Full wwPDB X-ray Structure Validation Report (i)

Sep 22, 2020 – 06:05 PM BST

|		
PDB ID	:	6STJ
Title	:	Selective Affimers Recognize BCL-2 Family Proteins Through Non-Canonical Structural Motifs
Authors	:	Hobor, F.; Miles, J.A.; Trinh, C.H.; Taylor, J.; Tiede, C.; Rowell, P.R.; Jackson, B.; Nadat, F.; Kyle, H.F.; Wicky, B.I.M.; Clarke, J.; Tomlinson, D.C.; Wilson, A.J.; Edwards, T.A.
Deposited on	:	2019-09-10
Resolution	:	2.20 Å(reported)

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
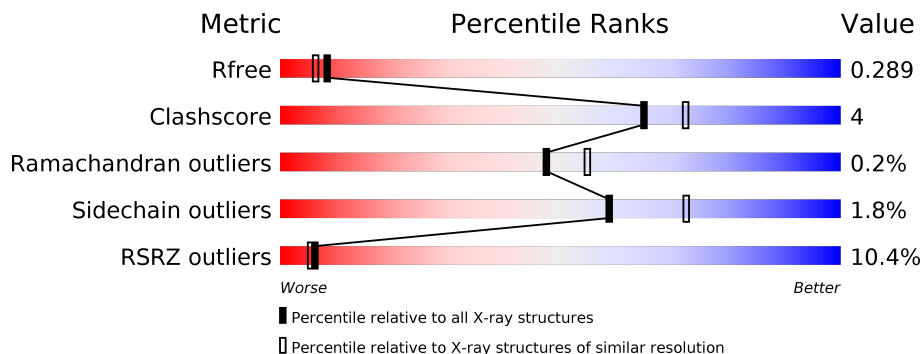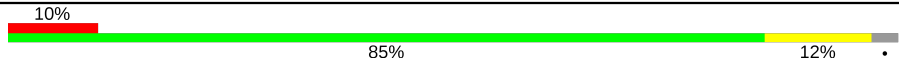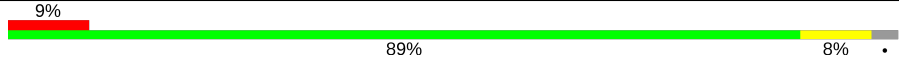

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	156	% 80% 10% 10%
1	B	156	% 83% 7% · 9%
1	C	156	21% 53% 13% 34%
1	D	156	19% 67% 19% · 14%
2	E	91	3% 88% 10% ·
2	F	91	4% 95% · ·

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	91	10% 85% 12%
2	H	91	9% 89% 8%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7308 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Induced myeloid leukemia cell differentiation protein Mcl-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	141	Total	C	N	O	S	0	0	0
			1142	720	209	210	3			
1	B	142	Total	C	N	O	S	0	0	0
			1151	726	211	211	3			
1	C	103	Total	C	N	O	S	0	0	0
			833	529	152	150	2			
1	D	134	Total	C	N	O	S	0	0	0
			1098	695	202	198	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	172	SER	-	expression tag	UNP Q07820
B	172	SER	-	expression tag	UNP Q07820
C	172	SER	-	expression tag	UNP Q07820
D	172	SER	-	expression tag	UNP Q07820

