STRUCTURAL BIOLOGY
COMMUNICATIONS

Volume 71 (2015)

Supporting information for article:

The role of flexibility and molecular shape in the crystallization of proteins by surface mutagenesis

Yancho D. Devedjiev

Table S1 Occurrence of Static Disorder (SD) and high Side Chain Entropy (SCE) in a set of 105 protein structures refined to high resolution (d >1.5 Å).

PDB ID	Surface exposed Residues No	SD surface exposed No / %	SD in contact Area %	SCE surface exposed No / %	SCE in contact area %	SD residues on flexible loops %
Polypeptide chain length 0 – 200 residues						
1GCI	187	21 / 11	43	85 / 45	29	43
1AB1	44	7 / 16	57	9 / 20	33	29
1G6X	55	12 / 22	83	17 / 31	65	25
1HJE	13	6 / 46	100	3 / 23	100	17
1M1Q	80	23 / 29	61	28 / 35	46	52
1UCS	56	5 / 9	89	22 / 39	54	20
2JFR	188	32 / 17	37	53 / 28	34	25
2OV0	88	8 / 9	62	28 / 32	79	0
2VB1	95	42 / 44	78	44 / 46	64	36
2VXN	189	22 / 12	59	73 / 36	26	14
2WFI	129	24 / 19	54	54 / 42	52	8
2WUR	190	22 / 12	54	69 / 36	41	27
2X46	128	7 / 5	71	43 / 34	61	43

2XOD	90	12 / 13	58	43 / 48	58	13
2XOM	114	21 / 18	67	52 / 45	58	9
2XU3	165	17 / 10	53	66 / 40	36	47
2Y78	103	16 / 15	50	33 / 32	30	19
2ZQ7	186	38 / 20	53	76 / 41	50	37
3A38	75	15 / 20	73	23 / 31	61	53
3AGN	104	8 / 8	37	30 / 29	63	62
3DVZ	61	13 / 21	54	30 / 43	67	15
3E7R	35	1 / 3	100	11 / 31	100	100
3FYM	70	5 / 7	60	37 / 51	73	0
3G21	71	17 / 24	59	25 / 35	76	59
3H31	67	17 / 25	88	23 / 34	87	65
3JUD	121	9 / 7	55	41 / 34	66	11
3KFF	123	22 / 18	54	51 / 42	44	4
3KLR	116	27 / 23	44	48 / 41	40	22
3LL1	107	23 / 22	48	40 / 38	65	26
3NED	178	33 / 18	33	75 / 42	39	24
3PUC	91	12 / 39	75	32 / 35	28	25
3QL9	107	1 / 1	100	46 / 43	59	0
3QPA	143	14 / 10	50	46 / 32	43	14

3SOJ	85	17 / 20	70	38 / 45	66	18
3W5H	211	38 / 18	47	79 / 37	38	21
4A02	132	2 / 1	100	45 / 34	42	0
4J4Z	126	9 / 7	78	46 / 36	50	11
4JLL	126	5 / 4	80	49 / 39	47	0
2GKG	96	22 / 23	64	35 / 36	60	14
4HS1	78	18 / 2	84	20 / 25	74	10
2FMA	54	2 / 4	100	17 / 31	94	0
1BZR	131	16 / 12	69	48 / 37	67	6
1CTJ	79	12 / 15	92	20 / 25	80	17
1KQ6	120	29 / 24	72	40 / 33	62	7
1MFM	123	14 /13	65	47 / 38	66	0
1PQ7	172	13 / 8	61	62 /36	64	8
1R6J	75	21 / 28	95	20 / 26	95	14
1YK4	50	14 / 28	86	15 / 27	86	21
2G58	109	15 / 14	47	37 / 34	73	13
2GKG	95	22 / 23	64	35 / 37	54	14
2O9S	60	7 / 12	86	22 / 37	95	14
2OFR	148	45 / 30	44	51 / 34	63	15
2PPN	93	15 / 16	67	31 / 33	77	13

