wwPDB X-ray Structure Validation Summary Report (i)

Jun 7, 2022 – 06:31 PM JST

PDB ID	:	7F4L
Title	:	Crystal structure of MTA1-p1-p2 complex
Authors	:	Chen, J.; Liu, L.
Deposited on	:	2021-06-21
Resolution	:	2.72 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

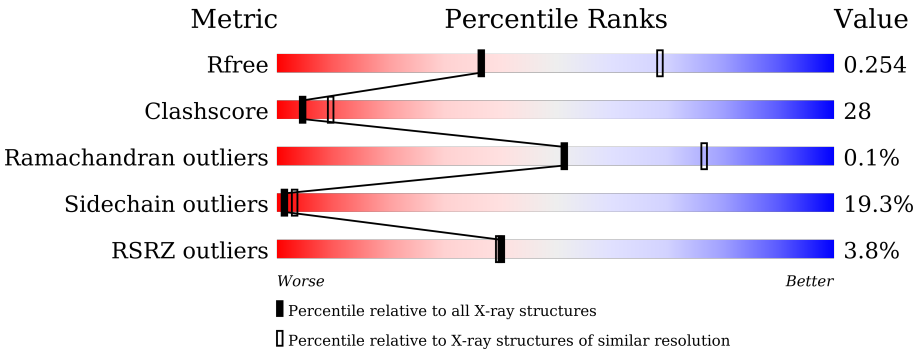
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

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 2.72 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	3359 (2.74-2.70)
Clashscore	141614	3686 (2.74-2.70)
Ramachandran outliers	138981	3622 (2.74-2.70)
Sidechain outliers	138945	3623 (2.74-2.70)
RSRZ outliers	127900	3276 (2.74-2.70)

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 <=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	142	4% 46% 37% 7% 9%
1	B	142	6% 49% 35% 8% 8%
2	C	247	3% 42% 34% 11% 13%
2	D	247	3% 41% 33% 13% 13%
3	E	309	9% 88%
3	F	309	6% 92%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6142 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 Transmembrane protein, putative.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	130	Total	C	N	O	S	Se	0	0	0
			1055	669	177	205	2	2			
1	A	129	Total	C	N	O	S	Se	0	0	0
			1053	667	177	205	2	2			

- Molecule 2 is a protein called MT-a70 family protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	215	Total	C	N	O	S	Se	0	0	0
			1721	1094	298	319	3	7			
2	D	216	Total	C	N	O	S	Se	0	0	0
			1739	1108	299	322	3	7			

- Molecule 3 is a protein called p1 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	37	Total	C	N	O	Se	0	0	0
			308	196	51	60	1			
3	F	25	Total	C	N	O		0	0	0
			208	134	37	37				

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	13	Total	O	0	0
			13	13		
4	C	11	Total	O	0	0
			11	11		
4	D	14	Total	O	0	0
			14	14		
4	E	3	Total	O	0	0
			3	3		

Continued on next page...

Continued from previous page...

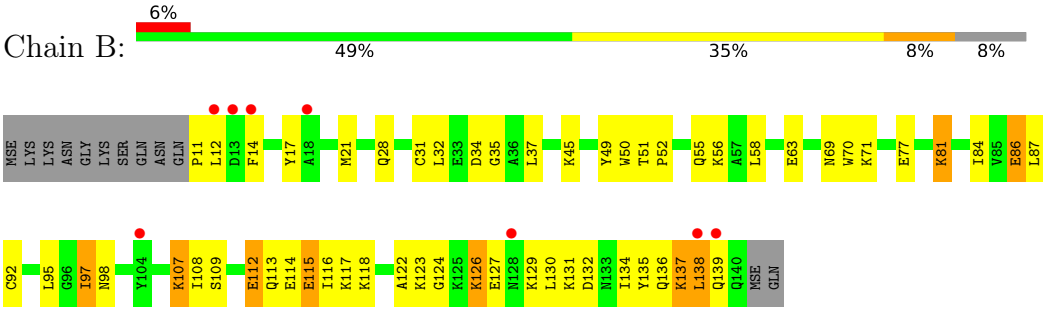
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	F	3	Total O 3 3	0	0
4	A	14	Total O 14 14	0	0

3 Residue-property plots (i)

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.