- Molecule 2 is a protein called Cystatin domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	89	Total	C	N	O	S	0	0	0
			760	497	124	136	3			
2	F	90	Total	C	N	O	S	0	0	0
			769	502	125	139	3			
2	G	88	Total	C	N	O	S	0	0	0
			752	493	122	134	3			
2	H	88	Total	C	N	O	S	0	0	0
			752	493	122	134	3			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	18	Total 18	O 18	0	0
3	B	11	Total 11	O 11	0	0
3	D	2	Total 2	O 2	0	0
3	E	6	Total 6	O 6	0	0
3	F	5	Total 5	O 5	0	0
3	G	4	Total 4	O 4	0	0
3	H	5	Total 5	O 5	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Induced myeloid leukemia cell differentiation protein Mcl-1
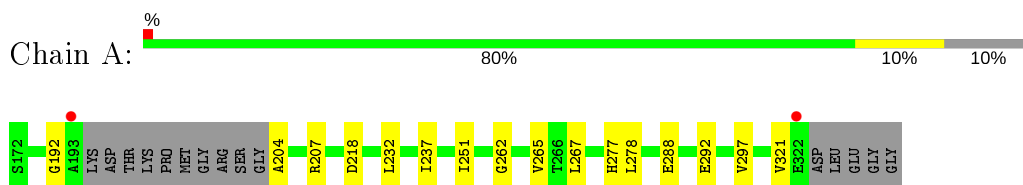
• Molecule 1: Induced myeloid leukemia cell differentiation protein Mcl-1
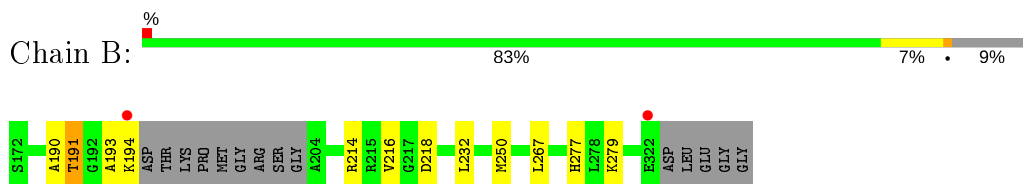
• Molecule 1: Induced myeloid leukemia cell differentiation protein Mcl-1
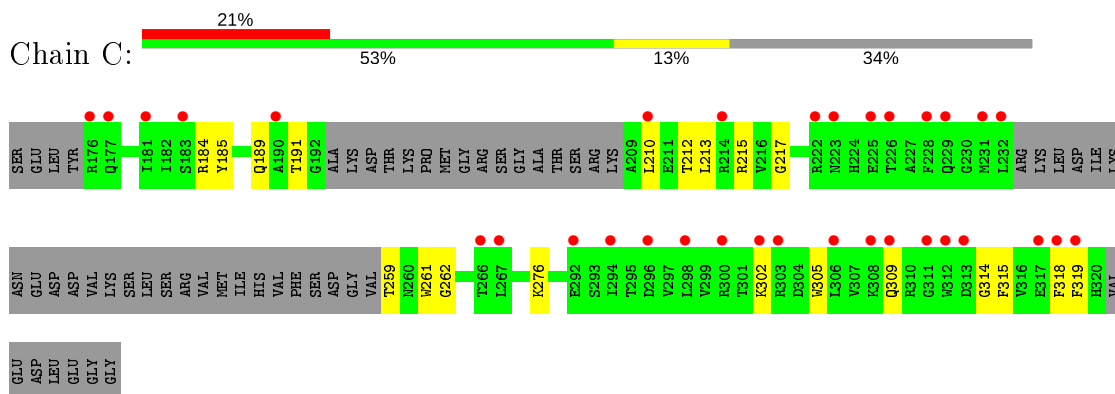
• Molecule 1: Induced myeloid leukemia cell differentiation protein Mcl-1

• Molecule 2: Cystatin domain-containing protein
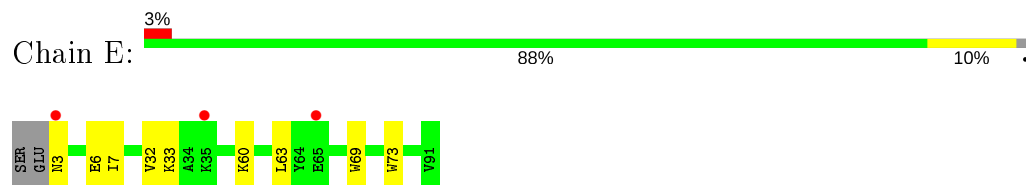
Chain E:

• Molecule 2: Cystatin domain-containing protein

Chain F:

• Molecule 2: Cystatin domain-containing protein

Chain G:

• Molecule 2: Cystatin domain-containing protein

Chain H:
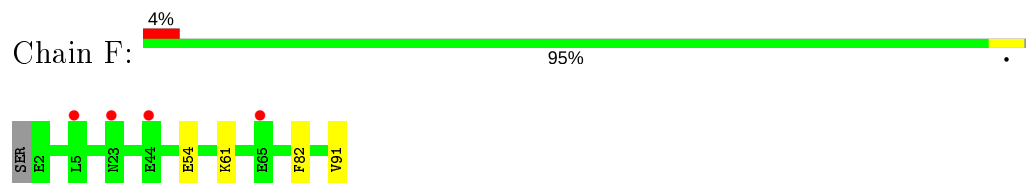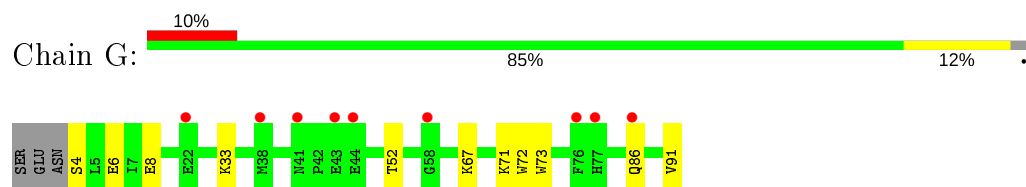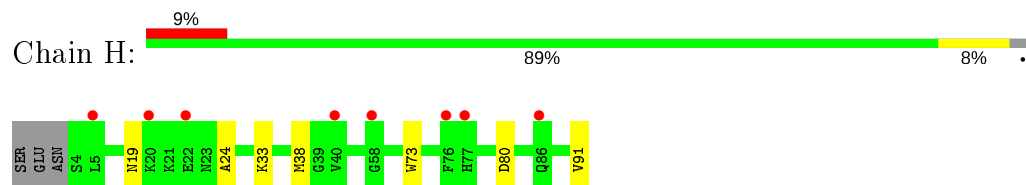
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	92.14Å 107.50Å 226.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	113.08 – 2.20 113.08 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.3 (113.08-2.20) 99.3 (113.08-2.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.79 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0253	Depositor
R, R_{free}	0.257 , 0.289 0.257 , 0.289	Depositor DCC
R_{free} test set	2850 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	46.1	Xtriage
Anisotropy	0.316	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 46.1	EDS
L-test for twinning [2]	$<|L|> = 0.50$, $<L^2> = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7308	wwPDB-VP
Average B, all atoms (Å^2)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 34.91 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 6.3396e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.66	0/1160	0.76	0/1560
1	B	0.66	0/1169	0.73	0/1571
1	C	0.67	0/847	0.71	0/1140
1	D	0.67	0/1115	0.74	0/1497
2	E	0.61	0/781	0.76	0/1050
2	F	0.62	0/790	0.73	0/1062
2	G	0.61	0/773	0.74	0/1039
2	H	0.61	0/773	0.74	0/1039
All	All	0.65	0/7408	0.74	0/9958

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1142	0	1149	11	0
1	B	1151	0	1162	9	0
1	C	833	0	827	12	0
1	D	1098	0	1111	18	0
2	E	760	0	745	6	0
2	F	769	0	751	2	0
2	G	752	0	739	6	1
2	H	752	0	739	4	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	18	0	0	0	0
3	B	11	0	0	0	0
3	D	2	0	0	0	0
3	E	6	0	0	0	0
3	F	5	0	0	0	0
3	G	4	0	0	0	0
3	H	5	0	0	0	0
All	All	7308	0	7223	59	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