2QCP	73	19 / 26	68	27 / 37	78	16
3O5P	113	18 / 16	61	41 / 37	58	22
3RCG	127	18 / 14	72	41 / 51	32	11
3UI6	83	13 / 16	92	32 / 38	69	8
3V1A	48	10 / 21	100	20 / 42	85	0
3VHG	162	13 / 8	69	49 / 30	51	8
3VOR	144	35 / 24	83	51 / 35	80	26
3WDN	108	28 / 26	57	49 / 45	43	4
3ZR8	59	11 / 17	91	20 / 34	100	18
3ZSJ	118	20 / 17	60	46 / 39	61	10
3ZZP	70	7 / 10	100	31 / 44	97	0
4ACJ	134	15 / 11	67	45 / 31	51	27
4AQO	80	8 / 10	75	29 / 38	86	0
4BCT	154	13 / 8	46	47 / 32	51	16
4BVM	113	31 / 27	46	52 / 46	52	0
4DP9	86	24 / 28	92	31 / 36	77	4
4DPB	87	25 / 29	80	32 / 37	84	12
4DRQ	110	12 / 11	58	42 / 39	69	8
4EA7	153	6 / 4	33	36 / 23	50	0
4EIC	79	6 / 8	83	32 / 41	47	0

4FPT	191	22 / 11	41	73 / 38	49	9
4FRC	199	18 / 9	39	79 / 40	44	22
4GA2	128	36 / 28	61	55 / 43	73	0
4HS1	77	20 / 30	80	19 / 25	89	10
4IAU	131	15 / 11	60	46 / 35	61	13
4KQP	183	32 / 17	59	70 / 39	54	22
4MZC	82	21 / 26	57	41 / 50	54	9
3E6Z	70	70 / 19	85	27 / 39	85	31
2IIM	58	2 /3	100	22 / 32	57	50
2FLA	75	7 / 9	57	24 / 31	83	0
2I4A	88	3 / 3	67	32 / 36	59	0
Average (%)		16	68	36	62	18
Polypeptide chain length 201 – 300 residues						
4HNO	210	19 / 9	63	93 / 42	53	5
4GNR	260	43 / 17	44	99 / 38	51	37
4F18	287	8 1/ 28	41	82 / 32	40	37
4F1U	286	90 / 31	45	83 / 29	42	45

4AWS	248	35 / 14	37	79 / 26	35	20
3U7C	202	23 / 11	39	72 / 35	60	0
3P8J	261	21 / 8	43	99 / 38	27	24
2FVY	231	22 / 9	41	90 / 39	35	0
4N1I	225	27 / 12	52	78 / 35	40	7
3O4P	233	39 / 18	54	93 / 40	34	18
3JYO	218	19 / 9	53	68 / 31	32	16
2Y61	202	20 / 10	40	72 / 36	35	25
2XJP	205	52 / 25	19	66 / 32	21	36
2QIK	224	9 / 4	22	82 / 36	40	4
1US0	221	68 / 30	34	106 / 47	41	47
4AWT	242	40 / 16	37	85 / 35	25	15
Average (%)		16	42	36	38	21
Polypeptide chain length 301 residues and up						
4NOV	236	13 / 5	31	46 / 19	26	38
4PMX	221	2 / 1	100	70 / 32	47	50
4WIH	289	8 / 3	0	99 / 34	33	5

4RJZ	305	20 / 7	55	81 / 27	44	12
4PVI	252	8 / 3	37	64 / 25	41	37
Average (%)		4	45	27	38	28
Grand total %		16	62	36	57	19

Table S2 Comparison of subset of SCE residues with higher than expected occurrence in crystal contacts (Price et al., 2009) with the group of small residues with high occurrence at contact area (Sieslik & Derewenda, 2009) in the set of 105 structures refined at high resolution (d >1.5Å).

