Full wwPDB X-ray Structure Validation Report ⓘ

Aug 25, 2020 – 02:36 PM BST

PDB ID	:	5JS4
Title	:	Crystal structure of phiAB6 tailspike
Authors	:	Lee, I.M.; Tu, I.F.; Huang, K.F.; Wu, S.H.
Deposited on	:	2016-05-07
Resolution	:	1.48 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see references ⓘ) were used in the production of this report:

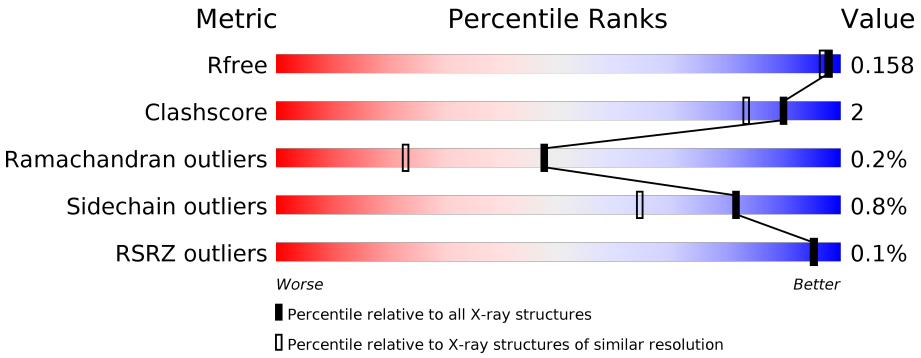
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13

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 1.48 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	4690 (1.50-1.46)
Clashscore	141614	4955 (1.50-1.46)
Ramachandran outliers	138981	4846 (1.50-1.46)
Sidechain outliers	138945	4844 (1.50-1.46)
RSRZ outliers	127900	4614 (1.50-1.46)

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	719	72% · 24%
1	B	719	73% · 24%
1	C	719	72% · 24%

2 Entry composition

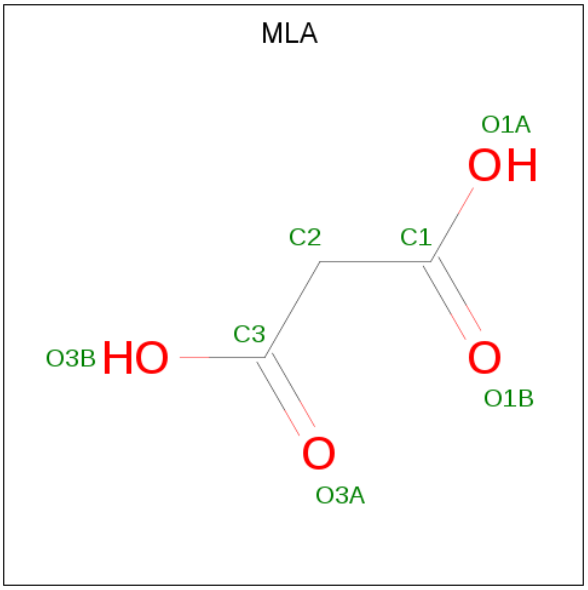
There are 3 unique types of molecules in this entry. The entry contains 15367 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 phiAB6 tailspike.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	547	Total	C	N	O	S	0	0	0
			4179	2632	710	817	20			
1	B	547	Total	C	N	O	S	0	0	0
			4179	2632	710	817	20			
1	C	547	Total	C	N	O	S	0	0	0
			4179	2632	710	817	20			

- Molecule 2 is MALONIC ACID (three-letter code: MLA) (formula: $C_3H_4O_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O		0	0
			7	3	4			
2	A	1	Total	C	O		0	0
			7	3	4			
2	A	1	Total	C	O		0	0
			7	3	4			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C O 7 3 4	0	0
2	B	1	Total C O 7 3 4	0	0
2	B	1	Total C O 7 3 4	0	0
2	C	1	Total C O 7 3 4	0	0
2	C	1	Total C O 7 3 4	0	0
2	C	1	Total C O 7 3 4	0	0

- Molecule 3 is water.