All (59) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:215:ARG:NH2	1:C:319:PHE:O	2.16	0.79
1:B:190:ALA:O	1:B:279:LYS:HD3	1.91	0.70
1:C:259:THR:HG23	1:C:302:LYS:HE3	1.83	0.60
1:D:254:PHE:CE1	1:D:267:LEU:HD22	2.38	0.58
1:A:321:VAL:HG22	2:E:69:TRP:CD2	2.40	0.57
2:G:4:SER:O	2:G:8:GLU:HG2	2.05	0.57
1:D:251:ILE:HD11	1:D:300:ARG:HD2	1.89	0.55
1:B:232:LEU:HD21	1:B:277:HIS:CG	2.42	0.55
1:D:262:GLY:HA3	2:H:73:TRP:CD1	2.41	0.54
1:C:262:GLY:HA3	2:G:73:TRP:CD1	2.42	0.54
2:G:6:GLU:OE2	2:G:72:TRP:HZ2	1.91	0.54
1:C:259:THR:HG22	1:C:305:TRP:CE2	2.43	0.54
1:D:174:LEU:HD21	1:D:306:LEU:HD13	1.90	0.53
1:A:237:ILE:HG22	1:A:237:ILE:O	2.09	0.52
1:A:321:VAL:HG22	2:E:69:TRP:CE3	2.47	0.50
1:D:235:LEU:HD13	1:D:246:LEU:CD1	2.41	0.49
1:A:267:LEU:HD23	1:A:267:LEU:C	2.32	0.49
1:D:240:GLU:HG3	1:D:286:CYS:SG	2.53	0.48
1:D:215:ARG:NH2	1:D:319:PHE:O	2.41	0.47
1:A:232:LEU:HD21	1:A:277:HIS:CG	2.50	0.47
1:D:261:TRP:CZ3	1:D:315:PHE:HB2	2.50	0.46
2:H:19:ASN:HB2	2:H:24:ALA:O	2.15	0.46
1:B:250:MET:HE1	1:B:267:LEU:HD23	1.96	0.46
1:B:191:THR:HG22	1:B:193:ALA:H	1.80	0.46
1:C:189:GLN:OE1	1:C:276:LYS:NZ	2.49	0.46
2:F:54:GLU:OE1	2:F:61:LYS:HB3	2.16	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:G:67:LYS:HD2	2:G:86:GLN:HB2	1.98	0.45
1:D:299:VAL:O	1:D:303:ARG:HB2	2.17	0.45
1:B:250:MET:CE	1:B:267:LEU:HD23	2.47	0.45
2:G:33:LYS:HG2	2:G:52:THR:HB	1.98	0.45
1:C:213:LEU:HD23	1:C:315:PHE:HZ	1.81	0.45
1:D:251:ILE:HD12	1:D:297:VAL:HG22	1.98	0.45
1:B:216:VAL:HA	2:F:82:PHE:CE1	2.52	0.45
1:C:184:ARG:HB3	1:C:210:LEU:HD21	1.98	0.45
1:C:185:TYR:OH	1:C:217:GLY:HA3	2.18	0.44
1:A:262:GLY:HA3	2:E:73:TRP:CD1	2.53	0.44
2:H:19:ASN:HB2	2:H:24:ALA:C	2.38	0.44
2:E:3:ASN:O	2:E:7:ILE:HG12	2.19	0.43
1:D:258:VAL:HB	1:D:263:ARG:NH1	2.34	0.43
1:A:288:GLU:O	1:A:292:GLU:HG3	2.19	0.43
1:C:318:PHE:O	2:G:71:LYS:HD3	2.19	0.43
1:B:194:LYS:HE3	1:B:214:ARG:HH12	1.83	0.42
2:E:32:VAL:HG21	2:E:63:LEU:HD21	2.00	0.42
1:D:185:TYR:OH	1:D:217:GLY:HA3	2.19	0.42
1:B:250:MET:HE1	1:B:267:LEU:CD2	2.49	0.42
1:A:321:VAL:HG13	2:E:69:TRP:CD1	2.55	0.42
1:C:212:THR:CG2	1:C:315:PHE:CE2	3.03	0.42
1:B:194:LYS:HE3	1:B:214:ARG:NH1	2.35	0.42
1:D:262:GLY:HA3	2:H:73:TRP:CG	2.55	0.42
1:C:309:GLN:O	1:C:314:GLY:HA3	2.20	0.41
1:D:216:VAL:O	1:D:220:VAL:HG23	2.20	0.41
1:A:204:ALA:O	1:A:207:ARG:HB3	2.20	0.41
1:C:261:TRP:HE1	1:C:309:GLN:HB3	1.86	0.41
1:D:254:PHE:HD2	1:D:302:LYS:HE2	1.86	0.41
1:D:262:GLY:O	1:D:265:VAL:HG12	2.20	0.41
1:A:251:ILE:HG21	1:A:297:VAL:CG1	2.51	0.41
1:D:312:TRP:O	1:D:316:VAL:HG23	2.20	0.40
1:A:237:ILE:HG22	1:A:278:LEU:HD21	2.04	0.40
1:D:251:ILE:HD13	1:D:300:ARG:NH1	2.36	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:G:91:VAL:OXT	2:H:91:VAL:O[4_555]	2.17	0.03

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	137/156 (88%)	132 (96%)	4 (3%)	1 (1%)	22	22
1	B	138/156 (88%)	133 (96%)	5 (4%)	0	100	100
1	C	97/156 (62%)	93 (96%)	3 (3%)	1 (1%)	15	14
1	D	128/156 (82%)	120 (94%)	8 (6%)	0	100	100
2	E	87/91 (96%)	85 (98%)	2 (2%)	0	100	100
2	F	88/91 (97%)	87 (99%)	1 (1%)	0	100	100
2	G	86/91 (94%)	82 (95%)	4 (5%)	0	100	100
2	H	86/91 (94%)	84 (98%)	2 (2%)	0	100	100
All	All	847/988 (86%)	816 (96%)	29 (3%)	2 (0%)	47	55

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	191	THR
1	A	192	GLY

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	124/135 (92%)	122 (98%)	2 (2%)	62	76
1	B	125/135 (93%)	123 (98%)	2 (2%)	62	76
1	C	88/135 (65%)	88 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	120/135 (89%)	117 (98%)	3 (2%)	47	60
2	E	79/81 (98%)	76 (96%)	3 (4%)	33	42
2	F	80/81 (99%)	79 (99%)	1 (1%)	69	81
2	G	78/81 (96%)	78 (100%)	0	100	100
2	H	78/81 (96%)	75 (96%)	3 (4%)	33	42
All	All	772/864 (89%)	758 (98%)	14 (2%)	59	72