	SCE* Q, R			Small G, A, S, T	
PDB ID	Surface exposed residues No	SCE exposed all %	SCE in contact area %	Small exposed %	Small in contact area %
Polypeptide chain length 0 – 200					
1GCI	187	10	55	40	27
1AB1	44	4	50	41	67
1G6X	55	13	86	24	61
1HJE	13	0	0	23	100
1M1Q	80	6	60	34	52
1UCS	56	5	33	23	61
2JFR	188	12	45	33	31
2OV0	88	3	100	32	43
2VB1	95	12	93	27	57
2VXN	189	12	35	29	37

2WFI	129	13	41	25	34
2WUR	190	7	54	21	22
2X46	128	2	100	40	43
2XOD	90	8	71	19	35
2XOM	114	8	67	29	42
2XU3	165	7	58	25	22
2Y78	108	8	33	34	43
2ZQ7	186	14	57	28	30
3A38	75	8	75	33	80
3AGN	104	8	75	33	10
3E7R	35	3	100	23	75
3FYM	70	11	91	19	46
3G21	71	15	72	20	45
3H31	67	10	86	28	68
3JUD	121	16	70	25	47
3KFF	121	5	33	19	50
3KLR	116	3	0	26	46
3LL1	106	5	60	47	68
3NED	178	6	54	23	37
3PUC	91	3	67	41	65

3QL9	107	12	77	19	45
3QPA	143	11	69	32	26
3UI4	85	9	50	25	57
4A02	132	10	31	29	39
4J4Z	126	10	85	32	27
4JLL	126	12	67	31	38
2GKG	96	10	64	36	60
4HS1	78	8	86	30	67
2FMA	54	5	100	13	71
1BZR	131	7	79	18	50
1CTJ	79	6	60	29	56
1KQ6	120	13	69	17	86
1MFM	123	6	50	32	56
1PQ7	172	8	46	51	47
1R6J	75	5	100	36	74
1YK4	50	2	100	22	73
2G58	109	5	100	19	53
2GKG	95	13	67	28	41
2O9S	61	15	100	21	61
2OFR	148	3	40	31	58

2PPN	92	12	82	27	44
2QCP	74	9	86	27	70
3O5P	113	4	60	26	57
3RCG	127	6	37	35	27
3UI6	83	10	75	28	48
3V1A	48	19	100	23	82
3VHG	162	9	73	34	36
3VOR	144	10	87	42	88
3WDN	108	3	33	27	38
3ZR8	59	8	100	29	100
3ZSJ	118	10	83	19	48
3ZZP	70	19	100	14	90
4ACJ	134	6	75	27	44
4AQO	89	2	100	33	86
4BCT	154	11	59	31	35
4BVM	126	8	50	26	45
4DP9	86	1	100	46	92
4DRQ	110	4	80	28	58
4EA7	153	8	61	33	43
4EIC	79	6	60	42	57

4FPT	191	8	75	25	42
4FRC	199	8	71	24	50
4GA2	128	12	80	16	48
4GNR	260	5	46	25	29
4HNO	210	9	58	21	41
4HS1	77	9	86	34	73
4IAU	131	12	81	35	56
4KQP	183	6	54	30	47
4MZC	82	4	33	18	53
3E6Z	70	11	100	27	47
2IIM	58	10	67	29	65
2FLA	75	8	67	36	78
2I4A	88	6	60	31	59
Average (%)		8	68	29	53
Polypeptide chain length 201 – 300					
1US0	221	11	42	18	39
2QIK	224	10	45	20	24
2XJP	205	5	27	36	19
2Y61	202	11	36	28	32

3JYO	217	11	58	31	41
3O4P	233	10	30	19	27
3W5H	211	9	50	17	24
4AWT	242	16	31	28	27
3P8J	261	12	34	31	16
3U7C	202	8	71	23	40
4AWS	248	11	39	29	26
4N1I	219	10	41	39	13
2FVY	231	8	53	21	31
4F1U	286	7	32	47	35
4F18	287	6	28	47	37
4GNR	260	5	46	25	29
4HNO	210	9	58	21	41
Average (%)		9	42	28	29
Polypeptide chain length 301 residues and up					
4NOV	236	4	55	24	16
4P\MX	221	17	76	30	36
4WIH	289	11	29	20	25
4RJZ	305	7	52	24	31

4PVI	252	10	54	23	30
Average (%)		10	53	24	28
Grand total %		9	63	28	48

* Side Chain Entropy

Table S3 Crystal forms of RhoGDI inhibitor show non-uniform changes in the shape of molecules in the asymmetric unit modulated by the solvent content and the crystallographic environment (for details see the text).

