQGADMVFVTSGMGGGTGTGAAPVVAKIAKEMGALTVGVVTRPFSFEGRKRQTQAAAGVEAMKAAVDTLIVIPNDRLLDIVDKSTPMMEAFKEADNVLRQ**GVQGISDLIAVSGE**VNLDFADVKTIMSNQGSALMGIGVSSGENRAVEAAKKAISSPLLETSIVGAQGVLMNITGGESLSLFEAQEAADIVQDAADEDVNMIFGTVINPELQDEIVVTVIATGFD	0.84
2p3vD	MDRLDFSIKLLRKVGHLLMIHWG**RVDNVEKKTGFKDIVT**EIDREAQRMIVDEIRKFFPDENIMAEEGIFEKGDRLWIIDPIDGTINFVHGLPNFSISLAYVENGEVKLGVVHAPALNETLYAEEGSGAFFNGERIRVSENASLEECVGSTGSYVDFTGKFIERMEKRTRRIRILGSAALNAAYVGAGRVDFFVTWRINPWDIAAGLIIVKEAGGMVTDFSGKEANAFSKNFIFSNGLIHDEVVKVVNEVVEEIGGK	0.94
4tsyD	**SADVAGAVIDGAGLG**FDVLKTVLEALGNVKRKIAVGIDNESGKTWTAMNTYFRSGTSDIVLPHKVAHGKALLYNGQKNRGPVATGVVGVIAYSMSDGNTLAVLFSVPYDYNWYSNWWNVRVYKGQKRADQRMYEELYYHRSPFRGDNGWHSRGLGYGLKSRGFMNSSGHAILEIHVTKA	0.95
3g0hA	SMEDRAAQSLLNKLIRSNLVDN**TNQVEVLQRDPN**SPLYSVKSFEELRLKPQLLQGVYAMGFNRPSKIQENALPLMLAEPPQNLIAQSQSGTGKTAAFVLAMLSQVEPANKYPQCLCLSPTYELALQTGKVIEQMGKFYPELKLAYAVRGNKLERGQKISEQIVIGTPGTVLDWCSKLKFIDPKKIKVFVLDEADVMIATQGHQDQSIRIQRMLPRNCQMLLFSATFEDSVWKFAQKVVPDPNVIKLKREEETLDTIKQYYVLCSSRDEKFQALCNLYGAITIAQAMIFCHTRKTASWLAAELSKEGHQVALLSGEMMVEQRAAVIERFREGKEKVLVTTNVCARGIDVEQVSVVINFDLPVDKDGNPDNETYLHRIGRTGRFGKRGLAVNMVDSKHSMNILNRIQEHFNKKIERLDTDDLDEIE	0.90
5lj3M	MTS**WRDKSAKVQVKESELPSSIPAQTGLTFNIWYNKWSQGFAGNT**RFVSPFALQPQLHSGKTRGDNDGQLFFCLFFAKGMCCLGPKCEYLHHIPDEEDIGKLALRTEALDCFGREKFADYREDMGGIGSFRKKNKTLYVGGIDGALNSKHLKPAQIESRIRFVFSRLGDIDRIRYVESKNCGFVKFKYQANAEFAKEAMSNQTLLLPSDKEWDDRREGTGLLVKWANEDPDPAAQKRLQEELKLESLNMMVHLINNNTNSAGTEVNNKNNERLDRTFPEASVDNVKKRLLPLDNGMESDDFIEKLKKVKKNISRENISSKPSVGKLGGPLLDYLSSDED	0.88
2pbkB	QGLYVGGFV**DVVSCPKLEQELYLDPDQVTDYL**PVTEPLPITIEHLPETEVGWTLGLFQVSHGIFCTGAITSPAFLELASRLADTSHVARAPVKNLPKEPLLEILHTWLPGLSLSSIHPRELSQTPSGPVFQHVSLCALGRRRGTVAVYGHDAEWVVSRFSSVSKSERAHILQHVSSCRLEDLSTPNFVSPLETLMAKAIDAGFIRDRLDLLKTDRGVASILSPVYLKA	
4rmbB	RIMKLDDERQTETYITQINPEGKEMYFASGLGNLYTIIGSDGTSGSPVNLLNAEVKILKTNSKNLTDSMDQNYDSPEFEDVTSQYSYTNDGSKITIDWKTNSISSTTSYVVLVKIPKQSGVLYSTVSDINQTYGSKYSYGHTN**ISGDSDANAEIKLLSLE**HHHHHH	

