



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 19, 2023 – 09:37 PM JST

PDB ID : 7BW2
Title : Crystal Structure of Cyanobacterial PSI Monomer from *T.elongatus* at 6.5 Å Resolution
Authors : Kurisu, G.; Coruh, O.; Tanaka, H.; Eithar, E.M.; Mian, Y.
Deposited on : 2020-04-13
Resolution : 6.50 Å (reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

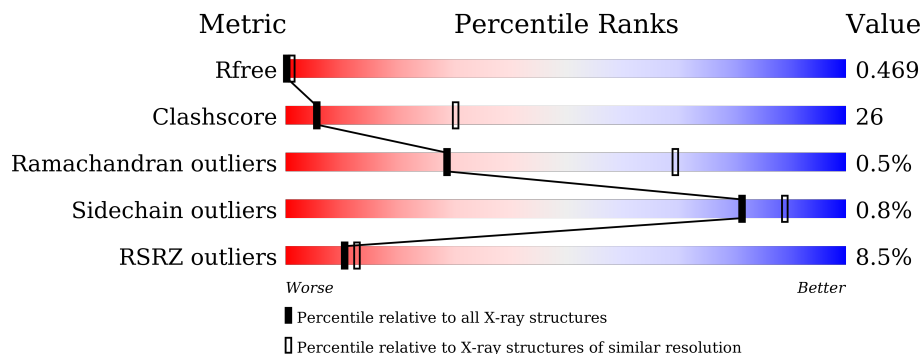
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 6.50 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	1000 (9.00-3.90)
Clashscore	141614	1064 (9.00-3.90)
Ramachandran outliers	138981	1012 (9.00-3.88)
Sidechain outliers	138945	1010 (9.00-3.84)
RSRZ outliers	127900	1002 (9.00-3.80)

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 of 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 $\leq 5\%$. 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	755	 5% 73% 25% ..
2	B	740	 11% 73% 27%
3	C	81	 15% 47% 49% ..
4	D	139	 12% 59% 39% ..
5	E	76	 5% 62% 29% 9%
6	F	164	 2% 43% 42% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	I	38	
8	J	41	
9	K	83	
10	L	155	
11	M	31	
12	X	39	

2 Entry composition [i](#)

There are 12 unique types of molecules in this entry. The entry contains 17345 atoms, of which 111 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 Photosystem I P700 chlorophyll a apoprotein A1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	743	5914	3807	111	991	979	26	0	0	0

- Molecule 2 is a protein called Photosystem I P700 chlorophyll a apoprotein A2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	739	5879	3867	986	1005	21	0	0	0

- Molecule 3 is a protein called Photosystem I iron-sulfur center.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	80	598	367	103	117	11	0	0	0

- Molecule 4 is a protein called Photosystem I reaction center subunit II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	138	1075	682	186	204	3	0	0	0

- Molecule 5 is a protein called Photosystem I reaction center subunit IV.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	E	69	539	342	93	104	0	0	0

- Molecule 6 is a protein called Photosystem I reaction center subunit III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	141	1065	680	184	197	4	0	0	0

- Molecule 7 is a protein called Photosystem I reaction center subunit VIII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	I	38	301	208	40	48	5	0	0	0

- Molecule 8 is a protein called Photosystem I reaction center subunit IX.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	J	41	338	231	51	54	2	0	0	0

- Molecule 9 is a protein called Photosystem I reaction center subunit PsaK.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
9	K	47	227	133	47	47	0	0	0

- Molecule 10 is a protein called Photosystem I reaction center subunit XI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	L	128	936	609	154	170	3	0	0	0

- Molecule 11 is a protein called Photosystem I reaction center subunit XII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	M	31	241	161	36	43	1	0	0	0

