Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 10:03 PM GMT

PDB ID	:	5C6M
Title	:	Crystal structure of deoxyribose-phosphate aldolase from Shewanella halifaxensis
Authors	:	Weiergraeber, O.H.; Dick, M.; Bramski, J.; Pietruszka, J.
Deposited on	:	2015-06-23
Resolution	:	1.76 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

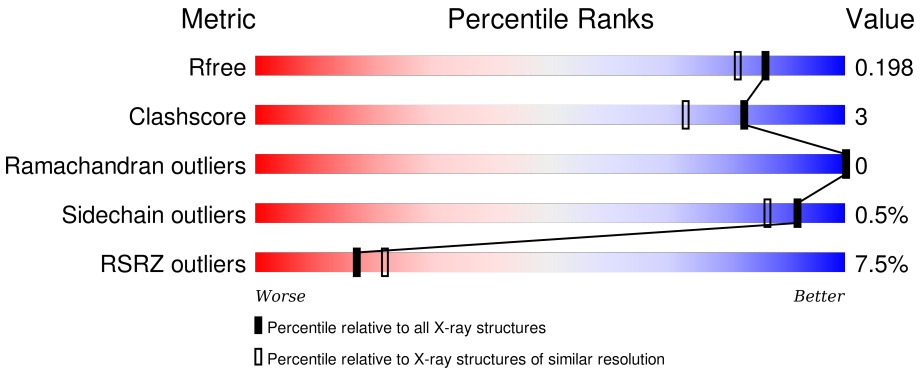
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20026982

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 1.76 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}	91344	1609 (1.76-1.76)
Clashscore	102246	1730 (1.76-1.76)
Ramachandran outliers	100387	1711 (1.76-1.76)
Sidechain outliers	100360	1711 (1.76-1.76)
RSRZ outliers	91569	1610 (1.76-1.76)

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. 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	265	% 86% 8% 6%
1	B	265	% 88% 6% 6%
1	C	265	3% 87% 6% 7%
1	D	265	23% 84% 10% 6%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8289 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 Deoxyribose-phosphate aldolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	248	Total	C	N	O	S	0	5	1
			1872	1181	312	370	9			
1	B	248	Total	C	N	O	S	0	9	0
			1896	1195	317	376	8			
1	C	246	Total	C	N	O	S	0	2	0
			1816	1149	298	361	8			
1	D	248	Total	C	N	O	S	0	1	0
			1777	1120	297	352	8			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	258	LEU	-	expression tag	UNP B0TQ91
A	259	GLU	-	expression tag	UNP B0TQ91
A	260	HIS	-	expression tag	UNP B0TQ91
A	261	HIS	-	expression tag	UNP B0TQ91
A	262	HIS	-	expression tag	UNP B0TQ91
A	263	HIS	-	expression tag	UNP B0TQ91
A	264	HIS	-	expression tag	UNP B0TQ91
A	265	HIS	-	expression tag	UNP B0TQ91
B	258	LEU	-	expression tag	UNP B0TQ91
B	259	GLU	-	expression tag	UNP B0TQ91
B	260	HIS	-	expression tag	UNP B0TQ91
B	261	HIS	-	expression tag	UNP B0TQ91
B	262	HIS	-	expression tag	UNP B0TQ91
B	263	HIS	-	expression tag	UNP B0TQ91
B	264	HIS	-	expression tag	UNP B0TQ91
B	265	HIS	-	expression tag	UNP B0TQ91
C	258	LEU	-	expression tag	UNP B0TQ91
C	259	GLU	-	expression tag	UNP B0TQ91
C	260	HIS	-	expression tag	UNP B0TQ91
C	261	HIS	-	expression tag	UNP B0TQ91
C	262	HIS	-	expression tag	UNP B0TQ91

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	263	HIS	-	expression tag	UNP B0TQ91
C	264	HIS	-	expression tag	UNP B0TQ91
C	265	HIS	-	expression tag	UNP B0TQ91
D	258	LEU	-	expression tag	UNP B0TQ91
D	259	GLU	-	expression tag	UNP B0TQ91
D	260	HIS	-	expression tag	UNP B0TQ91
D	261	HIS	-	expression tag	UNP B0TQ91
D	262	HIS	-	expression tag	UNP B0TQ91
D	263	HIS	-	expression tag	UNP B0TQ91
D	264	HIS	-	expression tag	UNP B0TQ91
D	265	HIS	-	expression tag	UNP B0TQ91

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Na 1 1	0	0
2	A	1	Total Na 1 1	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Cl 1 1	0	0
3	A	1	Total Cl 1 1	0	0
3	D	1	Total Cl 1 1	0	0
3	C	1	Total Cl 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	324	Total O 324 324	0	0
4	B	308	Total O 308 308	0	0
4	C	189	Total O 189 189	0	0

Continued on next page...