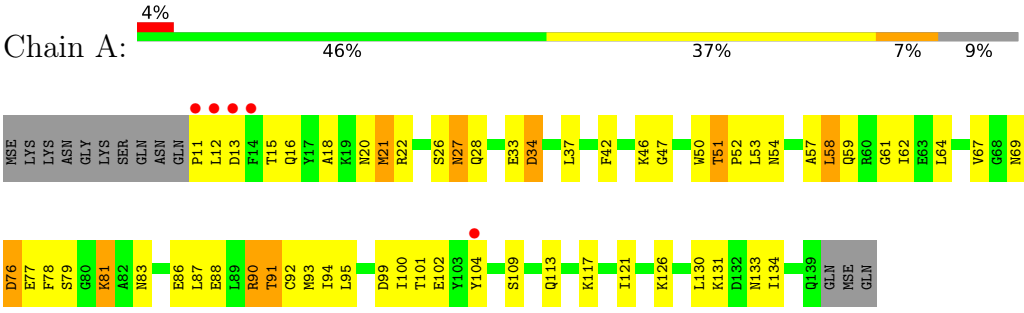
• Molecule 1: Transmembrane protein, putative

Chain B:



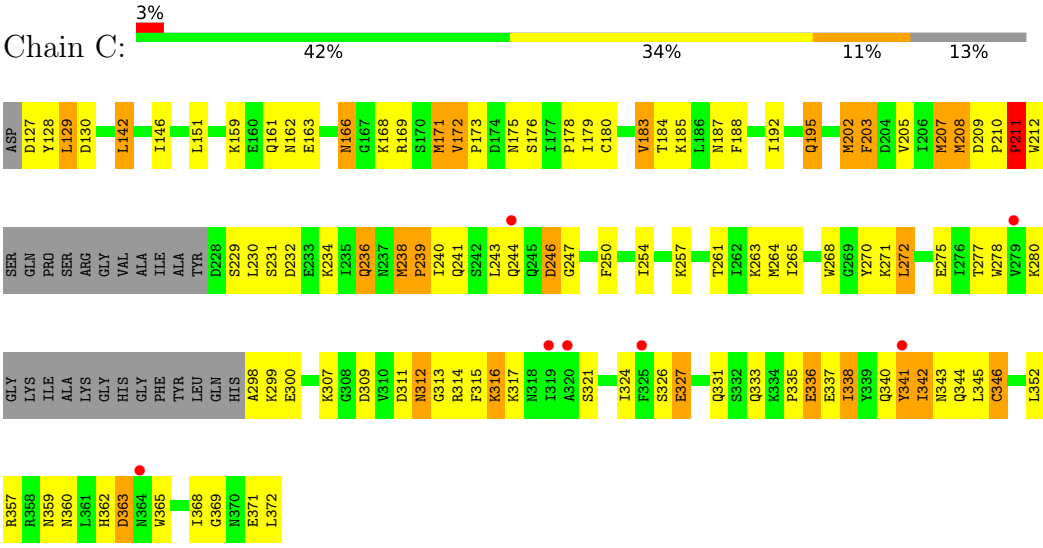
• Molecule 1: Transmembrane protein, putative

Chain A:



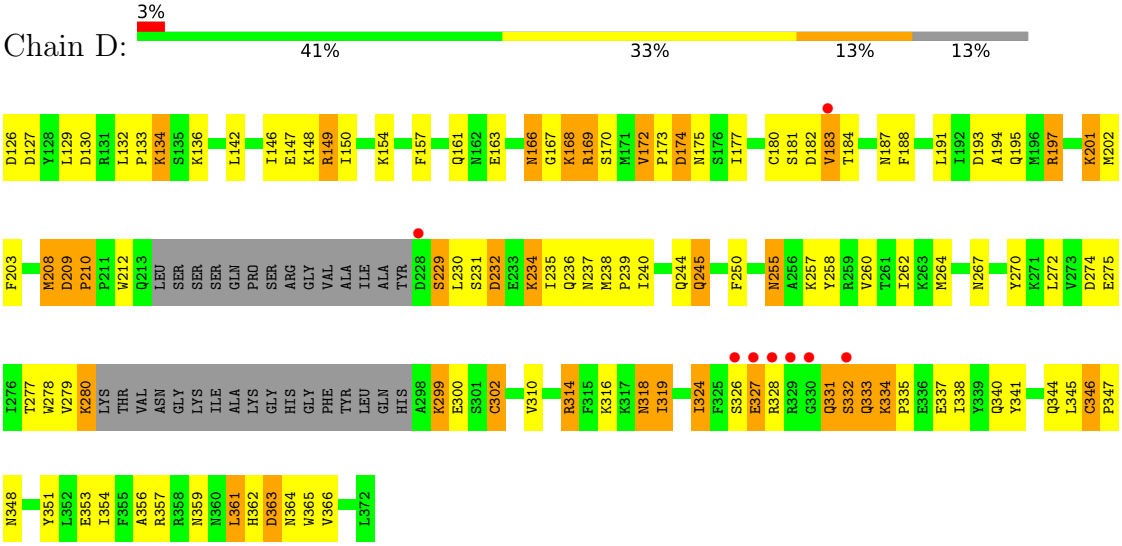
• Molecule 2: MT-a70 family protein

Chain C:



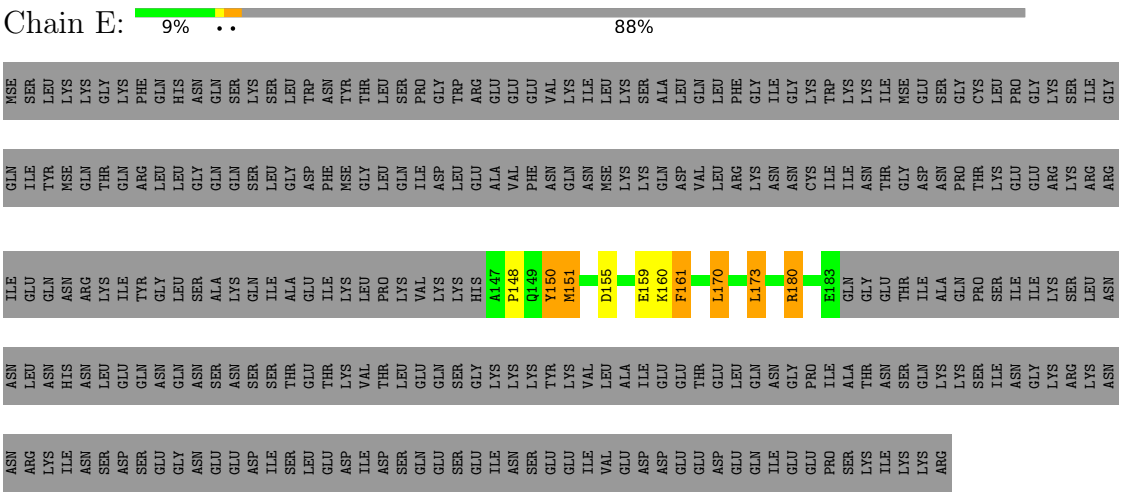
• Molecule 2: MT-a70 family protein

Chain D:



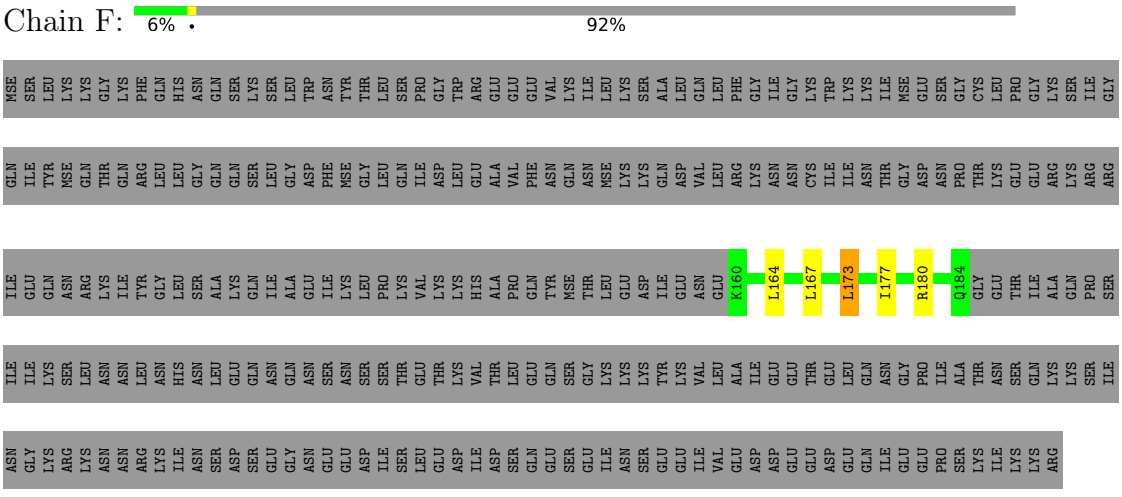
• Molecule 3: p1 protein

Chain E:



• Molecule 3: p1 protein

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	82.77Å 86.49Å 160.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.00 – 2.72 48.00 – 2.72	Depositor EDS
% Data completeness (in resolution range)	97.9 (48.00-2.72) 97.9 (48.00-2.72)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.50 (at 2.73Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.240 , 0.254 0.240 , 0.254	Depositor DCC
R_{free} test set	1554 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	38.8	Xtriage
Anisotropy	0.032	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 40.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.032 for k,h,-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	6142	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% of the height of the origin peak. No significant pseudotranslation is detected.*