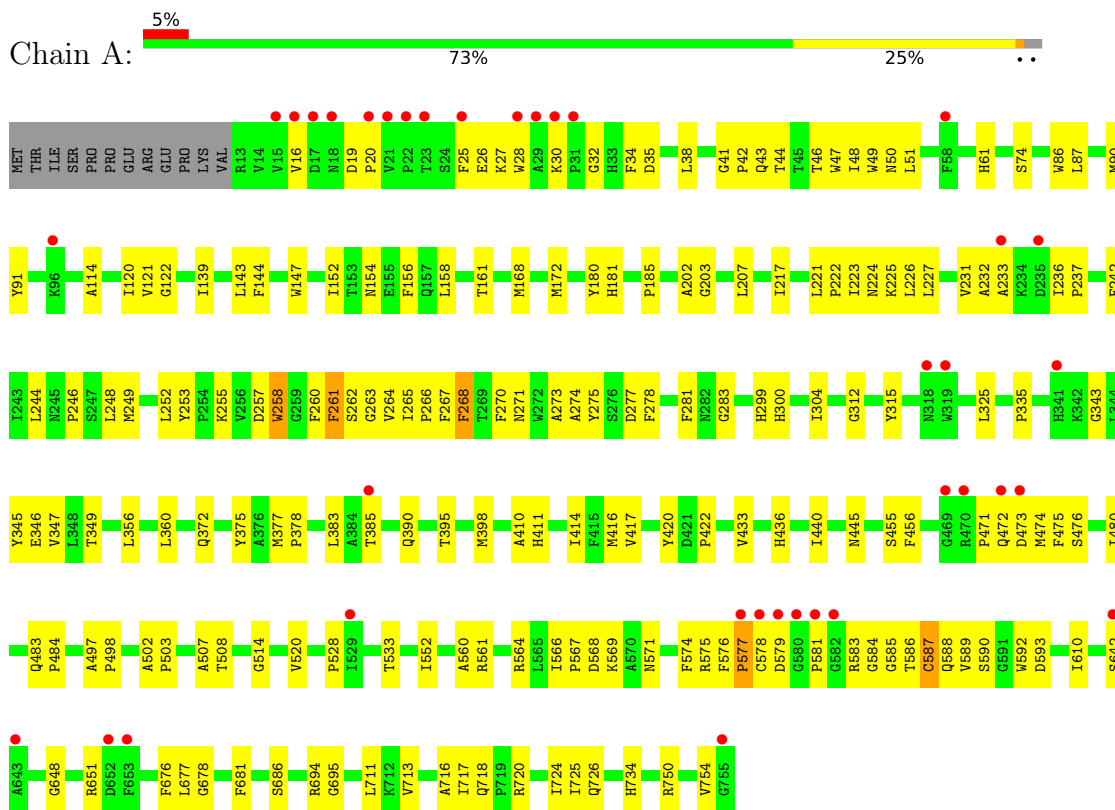
- Molecule 12 is a protein called Photosystem I 4.8K protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
12	X	29	232	163	34	35	0	0	0

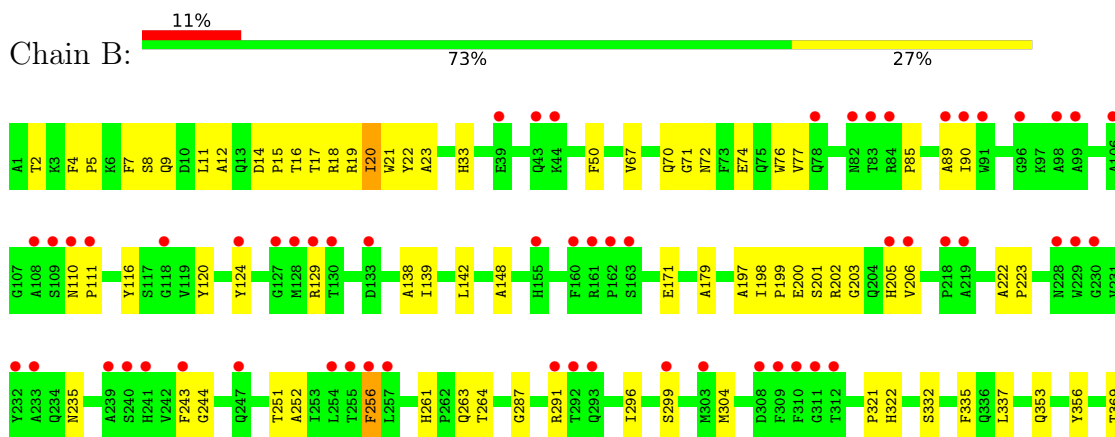
3 Residue-property plots

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: Photosystem I P700 chlorophyll a apoprotein A1