Continued from previous page...

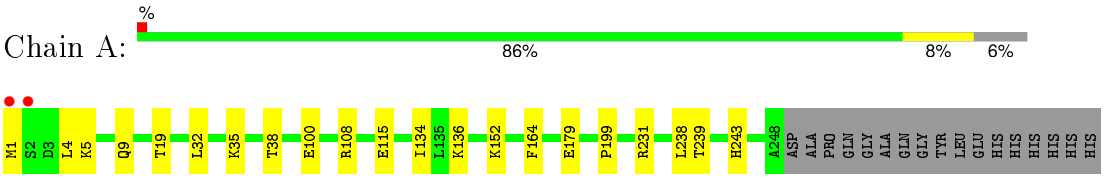
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	101	Total	O	0	0
			101	101		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

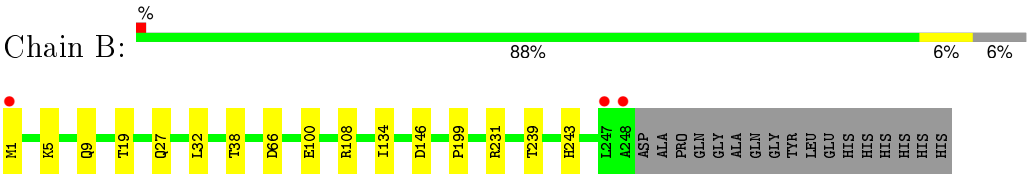
- Molecule 1: Deoxyribose-phosphate aldolase

Chain A:



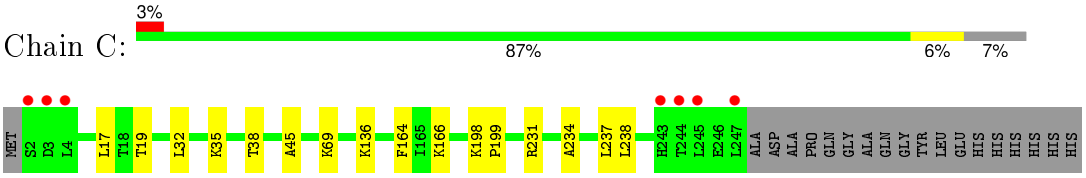
- Molecule 1: Deoxyribose-phosphate aldolase

Chain B:



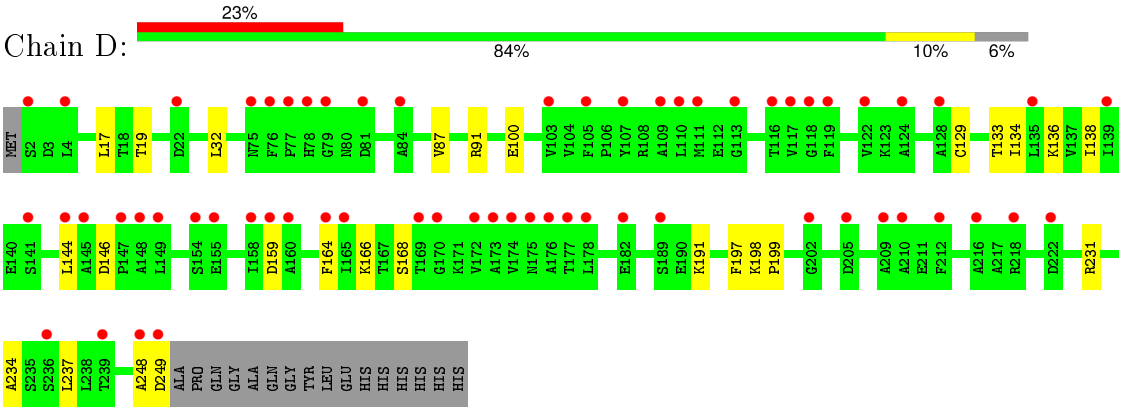
- Molecule 1: Deoxyribose-phosphate aldolase

Chain C:



- Molecule 1: Deoxyribose-phosphate aldolase

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	154.03Å 52.51Å 143.21Å 90.00° 122.34° 90.00°	Depositor
Resolution (Å)	48.70 – 1.76 48.70 – 1.76	Depositor EDS
% Data completeness (in resolution range)	99.5 (48.70-1.76) 99.5 (48.70-1.76)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.77 (at 1.76Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.168 , 0.198 0.168 , 0.198	Depositor DCC
R_{free} test set	4795 reflections (5.00%)	DCC
Wilson B-factor (Å^2)	23.4	Xtriage
Anisotropy	0.424	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 46.6	EDS
Estimated twinning fraction	0.017 for h,-k,-h-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	1 of 95913 reflections (0.001%)	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	8289	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.46% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NA, CL