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	218	ASP
1	A	265	VAL
1	B	191	THR
1	B	218	ASP
1	D	249	VAL
1	D	250	MET
1	D	251	ILE
2	E	6	GLU
2	E	33	LYS
2	E	60	LYS
2	F	91	VAL
2	H	33	LYS
2	H	38	MET
2	H	80	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	282	ASN
1	B	177	GLN
2	F	23	ASN
2	H	86	GLN

5.3.3   RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	141/156 (90%)	0.42	2 (1%)	75	73	34, 49, 68, 82	0
1	B	142/156 (91%)	0.36	2 (1%)	75	73	32, 49, 68, 91	0
1	C	103/156 (66%)	1.62	33 (32%)	0	0	63, 88, 130, 148	0
1	D	134/156 (85%)	1.39	30 (22%)	0	0	50, 80, 112, 128	0
2	E	89/91 (97%)	0.40	3 (3%)	45	43	41, 56, 75, 95	0
2	F	90/91 (98%)	0.48	4 (4%)	34	32	41, 58, 85, 102	0
2	G	88/91 (96%)	0.71	9 (10%)	6	6	45, 68, 99, 114	0
2	H	88/91 (96%)	0.74	8 (9%)	9	8	45, 64, 96, 112	0
All	All	875/988 (88%)	0.77	91 (10%)	6	5	32, 62, 107, 148	0

All (91) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	251	ILE	13.7
1	D	252	HIS	6.7
1	C	312	TRP	6.7
1	D	174	LEU	6.0
1	C	298	LEU	5.9
1	C	306	LEU	5.9
1	D	173	GLU	5.8
1	C	300	ARG	5.5
1	C	210	LEU	5.4
1	D	210	LEU	5.0
2	H	76	PHE	4.8
1	C	177	GLN	4.7
1	D	209	ALA	4.7
1	D	208	LYS	4.6
1	D	253	VAL	4.4
1	D	236	ASP	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	308	LYS	4.3
1	C	181	ILE	4.2
1	A	193	ALA	4.2
1	C	266	THR	4.0
1	C	317	GLU	3.9
1	D	207	ARG	3.9
2	G	58	GLY	3.9
1	C	226	THR	3.7
1	C	176	ARG	3.7
2	H	58	GLY	3.6
2	G	76	PHE	3.6
1	D	238	LYS	3.6
1	D	235	LEU	3.5
1	D	230	GLY	3.5
1	D	307	VAL	3.4
1	D	205	THR	3.4
1	C	228	PHE	3.4
1	D	312	TRP	3.2
1	C	190	ALA	3.2
1	C	214	ARG	3.2
1	C	313	ASP	3.2
2	G	77	HIS	3.2
1	C	302	LYS	3.1
2	G	22	GLU	3.0
1	A	322	GLU	3.0
1	D	249	VAL	3.0
1	C	318	PHE	3.0
1	C	222	ARG	3.0
1	D	225	GLU	2.9
2	F	65	GLU	2.9
2	G	41	ASN	2.9
1	B	322	GLU	2.9
1	C	232	LEU	2.8
1	C	225	GLU	2.8
1	D	177	GLN	2.8
2	F	5	LEU	2.8
2	H	20	LYS	2.8
1	C	319	PHE	2.7
2	H	40	VAL	2.7
1	C	231	MET	2.7
1	D	234	LYS	2.6
1	C	292	GLU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	294	ILE	2.5
1	C	267	LEU	2.5
1	C	303	ARG	2.5
1	C	309	GLN	2.5
1	D	231	MET	2.4
2	H	86	GLN	2.4
1	D	250	MET	2.4
2	G	43	GLU	2.4
2	F	23	ASN	2.4
1	C	229	GLN	2.3
2	G	38	MET	2.3
2	H	5	LEU	2.3
1	D	308	LYS	2.3
1	D	258	VAL	2.3
1	D	304	ASP	2.2
1	D	227	ALA	2.2
2	F	44	GLU	2.2
2	H	77	HIS	2.2
1	C	296	ASP	2.2
1	C	223	ASN	2.2
1	C	183	SER	2.2
2	H	22	GLU	2.2
2	G	86	GLN	2.2
1	D	179	LEU	2.1
2	E	35	LYS	2.1
2	G	44	GLU	2.1
1	B	194	LYS	2.1
2	E	65	GLU	2.1
1	D	248	ARG	2.1
2	E	3	ASN	2.1
1	C	311	GLY	2.0
1	D	246	LEU	2.0
1	D	317	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.