3kdsG	MATMYKPSGNKRVTFKDVGGAEEAIEELKEVVEFLKDPSKFNRIGARMPKGILLVGPPGTGATLLARAVAGEANVPFFHISGSDFVELFVGVGAARVRDLFAQAKAHAPCIVFIDEIDAVGRHRGAGLGGGHDEREQTLNQLLVEMDGFDSKEGIIVMAATNRPDILDPALLRPGRFDKKIVVDPPDMLGRKKILEIHTRNKPLAEDVNLEIIAKRTPGFVGADLENLVNEAALLAAREGRDKITMKDFEEAIDRVIAGPARKSLLISPAEKRIIAYHEAGHAVVSTVVPNGEPVHRISII**PRGYKALGYTLHLPEED**KYLVSRNELLDKLTALLGGRAAEEVVFGDVTSGAANDIERATEIARNMVCQLGMSEELGPLAWGKEEQEVFLGKEITRLRNYSEEVASKIDEEVKKIVTNCYERAKEIIRKYRKQLDNIVEILLEKETIEGDELRRILSEEFEKVVE	0.92
2wcdX	MHHHHHHTEIVADKTVEVVKNAIETADGALDLYNKYLDQVIPWQTFDETIKELSRFKQEYSQAASVLVGDIKTLLMDSQDKYFEATQTVYEWCGVATQLLAAYILLFDEYNEKKASAQKDILIKVLDDGITKLNEAQKSLLVSSQSFNNASGKLLALDSQLTNDFSEKSSYFQSQVDKIRKEAY**AGAAAGVVAGPFGLIIS**YSIAAGVVEGKLIPELKNKLKSVQNFFTTLSNTVKQANKDIDAAKLKLTTEIAAIGEIKTETETTRFYVDYDDLMLSLLKEAAKKMINTCNEYQKRHGKKTLFEVPEV	0.92
1kctA	EDPQGDAAQKTDTSHHDQDHPTFNKITPNLAEFAFSLYRQLAHQSNSTNIFFSPVSIAAAFAMLSLGAKGDTHDEILEGLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGNGLFLSEGLKLVDKFLEDVKKLYHSEAFTVNFGDTEEAKKQINDYVEKGTQGKIVDLVKELDRDTVFALVNYIFFKGKWERPFEVKDTEEEDFHVDQVTTVKVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNATAIFFLPDEGKLQHLENELTHDIITKFLENED**RRSASLHLPKLSITGTYDLKSVLGQLGITKVFSNGADLSGVTEEAPLKLSKAVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEV**KFNKPFVFLMIEQNTKSPLFMGKVVNPTQK	
3q8fA	EVKQENRLLNESESSSQGLLGYYFSDLNFQAPMVVTSSTTGDLSIPSSELENIPSENQYFQSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQYQRENPTEKGLDFKLYWTDSQNKKEVISSDNLQLPELKQKSSNSRKKRSTSAGPTVPDRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSTASDPYSDFEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSQTRTISKNTSTSRTHTSEVHGNAEVHASFFDIGGSVSAGFSNSNSSTVAI**DHSLSLAGERTWAETMGLNTAD**TARLNANIRYVNTGTAPIYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPSKNLAPIALNAQDDFSSTPITMNYNQFLELEKTKQLRLDTDQVYGNIATYNFENGRVRVDTGSNWSEVLPQIQETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNLQYQGKDITEFDFNFDQQTSQNIKNQLAELNATNIYTVLDKIKLNAKMNILIRDKRFHYDRNNIAVGADESVVKEAHREVINSSTEGLLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNISSLRQDGKTFIDFKKYNDKLPLYISNPNYKVNVYAVTKENTIINPSENGDTSTNGIKKILIFSKKGYEIG	
1jfkA	MAEALFKEIDVNGDGAVSYEEVKAFVSKKRAIKNEQLLQLIFKSID**ADGNGEIDQNEFAKFYGSIQGQDLSDDKIGLKVLYKLMDVDGDGKLTKEEVTSFFKKHGIEKVAEQVMKADANGDGYITLEEFLEFS**L	0.94
5l35G	PNNLDSNVSQIVLKKFLPGFMSDLVLAKTVDRQLLAGEINSSTGDSVSFKRPHQFSSLRTPTGDISGQNKNNLISGKATGRVGNYITVAVEYQQLEEAIKLNQLEEILAPVRQRIVTDLETELAHFMMNNGALSLGSPNTPITKWSDVAQTASFLKDLGVNEGENYAVMDPWSAQRLA**DAQTGLHASDQLVRTAWENAQIPTNFGG**IRALMSNGLASRTQGAFGGTLTVKTQPTVTYNAVKDSYQFTVTLTGATASVTGFLKAGDQVKFTNTYWLQQQTKQALYNGATPISFTATVTADANSDSGGDVTVTLSGVPIYDTTNPQYNSVSRQVEAGDAVSVVGTASQTMKPNLFYNKFFCGLGSIPLPKLHSIDSAVATYEGFSIRVHKYADGDANVQKMRFDLLPAYVCFNPHMGGQFFGNP	0.88
5i2sA	KFTIVFPHNQKGNWKNVPSNYHYCPSSSDLNWHNDLIGTALQVKMPKSHKAIQADGWMCHASKWVTTCDFRWYGPKYITHSIRSFTPSVEQCKESIEQTKQGTWLNPGFPPQSCGYATVTDAEAVIVQVTPHHVLVDEYTGEWVDSQFINGKCSNY	