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	0.35	0/1897	0.50	0/2574
1	B	0.36	0/1921	0.51	0/2607
1	C	0.27	0/1841	0.45	0/2507
1	D	0.25	0/1802	0.42	0/2455
All	All	0.31	0/7461	0.47	0/10143

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	1872	0	1889	13	0
1	B	1896	0	1912	9	0
1	C	1816	0	1801	8	0
1	D	1777	0	1723	15	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	1	0	0	0	0
4	A	324	0	0	4	0
4	B	308	0	0	4	0
4	C	189	0	0	0	0
4	D	101	0	0	0	0
All	All	8289	0	7325	45	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 3.

All (45) 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:A:9:GLN:NE2	4:A:401:HOH:O	2.21	0.73
1:B:108[B]:ARG:NH1	4:B:402:HOH:O	2.21	0.73
1:A:108[B]:ARG:NH1	4:A:404:HOH:O	2.23	0.71
1:C:19:THR:HG22	1:C:32:LEU:HD22	1.78	0.66
1:A:19:THR:HG22	1:A:32:LEU:HD22	1.82	0.61
1:B:27[B]:GLN:NE2	4:B:404:HOH:O	2.24	0.61
1:C:234:ALA:HB1	1:C:237:LEU:HB3	1.87	0.56
1:D:144:LEU:O	1:D:146:ASP:N	2.41	0.53
1:C:166:LYS:HD2	1:C:198:LYS:HD3	1.91	0.53
1:D:138:ILE:HG12	1:D:166:LYS:HD3	1.91	0.52
1:D:19:THR:HG22	1:D:32:LEU:HD22	1.92	0.52
1:B:19:THR:HG22	1:B:32:LEU:HD22	1.91	0.52
1:A:1:MET:HB3	1:A:4:LEU:HB2	1.93	0.51
1:D:129:CYS:HB3	1:D:133:THR:O	2.11	0.50
1:C:35:LYS:HE2	1:C:238:LEU:HD23	1.94	0.50
1:D:136:LYS:HG2	1:D:164:PHE:HB2	1.92	0.50
1:D:234:ALA:HB1	1:D:237:LEU:HB3	1.94	0.50
1:C:199:PRO:HD2	1:C:231:ARG:O	2.12	0.49
1:D:159:ASP:OD1	1:D:191:LYS:NZ	2.45	0.49
1:C:17:LEU:HG	1:C:32:LEU:HD11	1.96	0.47
1:B:100:GLU:HG2	1:B:134:ILE:HB	1.96	0.47
1:A:100:GLU:HG2	1:A:134:ILE:HG12	1.97	0.47
1:A:108[B]:ARG:NH1	4:A:412:HOH:O	2.47	0.47
1:C:45:ALA:HB2	1:C:69:LYS:HB2	1.97	0.46
1:D:166:LYS:HD2	1:D:198:LYS:HD3	1.96	0.46
1:B:199:PRO:HD2	1:B:231:ARG:O	2.16	0.46
1:D:87:VAL:HG12	1:D:91:ARG:HD2	1.98	0.46
1:A:199:PRO:HD2	1:A:231:ARG:O	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:17:LEU:HG	1:D:32:LEU:HD11	1.98	0.45
1:B:1:MET:O	1:B:5:LYS:N	2.38	0.45
1:B:146:ASP:OD1	4:B:403:HOH:O	2.21	0.45
1:A:1:MET:O	1:A:5:LYS:N	2.39	0.45
1:B:9:GLN:NE2	4:B:401:HOH:O	2.20	0.44
1:D:199:PRO:HD2	1:D:231:ARG:O	2.18	0.43
1:C:136:LYS:HG2	1:C:164:PHE:HB2	1.98	0.43
1:A:239:THR:HG22	1:A:243:HIS:CE1	2.54	0.43
1:A:136:LYS:HG2	1:A:164:PHE:HB2	2.02	0.42
1:A:115:GLU:CD	1:A:152:LYS:HE2	2.41	0.41
1:A:179:GLU:HG3	4:A:582:HOH:O	2.20	0.41
1:D:100:GLU:HG2	1:D:134:ILE:HG12	2.02	0.41
1:D:248:ALA:O	1:D:249:ASP:HB2	2.20	0.41
1:D:197:PHE:CE2	1:D:199:PRO:HD3	2.56	0.41
1:B:239[A]:THR:HG22	1:B:243:HIS:CE1	2.55	0.41
1:A:35:LYS:HE3	1:A:238:LEU:HD23	2.03	0.40
1:D:166:LYS:HE2	1:D:168:SER: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	251/265 (95%)	243 (97%)	8 (3%)	0	100	100
1	B	255/265 (96%)	247 (97%)	8 (3%)	0	100	100
1	C	246/265 (93%)	241 (98%)	5 (2%)	0	100	100
1	D	247/265 (93%)	237 (96%)	10 (4%)	0	100	100
All	All	999/1060 (94%)	968 (97%)	31 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains