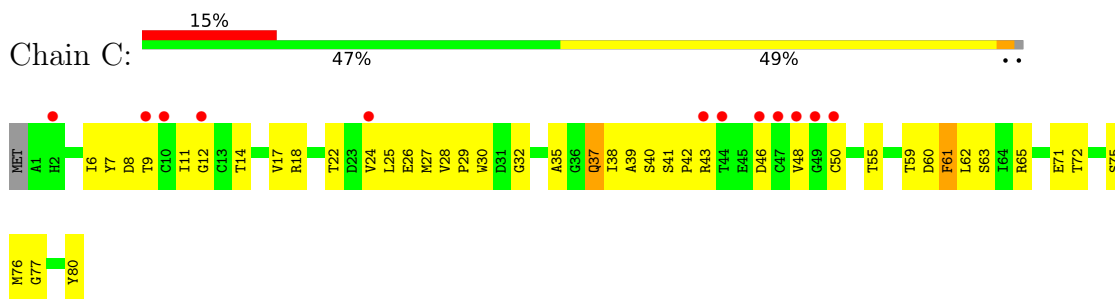


- Molecule 2: Photosystem I P700 chlorophyll a apoprotein A2

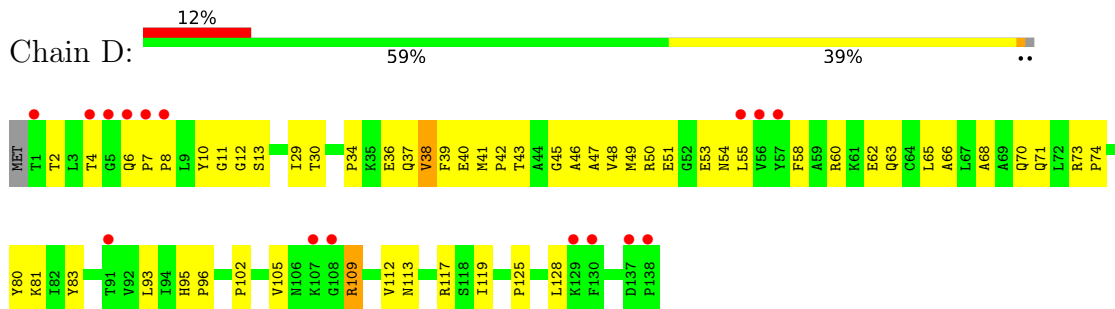




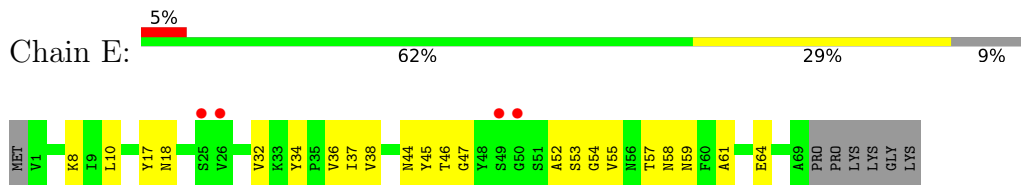
• Molecule 3: Photosystem I iron-sulfur center



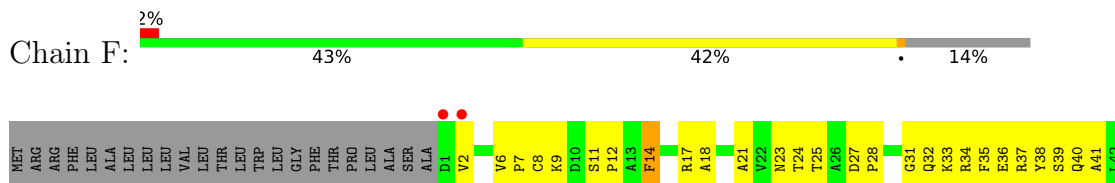
• Molecule 4: Photosystem I reaction center subunit II