5i2sA cont'd	ICPTVHNSTTWHSDYKVKGLCDSNLISMDITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVRLPSGVWFEMADKDLFAAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAGLPISPVDLSYLAPKNPGTGPAFTIINGTLKYFETRYIRVDIAAPILSRMVGMISGTTTERELWDDWAPYEDVEIGPNGVLRTSSGYKF**PLYMIGHGMLDSDLHLSSKAQVFEHPHIQDAASQLPDDES**	
5f5rB	AGHSTQTAEDKEEPLHSIISSTESVQ**GSTSKHEFQAETKKLLDIVARSLYS**EKEVFIRELISNASDALEKLRHKLVSDGQALPEMEIHLQTNAEKGTITIQDTGIGMTQEELVSNLGTIARSGSKAFLDALQNQAEASSKIIGQFGVGFYSAFMVADRVEVYSRSAAPGSLGYQWLSDGSGVFEIAEASGVRTGTKIIIHLKSDCKEFSSEARVRDVVTKYSNFVSFPLYLNGRRMNT	
2qqjA	GSHMFQCNVPLGMESGRIANEQISASSTYSDGRWTPQQSRLHGDDNGWTPNLDSNKEYLQVDLRFLTMLTAIATQGAISRETQNGYYVKSYKLEVSTNGEDWMVYRHGKNHKVFQANNDATEVVLNKLHAPLLTRFVRIRPQTWHSGIALRLELFGCRVTDA**PCSNMLGMLSGLIADSQISA**SSTQEYLWSPSAARLVSSRSGWFPRIPQAQPGEEWLQVDLGTPKTVKGVIIQGARGGDSITAVEARAFVRKFKVSYSLNGKDWEYIQDPRTQQPKLFEGNMHYDTPDIRRFDPIPAQYVRVYPERWSPAGIGMRLEVLGCDWT	
5jztG	AEPVYPDQLRLFSLGQGVCGDKYRPVNREEAQSVKSNIVGMMGQWQISGLANGWVIMGPGYNGEIKPGTASNTWCYPTNPVTGEIPTLSALDIPDGDEVDVQWRLVHDSANFIKPTSYLAHYLGYAWVGGNHSQYVGEDMDVTRDGDGWVIRGNNDGGCDGYRCGDKTAIKVSNFAYNLDPDSFKHGDVTQSDRQLVKTVVGWAVNDSDTPQSG**YDVTLRYDTATNWSKTNTYGLSEKVTTKNKFKWPLVGETELSIEIAANQSWASQNGGSTTTSLS**QSVRPTVPARSKIPVKIELYKADISYPYEFKADVSYDLTLSGFLRWGGNAWYTHPDNRPNWNHTFVIGPYKDKASSIRYQWDKRYIPGEVKWWDWNWTIQQNGLSTMQNNLARVLRPVRAGITGDFSAESQFAGNIEIGAPVPLAA	1.00
4pyjA	MGDTKEQRILNHVLQHAEPGNAQSVLEAIDTYCEQKEWAMNVGDKKGKIVDAVIQEHQPSVLLELGAYCGYSAVRMARLLSPGARLITIEINPDCAAITQRMVDFAGVKDKVTLVVGASQDIIPQLKKKYDVDTLDMVFLDHWKDRYLPDTLLLEECGLLRKGTVLLADNVICPGAPDFLAHVRGSSCFECTHYQSFLEYRE**VVDGLEKAIYKGPG**SEAGP	4.24
3mkoA	HHHHHHIEGRDEEFSDMLRLIDYNKAALSKFKQDVESALHVFKTTVNSLISDQLLMRN**HLRDLMGVPYCNYSKFWYLEHAKTGETSVPKCWLVTNGSYLNETHFSD**QIEQEADNMITEMLRKDYIKRQGSTPLALMD	1.02
4yypA	SAQLLKSVFVKNVGWATQLTSGAVWVQFNDGSQLVVQAGVSSISYTSPNGQTTRYGENEKLPDYIKQKL**QCLSSILLMFSNPTP**NFH	1.02
3l5nB	SNLDEDIIAEENIVSRSEFPESWLWNVEDLKEPPKNGISTKLMNIFLKDSITTWEILAVSMSDKKGICVADPFEVTVMQDFFIDLRLPYSVVRNEQVEIRAVLYNYRQNQELKVRVELLHNPAFCSLATTKRRHQQTVTIPPKSSLSVPYVIVPLKTGLQEVEVKAAVYHHFISDGVRKSLKVVPEGIRMNKTVAVRTLDPERLGREGVQKEDIPPADLSDQVPDTESETRILLQGTPVAQMTEDAVDAERLKHLIVTPSGCGEQNMIGMTPTVIAVHYLDETEQWEKFGLEKRQGALELIKKGYTQQLAFRQPSSAFAAFVKRAPSTWLTAYVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDGVFQEDAPVIHQEMIGGLRNNNEKDMALTAFVLISLQEAKDICEEQVNSLPGSITKAGDFLEANYMNLQRSYTVAIAGYALAQMGRLKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSYALLALLQLKDFDFVPPVVRWLNEQRYYGGGYGSTQATFMVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITHRIHWESASLLRSEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFDLKVTIKPAPETEKRPQDAKNTMILEICTRYRGDQDATMSILDISMMTGFAPDTDDLKQLANGVDRYISKYELDKAFSDRNTLIIYLDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYY**NLEESCTRFYHPEKEDGKLNKLCRDELCRCAEENCFI**QKSDDKVT	