PDB ID	Mutations	Crystallization conditions	Space group	Chain identifier Accessible area Å^2	Rmsd */ Maximum displacement Å	Rmsd **/ Maximum displacement Å	Reference
1RHO	none	2 M ammonium sulfate, 0.1 M Tris, pH 8.5	H32	A 7691 B 7699 C 7642	0.7 / 6.2		Keep et al., 1997
1KMT	E154A E155A	PEG 4000, Tris, pH 8.5, lithium sulfate, methylpentane diol	P1	A 7637 B 7728	0.8 / 6.1	1.0 / 9.3 1.1 / 9.4	Mateja et al., 2002
2JHV	E154A E155A	2.0 M ammonium sulfate, 0.2 M sodium chloride	C2	A 7680 B 7700 C 7676	0.6 / 5.0	1.1 / 5.6 1.1 / 5.8 1.1 / 6.4	Cooper et al., 2007

		0.1 M sodium cacodylate, pH 6.5		D 7702 E 7793 F 7815		1.0 / 6.5 1.1 / 6.7 1/1 / 6.0	
2JHW	E155A E157A	2M ammonium sulfate	H32	A 7680 B 7619	0.9 / 7.4	1.1 / 5.6 1.1 / 5.8	Cooper et al., 2007
1FSO	K135A K138A K141A F196I	Ammonium sulfate MES buffer, pH 6.0,	$R3_221$	A 7504		1.1 / 7.4	1FSO
1FST	K135A K138A K141A	PEG 3400, isopropanol, HEPES buffer, pH 7.0	H3	A 7729 B 7767	0.6 / 6.3	1.0 / 6.9 1.0 / 7.2	Longenecker et al., 2001
1FT0	K113A	Ammonium sulfate, NA/K Tartrate, Sodium citrate, pH 5.6	H32	A 7206 B 7231	0.9 / 5.7	1.1 / 8.1 0.9 / 9.0	Longenecker et al., 2001
2BXW	K135Y K138Y K141Y	4M sodium formate	$P3_121$	A 7710 B 7619	0.7 / 7.2	1.1 / 6.6	Cooper et al., 2007

Minimum value				7206	0.6 / 5.0	0.9 / 5.0	
Maximum value				7815	1.3 / 9.0	1.3 / 9.4	
Difference SASA (%)				8			

* rmsd and maximum displacement were calculated for all atoms between corresponding chains encompassing residues 67 to 202 in the asymmetric unit of the crystals

** rmsd and maximum displacement calculated by superimposition of the particular chain with molecule A from 1RHO

Table S4 SASA of the wt and surface mutated residues in RhoGDI (for details see the text).

PDB ID	Mutation	Molecule A	Molecule B	Molecule C	Molecule D	Molecule E	Molecule F
1RHO	Wt						
	K 113	40*	42	38			
	K 135	187	173	180			
	K 138	90	82	85			
	K 141	143	142	143			
	K 199	107	112	96			
	K 200	94	93	121			
	E 154	125	127	116			
	E 155	58	64	59			
	E 157	107	112	101			
1KMT	E → A						
	A 154	53	54				
	A 155	35	39				
2JHV	E → A						
	A 154	52	50	51	57	56	58
	A 155	27	31	28	30	32	37

2JHW	E → A A 155 A 157	 34 62	 37 54				
1FSO	K → A A 135 A 138 A 141	 101 73 60					
1FST	K → A A 135 A 138 A 141	 97 72 73	 103 61 67				
1FT0	K → A A 113	 0	 0				
2JHS	K → H H 135 H 138 H 141	 150 95 129					
2JHX	E → H H 155	 75	 67				

	H 157	92	98				
2JHZ	K → S						
	S 155	52	42				
	S 157	62	68				
2BXW	K → Y						
	Y 135	190	164				
	Y 138	114	110				
	Y 141	107	105				
2JI0	K → Y						
	Y 138	105					
	Y 141	99					
2JHT	Lys→T						
	T 135	110	136	132	114		
	T 138	81	97	76	77		
	T 141	71	76	76	73		
1QVY	Lys→R						
	Arg 199	129	123	116	117		
	Arg 200	168	167	174	172		