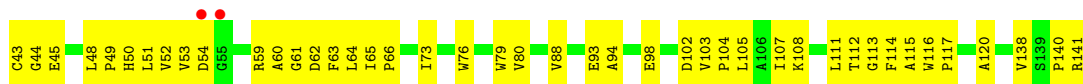


• Molecule 5: Photosystem I reaction center subunit IV



• Molecule 6: Photosystem I reaction center subunit III

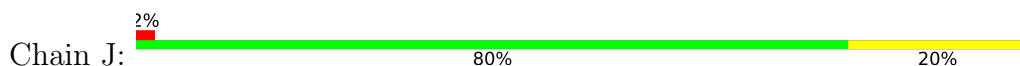




- Molecule 7: Photosystem I reaction center subunit VIII



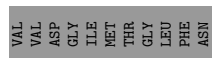
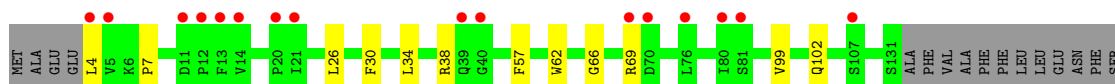
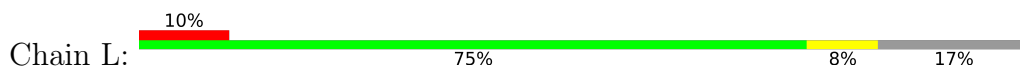
- Molecule 8: Photosystem I reaction center subunit IX



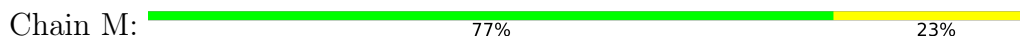
- Molecule 9: Photosystem I reaction center subunit PsaK



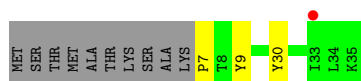
- Molecule 10: Photosystem I reaction center subunit XI



- Molecule 11: Photosystem I reaction center subunit XII



- Molecule 12: Photosystem I 4.8K protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	187.03Å 187.03Å 233.81Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.57 – 6.50 49.57 – 6.50	Depositor EDS
% Data completeness (in resolution range)	98.6 (49.57-6.50) 98.9 (49.57-6.50)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	8.76 (at 6.68Å)	Xtrriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.409 , 0.486 0.407 , 0.469	Depositor DCC
R_{free} test set	479 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	296.5	Xtrriage
Anisotropy	0.123	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 193.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtrriage
Estimated twinning fraction	0.000 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.71	EDS
Total number of atoms	17345	wwPDB-VP
Average B, all atoms (Å ²)	274.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.87% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.54	0/6003	0.65	2/8188 (0.0%)
2	B	0.55	0/6096	0.62	0/8332
3	C	0.70	0/608	0.78	0/824
4	D	0.59	0/1101	0.79	2/1492 (0.1%)
5	E	0.64	0/551	0.74	0/750
6	F	0.66	0/1087	0.71	0/1476
7	I	0.55	0/312	0.74	1/425 (0.2%)
8	J	0.47	0/350	0.62	0/477
9	K	0.86	0/225	0.87	0/307
10	L	0.58	0/960	0.61	0/1304
11	M	0.53	0/244	0.57	0/332
12	X	0.55	0/241	0.68	0/330
All	All	0.57	0/17778	0.67	5/24237 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	109	ARG	NE-CZ-NH2	-10.02	115.29	120.30
1	A	587	CYS	N-CA-CB	9.01	126.81	110.60
4	D	109	ARG	NE-CZ-NH1	5.91	123.25	120.30
7	I	20	TRP	C-N-CA	-5.45	108.07	121.70
1	A	587	CYS	CB-CA-C	-5.30	99.81	110.40

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	5803	111	5664	270	0
2	B	5879	0	5634	323	0
3	C	598	0	588	101	0
4	D	1075	0	1077	76	0
5	E	539	0	527	40	0
6	F	1065	0	1076	134	0
7	I	301	0	306	56	0
8	J	338	0	347	10	0
9	K	227	0	116	1	0
10	L	936	0	944	19	0
11	M	241	0	264	3	0
12	X	232	0	220	2	0
All	All	17234	111	16763	877	0

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 26.