3l5nB cont’d	LEERLDKACEPGVDYVYKTRLVKVQLSNDFDEYIMAIEQTIKSGSDEVQVGQQRTFISPIKCREALKLEEKKHYLMWGLSSDFWGEKPNLSYIIGKDTWVEHWPEEDECQDEENQKQCQDLGAFTESMVVFGCPN	
2lelA	VDMSNVVKTYDLQDGSKVHVFKDGKMGMENKFGKSMNMPEGKVMETRD**GTKIIMKGNEIFRLDEALRKGHS**EGG	0.92
3v0tA	MPRVKLGTQGLEVSKLGFGCMGLSGDYNDALPEEQGIAVIKEAFNCGITFFDTSDIYGENGSNEELLGKALKQLPREKIQVGTKFGIHEIGFSGVKAKGTPDYVRSCCEASLKRLDVDYIDLFYIHRIDTTVPIEITMGELKKLVEEGKIKYVGLSEASPDTIRRAHAVHPVTALQIEYSLWTRDIEDEIVPLCRQLGIGIVPY**SPIGRGLFWGKAIKE**SLPENSVLTSHPRFVGENLEKNKQIYYRIEALSQKHGCTPVQLALAWVLHQGEDVVPIPGTTKIKNLHNNVGALKVKLTKEDLKEISDAVPLDEVAGESIHEVIAVTNWKFANTPPLK	0.82
3t5oA	CFCDHYAWTQWTSCSKTCNSGTQSRHRQIVVDKYYQENFCEQICSKQETRECNWQRCPINCLLGDFGPWSDCDPCIEKQSKVRSVLRPSQFGGQPCTAPLVAFQPCIPSKLCKIEEADCKNKFRCDSGRCIARKLECNGENDCGDNSDERDCGRTKAVCTRKYNPIPSVQLMGNGFHFLAGEPRGEVLDNSFTGGICKTVKSSRTSNPYRVPANLENVGFEVQ**TAEDDLKTDFYKDLTS**LGHNENQQGSFSSQGGSSFSVPIFYSSKRSENINHNSAFKQAIQASHKKDSSFIRIHKVMKVLNFTTKAKDLHLSDVFLKALNHLPLEYNSALYSRIFDDFGTHYFTSGSLGGVYDLLYQFSSEELKNSGLTEEEAKHCVRIETKKRVLFAKKTKVEHRCTTNKLSEKHEGSFIQGAEKSISLIRGGRSEYGAALAWEKGSSGLEEKTFSEWLESVKENPAVIDFELAPIVDLVRNIPCAVTKRNNLRKALQEYAAKFDPCQCAPCPNNGRPTLSGTECLCVCQSGTYGENCEKQSPDYKSNAVDGQWGCWSSWSTCDATYKRSRTRECNNPAPQRGGKRCEGEKRQEEDCTFSIMENNGQPCINDDEEMKEVDLPEIEADSGCPQPVPPENGFIRNEKQLYLVGEDVEISCLTGFETVGYQYFRCLPDGTWRQGDVECQRTECIKPVVQEVLTITPFQRLYRIGESIELTCPKGFVVAGPSRYTCQGNSWTPPISNSLTCEKDTLTKLKGHCQLGQKQSGSECICMSPEEDCSHHSEDLCVFDTDSNDYFTSPACKFLAEKCLNNQQLHFLHIGSCQDGRQLEWGLERTRLSSNSTKKESCGYDTCYDWEKCSASTSKCVCLLPPQCFKGGNQLYCVKMGSSTSEKTLNICEVGTIRCANRKMEILHPGKCLA	
2h44A	EETRELQSLAAAVVPSAQTLKITDFSFSDFELSDLETALCTIRMFTDLNLVQNFQMKHEVLCRWILSVKKNYRKNVAYHNWRHAFNTAQCMFAALKAGKIQNKLTDLEILALLIAALSHDL**DHRGVNNSYIQRSEHPLAQLYC**HSIMEHHHFDQCLMILNSPGNQILSGLSIEEYKTTLKIIKQAILATDLALYIKRRGEFFELIRKNQFNLEDPHQKELFLAMLMTACDLSAITKPWPIQQRIAELVATEFFDQGDRERKELNIEPTDLMNREKKNKIPSMQVGFIDAICLQLYEALTHVSEDCFPLLDGCRKNRQKWQALAEQQ	0.88
2k42A	GHMS**GFKHVSHVGWDPQNGFDVNN**LDPDLRSLFSRAGISEAQLTDAETSKLIYDFIEDQGGLEAVRQEMRRQ	0.96
1xezA	GAMGSNINEPSGEAADIISQVADSHAIKYYNAADWQAEDNALPSLAELRDLVINQQKRVLVDFSQISDAEGQAEMQAQFRKAYGVGFANQFIVITEHKGELLFTPFDRTEEIDPALLEAPRTAALLGASGFASPAPANSETNTLPHVAFYISVNRAISDEECTFNNSWLWKNEKGSRPFCKDANISLIYRVNLERSLQYGIVGSATPDAKIVRISLDDDSTGAGIHLNDQLGYRQFGASYTTLDAYFREWSTDAIAQDYRFVFNASNNKAQILKTFPVDNINEKFERKEVSGFELGV**TGGVEVSGDGPKAKLEAR**ASYTQSRWLTYNTQDYRIERNAKNAQAVSFTWNRQQYATAESLLNRSTDALWVNTYPVDVNRISPLSYASFVPKMDVIYKASATETGSTDFIIDSSVNIRPIYNGAYKHYYVVGAHQSYHGFEDTPRRRITKSASFTVDWDHPVFTGGRPVNLQLASFNNRCIQVDAQGRLTANMCDSQQSAQSFIYDQLGRYVSASNTKLCLDGAALDALQPCNQNLTQRWEWRKGTDELTNVYSGESLGHDKQTGELGLYASSNDAVSLRTITAYTDVFNAQESSPILGYTQGKMNQQRVGQDNRLYVRAGAAIDAL	