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	195/205 (95%)	194 (100%)	1 (0%)	92	87
1	B	197/205 (96%)	195 (99%)	2 (1%)	82	69
1	C	184/205 (90%)	183 (100%)	1 (0%)	92	87
1	D	173/205 (84%)	173 (100%)	0	100	100
All	All	749/820 (91%)	745 (100%)	4 (0%)	92	87

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	THR
1	B	38	THR
1	B	66	ASP
1	C	38	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	243	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

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	248/265 (93%)	-0.23	2 (0%) 87 91	15, 23, 40, 72	0
1	B	248/265 (93%)	-0.24	3 (1%) 81 86	15, 24, 39, 73	0
1	C	246/265 (92%)	-0.05	7 (2%) 56 62	22, 37, 65, 87	0
1	D	248/265 (93%)	1.30	62 (25%) 1 1	37, 66, 91, 111	0
All	All	990/1060 (93%)	0.20	74 (7%) 17 22	15, 32, 83, 111	0

All (74) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	174	VAL	8.0
1	D	176	ALA	7.3
1	B	1	MET	7.1
1	D	178	LEU	6.6
1	D	144	LEU	6.2
1	D	105	PHE	6.2
1	C	3	ASP	5.5
1	C	247	LEU	5.5
1	D	116	THR	5.5
1	D	172	VAL	5.3
1	D	2	SER	5.3
1	D	249	ASP	5.2
1	C	2	SER	5.1
1	A	1	MET	5.0
1	D	148	ALA	4.8
1	D	145	ALA	4.5
1	D	160	ALA	4.4
1	D	79	GLY	4.4
1	D	173	ALA	4.4
1	D	81	ASP	4.2
1	D	122	VAL	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	77	PRO	3.6
1	D	107	TYR	3.5
1	D	119	PHE	3.5
1	D	78	HIS	3.4
1	C	4	LEU	3.4
1	D	177	THR	3.4
1	D	111	MET	3.3
1	D	109	ALA	3.2
1	D	110	LEU	3.2
1	D	169	THR	3.2
1	D	209	ALA	3.2
1	D	149	LEU	3.2
1	D	158	ILE	3.1
1	D	159	ASP	3.1
1	D	113	GLY	3.0
1	C	243	HIS	3.0
1	A	2	SER	3.0
1	B	247	LEU	3.0
1	D	128	ALA	2.8
1	C	244	THR	2.7
1	D	189	SER	2.7
1	D	139	ILE	2.6
1	D	165	ILE	2.6
1	D	124	ALA	2.6
1	C	245	LEU	2.6
1	D	141	SER	2.5
1	D	164	PHE	2.5
1	D	4	LEU	2.5
1	D	210	ALA	2.5
1	D	135	LEU	2.5
1	B	248	ALA	2.4
1	D	236	SER	2.4
1	D	248	ALA	2.4
1	D	202	GLY	2.4
1	D	222	ASP	2.4
1	D	103	VAL	2.3
1	D	154	SER	2.3
1	D	84	ALA	2.3
1	D	147	PRO	2.2
1	D	118	GLY	2.2
1	D	212	PHE	2.2
1	D	216	ALA	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	175	ASN	2.1
1	D	75	ASN	2.1
1	D	182	GLU	2.1
1	D	76	PHE	2.1
1	D	117	VAL	2.1
1	D	155	GLU	2.0
1	D	205[A]	ASP	2.0
1	D	218	ARG	2.0
1	D	22	ASP	2.0
1	D	239	THR	2.0
1	D	170	GLY	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 carbohydrates in this entry.

6.4 Ligands (i)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
3	CL	A	302	1/1	0.99	0.04	-	23,23,23,23	0
3	CL	B	302	1/1	0.99	0.05	-	24,24,24,24	0
2	NA	A	301	1/1	0.97	0.07	-	35,35,35,35	0
3	CL	D	301	1/1	0.92	0.08	-	41,41,41,41	0
3	CL	C	301	1/1	0.96	0.07	-	43,43,43,43	0
2	NA	B	301	1/1	0.98	0.06	-	35,35,35,35	0

6.5 Other polymers (i)

There are no such residues in this entry.