wwPDB X-ray Structure Validation Summary Report (i)

Sep 23, 2020 – 04:08 PM BST

PDB ID	:	6SBA
Title	:	Crystal Structure of mTEAD with a VGL4 Tertiary Structure Mimetic
Authors	:	Adihou, H.; Grossmann, T.N.; Waldmann, H.; Gasper, R.
Deposited on	:	2019-07-19
Resolution	:	1.30 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:

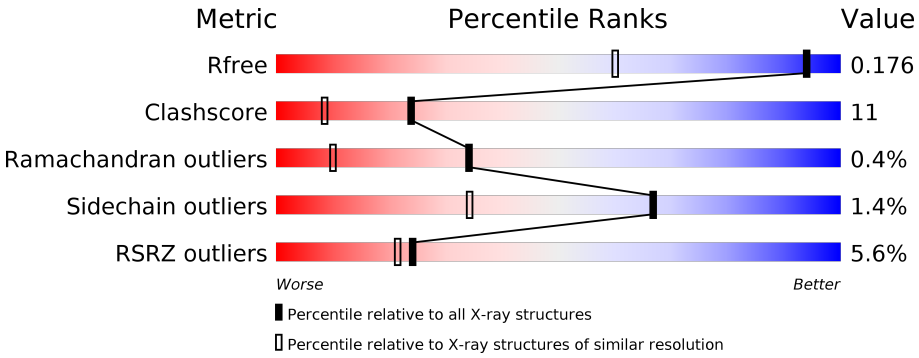
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.6

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 1.30 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	1058 (1.30-1.30)
Clashscore	141614	1101 (1.30-1.30)
Ramachandran outliers	138981	1058 (1.30-1.30)
Sidechain outliers	138945	1058 (1.30-1.30)
RSRZ outliers	127900	1029 (1.30-1.30)

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	221	5% 77% 18%
2	B	20	10% 75% 25%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4202 atoms, of which 1995 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 Transcriptional enhancer factor TEF-3.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	212	Total	C	H	N	O	S	0	22	0
			3692	1230	1807	305	339	11			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	207	GLY	-	expression tag	UNP Q62296
A	208	PRO	-	expression tag	UNP Q62296

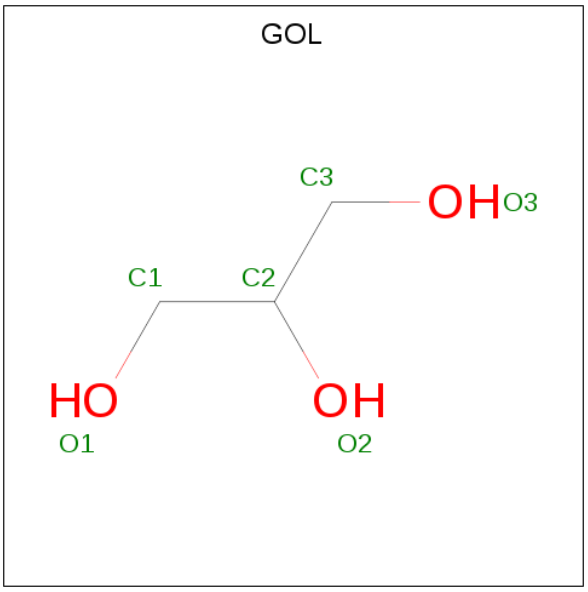
- Molecule 2 is a protein called Vestigial like 4 (Drosophila), isoform CRA_a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	20	Total	C	H	N	O	0	0	0
			304	100	149	26	29			

There is a discrepancy between the modelled and reference sequences:

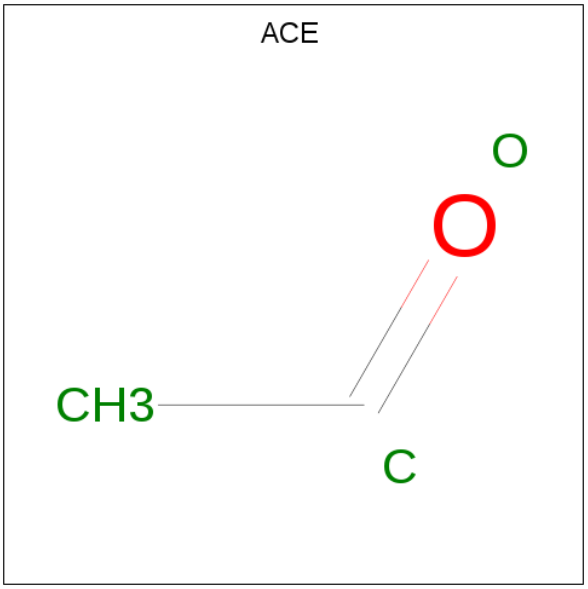
Chain	Residue	Modelled	Actual	Comment	Reference
B	10	GLU	ASP	conflict	UNP Q3TQI9

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 14 3 8 3	0	0
3	A	1	Total C H O 14 3 8 3	0	0
3	A	1	Total C H O 14 3 8 3	0	0
3	A	1	Total C H O 14 3 8 3	0	0
3	A	1	Total C H O 13 3 7 3	0	0

- Molecule 4 is ACETYL GROUP (three-letter code: ACE) (formula: C_2H_4O).