1xezA cont'd	GSASDLLVGGNGGSLSSVDLSGVKSITATSGDFQYGGQQLVALTFTYQDGRQQTVGSKAYVTNAHEDRFDLPDAAKITQLKIWADDWLVKGVQFDLN	
2z9oB	MRGSHHHHHHGSIEGRAETAVINHKKRKNSPRIVQSNDLTEAAYSLSRDQKRMLYLFVDQIRKSDGTLQEHDGICEIHVAKYAEIFGLTSAEASKDIRQALKSFAGKEVVFYRPEEDAGDEKGYESFPWFIKRAHSPSRGLYSVHINPYLIPFF**IGLQNRFTQFRLSETKE**ITNPYAMRLYESLCQYRKPDGSGIVSLKIDWIIERYQLPQSYQRMPDFRRRFLQVCVNEINSRTPMRLSYIEKKKGRQTTHIVFSFRDITSMTTG	
5c3iF	MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTV**ALREIRRYQKSTELLI**RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA	2.50
2kkwA	**MDVFMKGLSKAKEGVVAAAEKTKQGVAEAAGKTKEGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAVVTGVTAVAQKTVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEP**EA	1.04
4w50B	GPGNEVTLLDSRSVQGELGWIASPLEGGWEEVSIMDEKNTPIRTYQVCNVMEPSQNNWLRTDWITREGAQRVYIEIKFTLRDCNSLPGVMGTCKETFNLYYYESDNDKERFIRENQFVKIDTIAADE**SFTQVDIGDRIMKL**NTEIRDVGPLSKKGFYLAFQDVGACIALVSVRVFYKKA	