The worst 5 of 877 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)
2:B:562:ALA:HB2	2:B:579:TRP:CB	1.34	1.53
1:A:44:THR:CG2	1:A:717:ILE:HB	1.33	1.52
3:C:26:GLU:C	4:D:109:ARG:HH22	1.07	1.51
1:A:264:VAL:HG23	1:A:268:PHE:CE1	1.48	1.48
7:I:20:TRP:HA	7:I:23:PRO:CG	1.44	1.47

There are no symmetry-related clashes.

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	741/755 (98%)	697 (94%)	40 (5%)	4 (0%)	29	69
2	B	737/740 (100%)	703 (95%)	34 (5%)	0	100	100
3	C	78/81 (96%)	71 (91%)	6 (8%)	1 (1%)	12	48
4	D	136/139 (98%)	124 (91%)	10 (7%)	2 (2%)	10	46
5	E	67/76 (88%)	60 (90%)	6 (9%)	1 (2%)	10	46
6	F	139/164 (85%)	128 (92%)	9 (6%)	2 (1%)	11	46
7	I	36/38 (95%)	32 (89%)	3 (8%)	1 (3%)	5	30
8	J	39/41 (95%)	37 (95%)	2 (5%)	0	100	100
9	K	43/83 (52%)	39 (91%)	3 (7%)	1 (2%)	6	34
10	L	126/155 (81%)	121 (96%)	5 (4%)	0	100	100
11	M	29/31 (94%)	27 (93%)	2 (7%)	0	100	100
12	X	27/39 (69%)	24 (89%)	3 (11%)	0	100	100
All	All	2198/2342 (94%)	2063 (94%)	123 (6%)	12 (0%)	29	69

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	261	PHE
4	D	38	VAL
7	I	36	GLY
9	K	41	PRO
3	C	61	PHE

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	591/603 (98%)	587 (99%)	4 (1%)	84	90
2	B	595/597 (100%)	591 (99%)	4 (1%)	84	90
3	C	67/68 (98%)	66 (98%)	1 (2%)	65	80
4	D	115/116 (99%)	114 (99%)	1 (1%)	78	88

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	E	59/65 (91%)	59 (100%)	0	100	100
6	F	109/128 (85%)	108 (99%)	1 (1%)	78	88
7	I	32/32 (100%)	31 (97%)	1 (3%)	40	62
8	J	36/36 (100%)	36 (100%)	0	100	100
10	L	98/120 (82%)	97 (99%)	1 (1%)	76	86
11	M	26/26 (100%)	25 (96%)	1 (4%)	33	57
12	X	20/31 (64%)	20 (100%)	0	100	100
All	All	1748/1822 (96%)	1734 (99%)	14 (1%)	81	89

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	575	ASP
3	C	37	GLN
11	M	5	ASP
7	I	20	TRP
10	L	57	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	497	ASN
2	B	614	GLN
6	F	23	ASN
4	D	113	ASN
2	B	43	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	743/755 (98%)	0.20	37 (4%) 28 28	120, 234, 334, 449	0
2	B	739/740 (99%)	0.50	80 (10%) 5 9	138, 283, 382, 480	0
3	C	80/81 (98%)	0.67	12 (15%) 2 5	172, 259, 413, 481	0
4	D	138/139 (99%)	0.56	16 (11%) 4 8	152, 275, 362, 380	0
5	E	69/76 (90%)	0.26	4 (5%) 23 23	206, 290, 382, 410	0
6	F	141/164 (85%)	-0.10	4 (2%) 53 46	178, 275, 373, 446	0
7	I	38/38 (100%)	0.86	8 (21%) 1 3	274, 349, 476, 481	0
8	J	41/41 (100%)	-0.15	1 (2%) 59 52	165, 241, 310, 358	0
9	K	47/83 (56%)	1.04	9 (19%) 1 3	184, 298, 338, 354	0
10	L	128/155 (82%)	0.62	16 (12%) 3 8	259, 361, 556, 628	0
11	M	31/31 (100%)	0.12	0 100 100	210, 350, 450, 487	0
12	X	29/39 (74%)	-0.01	1 (3%) 45 40	166, 262, 394, 411	0
All	All	2224/2342 (94%)	0.36	188 (8%) 10 13	120, 268, 391, 628	0

The worst 5 of 188 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	110	ASN	7.8
3	C	47	CYS	7.6
2	B	566	ASP	7.4
2	B	91	TRP	7.4
4	D	6	GLN	6.9

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.