[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

5.1 Standard geometry

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	0.90	1/1067 (0.1%)	0.79	2/1426 (0.1%)
1	B	0.79	1/1069 (0.1%)	0.75	2/1429 (0.1%)
2	C	0.84	3/1747 (0.2%)	0.84	6/2343 (0.3%)
2	D	0.90	3/1766 (0.2%)	0.91	8/2368 (0.3%)
3	E	0.89	0/310	0.98	2/416 (0.5%)
3	F	0.82	0/210	0.71	0/283
All	All	0.86	8/6169 (0.1%)	0.84	20/8265 (0.2%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	302	CYS	CB-SG	-6.58	1.71	1.82
1	B	86	GLU	CB-CG	-5.60	1.41	1.52
2	D	133	PRO	N-CD	5.36	1.55	1.47
2	C	239	PRO	N-CD	5.23	1.55	1.47
1	A	52	PRO	N-CD	5.16	1.55	1.47

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	203	PHE	N-CA-CB	-12.74	87.66	110.60
3	E	150	TYR	N-CA-CB	-11.57	89.78	110.60
2	C	208	MSE	N-CA-CB	-6.67	98.59	110.60
2	C	346	CYS	C-N-CD	6.06	141.12	128.40
2	D	210	PRO	C-N-CD	6.01	141.02	128.40

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	1053	0	1041	54	0
1	B	1055	0	1039	54	0
2	C	1721	0	1681	104	0
2	D	1739	0	1706	138	0
3	E	308	0	305	13	0
3	F	208	0	213	4	0
4	A	14	0	0	1	0
4	B	13	0	0	1	0
4	C	11	0	0	2	0
4	D	14	0	0	4	0
4	E	3	0	0	1	0
4	F	3	0	0	0	0
All	All	6142	0	5985	342	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 28.

The worst 5 of 342 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:C:192:ILE:HG23	2:C:202:MSE:CE	1.37	1.51
2:C:213:GLN:HG3	2:C:229:SER:OG	1.30	1.28
1:B:131:LYS:HD2	1:B:136:GLN:OE1	1.25	1.27
2:C:336:GLU:OE2	2:C:362:HIS:CE1	1.95	1.19
1:B:131:LYS:CD	1:B:136:GLN:OE1	1.92	1.17

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	127/142 (89%)	125 (98%)	2 (2%)	0	100 100
1	B	128/142 (90%)	126 (98%)	2 (2%)	0	100 100
2	C	209/247 (85%)	206 (99%)	3 (1%)	0	100 100
2	D	210/247 (85%)	203 (97%)	7 (3%)	0	100 100
3	E	35/309 (11%)	34 (97%)	0	1 (3%)	4 10
3	F	23/309 (7%)	21 (91%)	2 (9%)	0	100 100
All	All	732/1396 (52%)	715 (98%)	16 (2%)	1 (0%)	51 77

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	148	PRO

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	113/123 (92%)	91 (80%)	22 (20%)	1 3
1	B	112/123 (91%)	89 (80%)	23 (20%)	1 2
2	C	185/212 (87%)	151 (82%)	34 (18%)	1 3
2	D	188/212 (89%)	148 (79%)	40 (21%)	1 2
3	E	33/276 (12%)	28 (85%)	5 (15%)	3 6
3	F	22/276 (8%)	20 (91%)	2 (9%)	9 21
All	All	653/1222 (53%)	527 (81%)	126 (19%)	1 3

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	127	ASP
1	A	58	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	184	THR
1	A	34	ASP
1	A	101	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	359	ASN
1	A	54	ASN
2	C	331	GLN
2	C	360	ASN
2	C	362	HIS

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

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	127/142 (89%)	0.34	5 (3%) 39 39	9, 30, 54, 83	0
1	B	128/142 (90%)	0.41	8 (6%) 20 19	10, 32, 59, 81	0
2	C	208/247 (84%)	0.41	7 (3%) 45 45	11, 34, 65, 80	0
2	D	209/247 (84%)	0.22	8 (3%) 40 40	7, 29, 72, 92	0
3	E	36/309 (11%)	-0.00	0 100 100	6, 13, 33, 47	0
3	F	25/309 (8%)	-0.11	0 100 100	7, 14, 39, 52	0
All	All	733/1396 (52%)	0.31	28 (3%) 40 40	6, 30, 64, 92	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	11	PRO	7.8
1	A	12	LEU	6.6
2	D	326	SER	6.4
1	B	12	LEU	5.5
2	D	329	ARG	4.4

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.