[4] Pdb ID+chain correspond to column 2 of Table S1

[5] Amino acid sequences of PDB chains from column 2 taken from their FASTA files. Bold represents identified fold-switching region, and yellow highlighting represents regions with independent folding cooperativity identified by SEED. Bold regions that are not highlighted were not recognized as independent cooperative folding units. Those not recognized in both conformations are highlighted in red.

[6] Qualifying ratios calculated by SEED (**Methods**), a measure of independent folding cooperativity. No score is reported for sequences in which SEED could not find a cooperative folding unit.

Table S3A—Expected fold switchers[1,2]

Viral fusion proteins p < 0.0009 (hypergeometric test)
4D6WA
1OK8A
4MMSA
2IEQA
2FYZB
4G2KA
1Y4MA
4GIPA
3DUZA
3FVCA
5H9CA
4O5NA
3KU3A
4JF3A
1G5GA
5L1XA
4H32A
5J81A
3VOPA
3WMIA
4XYPA
3N27A
1ZTMA
1WDGA
3G9RA
5G47A
5LJYH
2YQ2A
4NKJA
Pore proteins p <0.002 (hypergeometric test)
4PKMA
2CA5A
2YGTA
3ZJXA
4MKOA
5IMYA
3HVNA
1LKFA
2QK7A

4I0NA
1CBYA
4Q7GA
4K1PA
3EB7A
1PP0A
5DI0A
3KOGA
3X0TA
4OV8A
4PMKA
2C9KA
1S3RA
5GHEA
4IJYA
2NRJA
3ROHA
3TULA
3WA1A
Prion/amyoid proteins ($p < 3.2*10^{-5}$)
2MFZA
2KJ3A
4UNUA
2FMCA
2MZ7A
2H8NA
2LSHA
2N4OA
2KTMA
4HLSA
4O9LA
1G96A
“Structural Rearrangement” in Abstract (p < 0.016, hypergeometric test)
3DRNA
2D28C
5CB7A
4A37A
4CVNE
1CD3B
5HSQA

4NKJA
1S5PA
3TKAA
Other (no p-values quantified since PDB was not searched exhaustively)
1F3MA
2X9CA
3J9ED
2MAMA
1WWJA
2HUEB
4MI5A
3O26A
2C9KA
4A37A
2KXOA
2LE3A
5SUZA

[1] Proteins are categorized by type, e.g. viral fusion proteins, pores, etc.

[2] Black PDB IDS were identified by our method; red were not. Light gray/red were among the 16 experimentally-supported protein fold switchers. Light colors indicate that they were not double-counted.

Table S3B: Experimentally-supported protein fold switchers

PDB ID+Chain [1]	Justification	DOI(s)
2KXOA	"These results identify the MinD-dependent conformational changes in MinE that convert it from a latent to an active form and lead to a model of how MinE persists at the MinD-membrane surface."	10.1016/j.cell.2011.06.042 10.1073/pnas.1007141107
2LSHA	"We demonstrate that DewA populates two conformations in solution"	10.1016/j.jmb.2012.10.021
4OV8A	"The major conformational changes in PlyB are a ~70° opening of the bent and distorted central β-sheet of the MACPF domain, accompanied by extrusion and refolding of two α-	10.1371/journal.pbio.1002049

	helical regions into transmembrane β-hairpins (TMH1 and TMH2)."	
2MZ7A	"While Tau is highly flexible in solution and adopts a β-sheet structure in amyloid fibrils, in complex with microtubules the conserved hexapeptides at the beginning of the Tau repeats two and three convert into a hairpin conformation. Thus, binding to microtubules stabilizes a unique conformation in Tau."	10.1002/anie.201501714
4PMKA	"In solution, kissper is highly flexible and displays pore-forming activity in synthetic lipid-bilayers"	10.1016/j.jsb.2014.07.005
2N4OA	"We observe fast-timescale dynamics of two inter-cysteine segments in the protein that might foreshadow the conformational change that occurs when this protein self-assembles into amyloid-like fibrillar structures."	10.1038/srep25288
2KTMA	"We also prove that the isolated H2H3 is highly fibrillogenic and forms amyloid fibers morphologically similar to those obtained for the full-length protein."	10.1074/jbc.M110.111815
2LE3A	"Here, we show that the N-terminal regulatory domain (N) of CPT1A can adopt two complex amphiphilic structural states, termed Nα and Nβ, that interchange in a switch-like manner in	10.1074/jbc.M111.306951