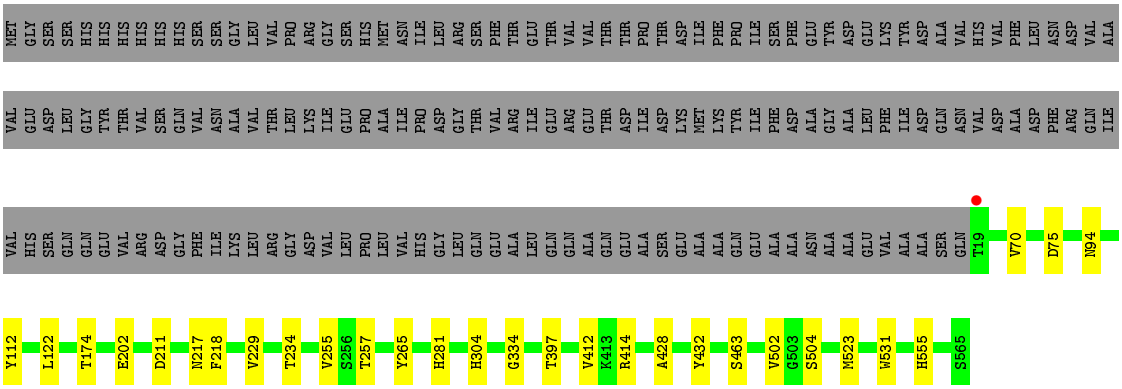
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	931	Total O 931 931	0	0
3	B	902	Total O 902 902	0	0
3	C	934	Total O 934 934	0	0

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.

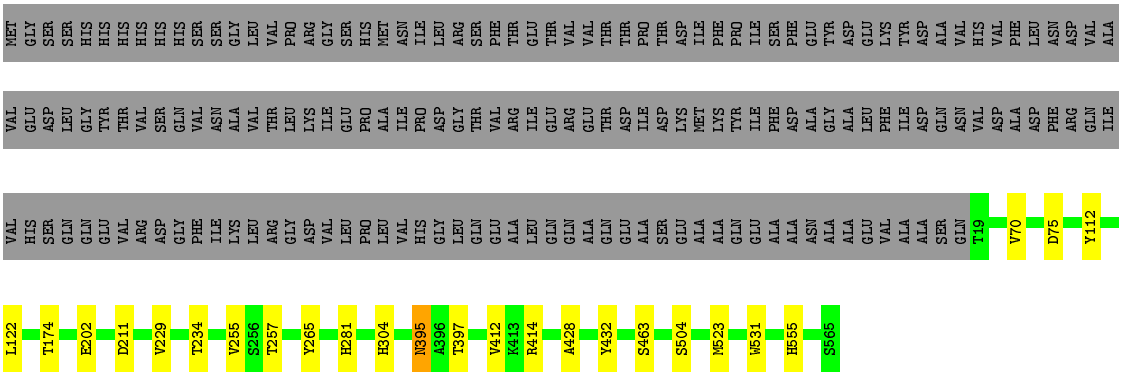
- Molecule 1: phiAB6 tailspike

Chain A:



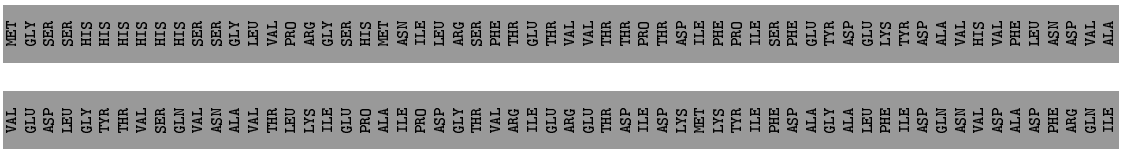
- Molecule 1: phiAB6 tailspike

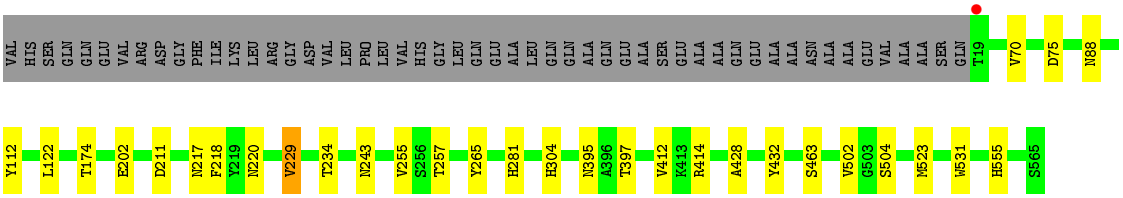
Chain B:



- Molecule 1: phiAB6 tailspike

Chain C:




VAL HIS SER GLN GLN GLU VAL ARG ASP GLY PHE ILE LYS LEU ARG GLY ASP VAL LEU PRO LEU VAL HIS GLY LEU GLN GLU ALA LEU GLN GLN ALA GLN GLU ALA SER GLU ALA ALA GLN GLU ALA ALA ASN ALA ALA GLU VAL ALA ALA SER GLN
T19 V70 D75 N88
Y112 L122 T174 E202 D211 N217 F218 Y219 N220 V229 T234 N243 V255 S256 T257 Y265 H281 H304 N395 A396 T397 V412 K413 R414 A428 Y432 S463 V502 Q503 S504 M523 W531 H555 S565