SASA is presented in $Å^2$

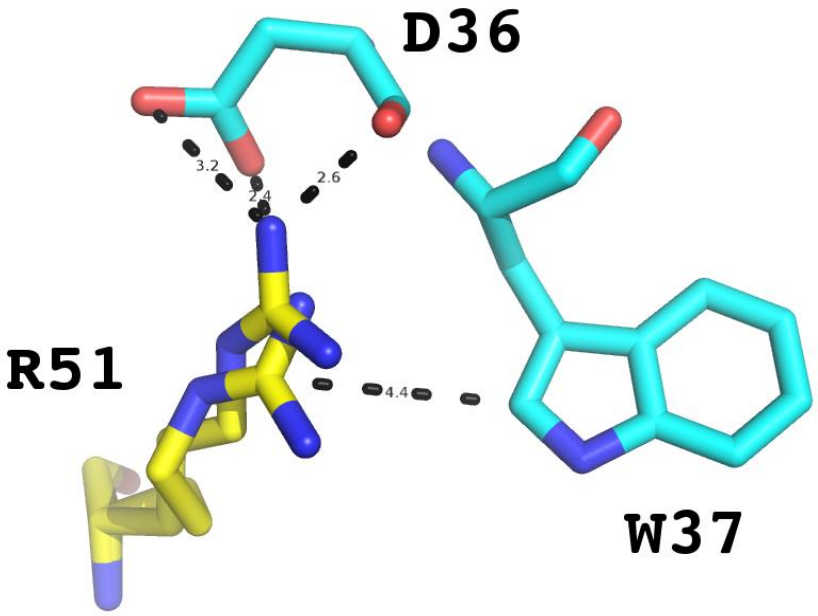

Figure S1 In rubredoxin (PDB entry 1YK4) R51 displays alternative conformations. One is stabilized by a salt bridge with side chain of symmetry-related D36^{sym} and H-bond between NH1 atom of R51 and the main chain oxygen from the symmetry-related D36. The second alternative conformation is stabilized by van der Waals contacts with W37^{sym}.

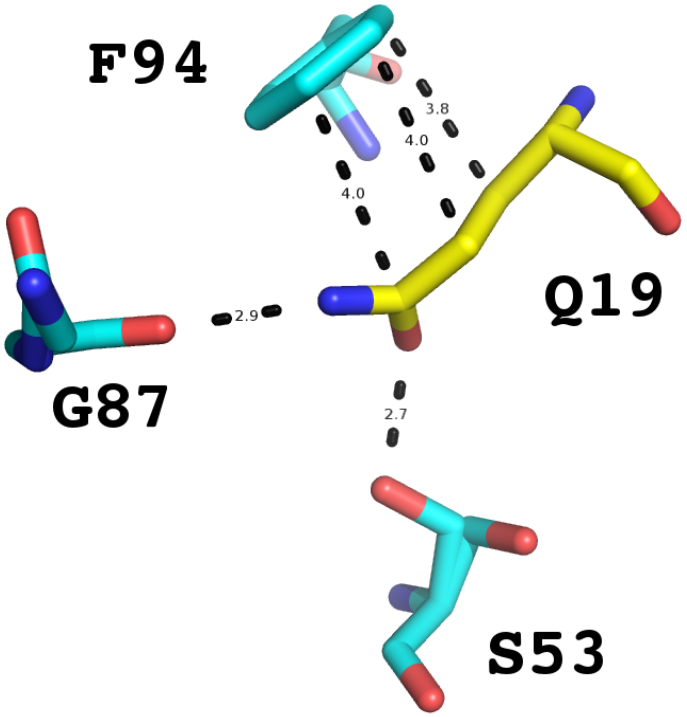

Figure S2 In PIASE domain of human cyclophilin (PDB entry 2WFI) Q19 (yellow bonds) is involved in contact with symmetrically related molecule (cyan bonds) via a set of intermolecular interactions. NE2 atom forma a H-bonds with main chain oxygen from $G87^{sym}$, OE1 is H-bonded to OG atom of the symmetrically-related $S53^{sym}$, while all side chain atoms of Q19 are positioned at Van der Waals distance to the rim of benzyl ring of symmetrically-related $F94^{sym}$.