	response to offered binding surface curvature.”	
2X9CA	“Here we show that the TTSS needle protomer refolds spontaneously to extend the needle from the distal end….We show that the protomer partially refolds from alpha-helix into beta-strand conformation to form the TTSS needle. Reconstitution experiments show that needle growth does not require ATP.”	10.1038/nsmb.1822
3J9ED	“Exposing BTV to low pH detaches VP2 and dramatically refolds the dagger and unfurling domains of VP5.” “This barb-like structure is long (~150 Å) and flexible and thus cannot be resolved in the three-dimensional (3D) density map at 9-Å resolution”	10.1038/nsmb.3134
5SUZA	““Here, we use structural analysis, as well as biophysical and cell-based assays, to show that the DEP domain of Dishevelled undergoes a conformational **switch**, from monomeric to swapped dimer, to trigger DIX-dependent polymerization and signaling to β-catenin”	10.1016/j.molcel.2016.08.026
4HLSA	“[Mutations] also alter the β-state-misfolding propensity of PrP; the serine mutations in hamster PrP decrease the propensity up to 35%, whereas the asparagine mutations in	10.1371/journal.pone.0063047

	rabbit PrP increase it up to 42%. Rapid dilution of rabbit and hamster into β-state buffer conditions causes quick conversion to β-state monomers.”	
1S5PA	“Moreover, the endothermic nature of the binding reaction suggests that cobB binding to the acetylated Acs protein target is dominated by an entropic contribution involving a burial of hydrophobic surface and/or structural rearrangement involving cobB, Acs or both.”	10.1016/j.jmb.2004.01.060
3TKAA	““It indicates that the complex is not in a catalytically active state, and structural rearrangement of RsmH or the nucleotides neighboring C1402 may be necessary to trigger catalysis. Although there is only one molecule in the asymmetric unit of the crystals, RsmH can form a compact dimer across a crystallographic twofold axis. Further analysis of RsmH by small-angle X-ray scattering (SAXS) also revealed the dimer in solution, but with a more flexible conformation than that in crystal, likely resulting from the absence of the substrate.”	10.1016/j.jsb.2012.04.011
3GAXA	“The crystal structure of human cystatin C, a protein with amyloidogenic properties and a potent inhibitor of cysteine	10.1038/86188 10.1111/j.1742-4658.2010.07596.x

	proteases, reveals how the protein refolds to produce very tight two-fold symmetric dimers while retaining the secondary structure of the monomeric form.” Note: 3GAXA was engineered to not switch folds. However, it is thought to represent the monomeric form of naturally-occurring cystatin C, which does switch folds. This is the only monomeric form of cystatin C currently available in the PDB.	

[1] Black PDB IDS were identified by our method; red were not.

Table S3C—Proteins expected to not switch folds.

Protein Family	PDBs not identified as fold switchers (true negatives)	PDBs+chains identified as fold switchers (false positives)
Green fluorescent proteins	2IB5, 5DPG, 4Z4K, 5DPH, 5DU0, 2HPW, 4Z4M, 5DTX, 2VZX, 5EJU, 2ZO6, 5DTZ, 4HVF, 5EHU, 2ZMU, 5DY6, 5DPI, 5DPJ	
Staphylococcal nucleases	4EQP, 5KEE, 5JOB	
Maltose binding proteins	2GH9, 2ELJ, 1EU8A	
WW domains	2JXW, 2YSE, 2DWV, 2KPZ, 1I5H, 2YSG, 2DMV, 2EZ5W, 2YSF, 2YSH, 2DK7, 1O6W, 2JV4, 2JX8, 1WMV, 2YSB, 3L4H, 1TK7, 2YSD, 4REX, 2L4JA,	2M8IA, 2MDIA, 1EG3A
Villin headpieces	1QZP	2K6MA
Trp-cage	2JOF	
BBA	1FME	
NTL9	2HBA	
BBL		2F60KA
Protein B	1Y71	
Homeodomains	2CXQ, 1Y66, 2YS9, 1W0T, 3MGQ, 2KMU, 2VI6, 2HI3,	1JGGA, 3NARA, 5JLWA, 1B72A, 3A02, 3A03, 1BW5, 2K40A, 2RSDA