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	135.02Å 77.98Å 248.08Å 90.00° 100.46° 90.00°	Depositor
Resolution (Å)	19.98 – 1.48 19.69 – 1.48	Depositor EDS
% Data completeness (in resolution range)	93.9 (19.98-1.48) 93.8 (19.69-1.48)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.17 (at 1.48Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.136 , 0.156 0.137 , 0.158	Depositor DCC
R_{free} test set	19849 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	12.9	Xtriage
Anisotropy	0.303	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 40.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.487 for 1/2*h+3/2*k,1/2*h-1/2*k,-1/2*h-1/2*k-l 0.487 for 1/2*h-3/2*k,-1/2*h-1/2*k,-1/2*h+1/2*k-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	15367	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.84% 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

Bond lengths and bond angles in the following residue types are not validated in this section: MLA

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	0.27	0/4272	0.54	0/5802
1	B	0.28	0/4272	0.54	0/5802
1	C	0.27	0/4272	0.54	0/5802
All	All	0.27	0/12816	0.54	0/17406

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

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	4179	0	4026	18	0
1	B	4179	0	4026	17	0
1	C	4179	0	4026	18	0
2	A	21	0	6	0	0
2	B	21	0	6	0	0
2	C	21	0	6	0	0
3	A	931	0	0	3	0
3	B	902	0	0	2	0
3	C	934	0	0	3	0
All	All	15367	0	12096	44	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 2.

All (44) 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)
3:B:703:HOH:O	1:C:504:SER:HB3	1.79	0.82
1:C:220:ASN:HD21	1:C:243:ASN:HD22	1.36	0.72
3:A:707:HOH:O	1:B:504:SER:HB3	1.88	0.71
1:A:504:SER:HB3	3:C:703:HOH:O	1.94	0.66
1:A:94:ASN:HD21	1:C:88:ASN:HD21	1.46	0.63
1:C:174:THR:HG21	1:C:202:GLU:HG3	1.87	0.57
1:A:174:THR:HG21	1:A:202:GLU:HG3	1.87	0.56
1:B:174:THR:HG21	1:B:202:GLU:HG3	1.88	0.56
1:A:463:SER:OG	1:A:555:HIS:HE1	1.91	0.53
1:B:463:SER:OG	1:B:555:HIS:HE1	1.93	0.51
1:B:112:TYR:CE1	1:B:122:LEU:HA	2.46	0.51
1:C:463:SER:OG	1:C:555:HIS:HE1	1.93	0.51
1:C:112:TYR:CE1	1:C:122:LEU:HA	2.45	0.51
1:A:112:TYR:CE1	1:A:122:LEU:HA	2.46	0.50
1:B:523:MET:HG3	1:B:531:TRP:CE2	2.46	0.50
1:C:523:MET:HG3	1:C:531:TRP:CE2	2.48	0.49
1:B:257:THR:HA	1:B:281:HIS:O	2.13	0.49
1:A:257:THR:HA	1:A:281:HIS:O	2.13	0.49
1:B:211:ASP:HA	1:B:234:THR:O	2.12	0.49
1:A:523:MET:HG3	1:A:531:TRP:CE2	2.47	0.49
1:C:257:THR:HA	1:C:281:HIS:O	2.13	0.49
1:A:211:ASP:HA	1:A:234:THR:O	2.13	0.49
1:C:211:ASP:HA	1:C:234:THR:O	2.15	0.46
3:A:722:HOH:O	1:C:304:HIS:HD2	2.00	0.44
1:A:75:ASP:HA	1:B:70:VAL:HG22	2.01	0.43
1:B:75:ASP:HA	1:C:70:VAL:HG22	2.01	0.43
1:A:70:VAL:HG22	1:C:75:ASP:HA	2.00	0.42
1:A:504:SER:OG	1:C:502:VAL:HG23	2.19	0.42
1:A:502:VAL:HG23	1:B:504:SER:OG	2.19	0.42
1:A:414:ARG:HA	1:A:428:ALA:O	2.20	0.42
1:C:414:ARG:HA	1:C:428:ALA:O	2.20	0.42
1:A:304:HIS:HD2	3:B:724:HOH:O	2.04	0.41
3:A:722:HOH:O	1:C:281:HIS:HD2	2.04	0.41
1:A:412:VAL:O	1:B:397:THR:HA	2.20	0.41
1:A:304:HIS:HA	1:A:334:GLY:O	2.21	0.41
1:B:304:HIS:HD2	3:C:734:HOH:O	2.04	0.41
1:B:414:ARG:HA	1:B:428:ALA:O	2.21	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:395:ASN:HD22	1:B:395:ASN:C	2.25	0.40
1:B:281:HIS:HD2	3:C:734:HOH:O	2.04	0.40
1:A:217:ASN:O	1:A:218:PHE:HB2	2.22	0.40
1:B:463:SER:OG	1:B:555:HIS:CE1	2.74	0.40
1:C:217:ASN:O	1:C:218:PHE:HB2	2.22	0.40
1:B:412:VAL:O	1:C:397:THR:HA	2.22	0.40
1:A:397:THR:HA	1:C:412:VAL:O	2.21	0.40