	3NAU, 1WI3, 1K61A, 1MIJ, 2L7FP, 2M34, 2M7B	
Protein G	1GB1	
A3D	2A3D	
Lambda repressor	1LMB	
Ankyrin repeat proteins	2JABA, 5CECB, 1SW6A, 4DUIA, 4N5QA, 1IHBA, 1BD8A, 2DZNA, 5D66A, 4HI8A, 4HLLA, 3EU9A, 2L6BA, 5AARA, 2ZGDA, 1YMPA, 3LJNA, 1UOHA, 5BXOA, 4QFVA, 5EIDA, 3HRAA, 3C5RA, 3W9FA	1YCSB, 2RFMA, 401PA, 3B95A, 2F8YA, 2FO1E, 1DCQA, 2PNNA, 2ETBA, 2DZNA, 1OT8A, 2RFAA
Leucine rich repeat proteins	3VQ2C, 4OJUA, 4GT6A, 2RA8A, 4FMZA, 2P1MB, 2HR7A, 3BZ5A, 2LZ0A	4K5UA, 5HZLB, 3B2DA, 3WPFA, 3OGKB, 3T6QA, 3G06A, 1P9AG, 3WPCA, 3WN4A, 3WO9A, 4XOSA, 4XSQA, 3UN9A, 1H6UA, 5IL7A, 5GR9B, 4R58A, 5A5CA
Hemoglobins	3BOMA, 2GDMA, 3BCQA, 5AB8A, 2IG3A, 4HRRB, 1JF3A, 1WXRA, 1DLWA, 1WMUA, 1S69A, 1DLYA, 4NK2A, 2BK9A, 2OIFA, 4MAXA, 3AQ9A, 2VYWA, 2C0KA, 1UX8A, 2ZS0D, 1X3KA, 5D1VA, 1X46A, 2XYKA, 1TU9A, 4XDIA	3D1KB, 4H2LB, 1HBRA, 2BMMA,3MKBB, 1YHUC, 2AA1B, 1GCVB, 3AT5A,2QRWA, 3BJ1A
Ubiquitin	1WE7A	
Glutathione S-transferase	5EURA, 4HI7A, 4PNGA, 1R5AA, 4GCIA, 1V2AA, 1YQ1A, 4IVFA, 3RBTA, 3BBYA, 1JLVA, 1OYJA, 5ECKB, 5HFKA, 5ELGA, 4L8EA, 4MP4A, 3IBHA, 1GWCA, 3H1NA, 4PXOA, 4NAXA, 4MPGA, 1K3YA, 3LXZA, 3EINA, 4MK3A, 4KH7A, 4MF7A, 4NHWA, 4MF5A,	4HOJA, 4ID0A, 4KGIA, 4IELA, 4KF9A, 4KDXA, 1B48A, 1F3AA, 1GSUA
Chymotrypsin inhibitor	4TLPA, 1JXCA	2M99A

SI References

1. Kabsch W & Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22(12):2577-2637.
2. Fitzkee NC, Fleming PJ, & Rose GD (2005) The Protein Coil Library: a structural database of nonhelix, nonstrand fragments derived from the PDB. *Proteins* 58(4):852-854.
3. Kortemme T, Morozov AV, & Baker D (2003) An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein-protein complexes. *J Mol Biol* 326(4):1239-1259.
4. Wang G & Dunbrack RL, Jr. (2003) PISCES: a protein sequence culling server. *Bioinformatics* 19(12):1589-1591.
5. Porter LL & Rose GD (2012) A thermodynamic definition of protein domains. *Proc Natl Acad Sci U S A* 109(24):9420-9425.
6. Yang Y, *et al.* (2017) SPIDER2: A Package to Predict Secondary Structure, Accessible Surface Area, and Main-Chain Torsional Angles by Deep Neural Networks. *Methods Mol Biol* 1484:55-63.
7. Koehnke J, *et al.* (2010) Splice form dependence of beta-neurexin/neuroligin binding interactions. *Neuron* 67(1):61-74.

8. Schellenberg MJ, *et al.* (2010) Context-dependent remodeling of structure in two large protein fragments. *J Mol Biol* 402(4):720-730.

9. Alexander PA, He Y, Chen Y, Orban J, & Bryan PN (2009) A minimal sequence code for switching protein structure and function. *Proc Natl Acad Sci U S A* 106(50):21149-21154.

10. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

11. Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* 27(10):527-533.