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	545/719 (76%)	526 (96%)	18 (3%)	1 (0%)	47	23
1	B	545/719 (76%)	526 (96%)	18 (3%)	1 (0%)	47	23
1	C	545/719 (76%)	526 (96%)	17 (3%)	2 (0%)	34	13
All	All	1635/2157 (76%)	1578 (96%)	53 (3%)	4 (0%)	47	23

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	255	VAL
1	B	255	VAL
1	C	255	VAL
1	C	229	VAL

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	452/597 (76%)	449 (99%)	3 (1%)	84	68
1	B	452/597 (76%)	448 (99%)	4 (1%)	78	59
1	C	452/597 (76%)	448 (99%)	4 (1%)	78	59
All	All	1356/1791 (76%)	1345 (99%)	11 (1%)	81	64

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	229	VAL
1	A	265	TYR
1	A	432	TYR
1	B	229	VAL
1	B	265	TYR
1	B	395	ASN
1	B	432	TYR
1	C	229	VAL
1	C	265	TYR
1	C	395	ASN
1	C	432	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (29) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	88	ASN
1	A	94	ASN
1	A	184	GLN
1	A	235	ASN
1	A	263	ASN
1	A	267	GLN
1	A	281	HIS
1	A	287	ASN
1	A	304	HIS
1	A	555	HIS
1	B	94	ASN
1	B	232	ASN
1	B	263	ASN
1	B	281	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	304	HIS
1	B	314	ASN
1	B	395	ASN
1	B	448	ASN
1	B	555	HIS
1	C	52	ASN
1	C	184	GLN
1	C	209	ASN
1	C	220	ASN
1	C	281	HIS
1	C	287	ASN
1	C	304	HIS
1	C	395	ASN
1	C	448	ASN
1	C	555	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

9 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	MLA	B	603	-	0,6,6	0.00	-	0,7,7	0.00	-
2	MLA	A	603	-	0,6,6	0.00	-	0,7,7	0.00	-
2	MLA	B	602	-	0,6,6	0.00	-	0,7,7	0.00	-
2	MLA	B	601	-	0,6,6	0.00	-	0,7,7	0.00	-
2	MLA	A	601	-	0,6,6	0.00	-	0,7,7	0.00	-
2	MLA	C	601	-	0,6,6	0.00	-	0,7,7	0.00	-
2	MLA	A	602	-	0,6,6	0.00	-	0,7,7	0.00	-
2	MLA	C	602	-	0,6,6	0.00	-	0,7,7	0.00	-
2	MLA	C	603	-	0,6,6	0.00	-	0,7,7	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	MLA	B	603	-	-	0/0/4/4	-
2	MLA	A	603	-	-	0/0/4/4	-
2	MLA	B	602	-	-	0/0/4/4	-
2	MLA	B	601	-	-	0/0/4/4	-
2	MLA	A	601	-	-	0/0/4/4	-
2	MLA	C	601	-	-	0/0/4/4	-
2	MLA	A	602	-	-	0/0/4/4	-
2	MLA	C	602	-	-	0/0/4/4	-
2	MLA	C	603	-	-	0/0/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	547/719 (76%)	-0.32	1 (0%) 95 95	7, 14, 23, 42	0
1	B	547/719 (76%)	-0.35	0 100 100	7, 14, 23, 42	0
1	C	547/719 (76%)	-0.33	1 (0%) 95 95	7, 14, 22, 43	0
All	All	1641/2157 (76%)	-0.33	2 (0%) 95 95	7, 14, 23, 43	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	19	THR	3.3
1	C	19	THR	2.8

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MLA	B	601	7/7	0.84	0.18	24,26,33,37	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MLA	B	602	7/7	0.86	0.14	20,24,28,31	0
2	MLA	B	603	7/7	0.88	0.15	20,24,29,33	0
2	MLA	A	602	7/7	0.89	0.15	21,24,26,31	0
2	MLA	A	603	7/7	0.91	0.15	19,24,29,33	0
2	MLA	C	603	7/7	0.91	0.16	20,24,31,34	0
2	MLA	C	602	7/7	0.92	0.15	21,23,26,30	0
2	MLA	C	601	7/7	0.94	0.13	25,26,32,33	0
2	MLA	A	601	7/7	0.94	0.11	24,24,32,33	0

6.5 Other polymers (i)

There are no such residues in this entry.