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

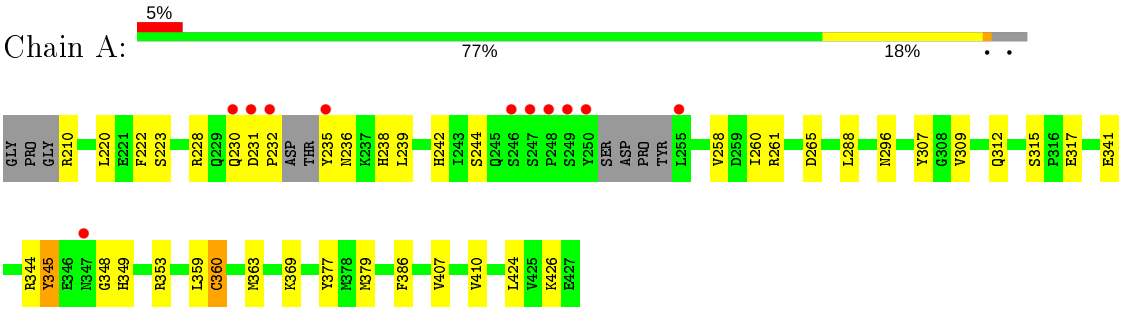
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	127	Total O 127 127	0	0
5	B	7	Total O 7 7	0	0

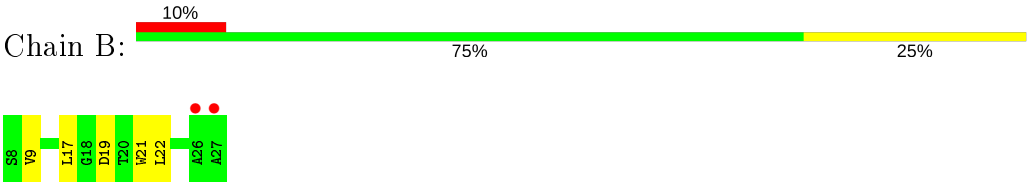
3 Residue-property plots (i)

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: Transcriptional enhancer factor TEF-3



- Molecule 2: Vestigial like 4 (Drosophila), isoform CRA_a



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.50Å 65.11Å 74.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.14 – 1.30 49.14 – 1.30	Depositor EDS
% Data completeness (in resolution range)	97.9 (49.14-1.30) 97.9 (49.14-1.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.16 (at 1.30Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.151 , 0.175 0.152 , 0.176	Depositor DCC
R_{free} test set	3359 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	19.8	Xtriage
Anisotropy	0.304	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 47.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	4202	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.46% 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

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, P1L, ACE

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	1.21	11/1963 (0.6%)	1.02	3/2647 (0.1%)
2	B	1.10	1/158 (0.6%)	1.11	2/211 (0.9%)
All	All	1.20	12/2121 (0.6%)	1.03	5/2858 (0.2%)

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	341	GLU	CD-OE1	13.17	1.40	1.25
1	A	223	SER	CB-OG	-7.38	1.32	1.42
1	A	341	GLU	CD-OE2	6.02	1.32	1.25
1	A	296[A]	ASN	CB-CG	5.82	1.64	1.51
1	A	296[B]	ASN	CB-CG	5.82	1.64	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	19	ASP	CB-CG-OD1	9.50	126.85	118.30
2	B	19	ASP	CB-CG-OD2	-7.45	111.59	118.30
1	A	341	GLU	OE1-CD-OE2	6.92	131.61	123.30
1	A	265	ASP	CB-CG-OD2	-6.47	112.48	118.30
1	A	344	ARG	NE-CZ-NH2	-5.46	117.57	120.30

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	1885	1807	1880	44	0
2	B	155	149	148	2	0
3	A	30	39	40	3	0
4	B	3	0	3	0	0
5	A	127	0	0	4	1
5	B	7	0	0	0	0
All	All	2207	1995	2071	46	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 11.

The worst 5 of 46 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:360[B]:P1L:H171	1:A:363[B]:MET:HE1	1.65	0.78
1:A:349:HIS:ND1	5:A:601:HOH:O	2.19	0.74
1:A:360[B]:P1L:H8C2	1:A:363[B]:MET:SD	2.29	0.73
1:A:232:PRO:HB3	1:A:236:ASN:H	1.55	0.69
1:A:360[B]:P1L:H112	1:A:363[B]:MET:SD	2.32	0.69

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)
5:A:617:HOH:O	5:A:632:HOH:O[4_445]	2.15	0.05

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	226/221 (102%)	220 (97%)	5 (2%)	1 (0%)	34 10
2	B	18/20 (90%)	18 (100%)	0	0	100 100
All	All	244/241 (101%)	238 (98%)	5 (2%)	1 (0%)	34 10

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	230	GLN

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	215/201 (107%)	212 (99%)	3 (1%)	67 34
2	B	14/15 (93%)	13 (93%)	1 (7%)	14 0
All	All	229/216 (106%)	225 (98%)	4 (2%)	67 26

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

Mol	Chain	Res	Type
1	A	210	ARG
1	A	261[A]	ARG
1	A	261[B]	ARG
2	B	22	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	$\#\|Z\|>2$
1	P1L	A	360[B]	-	21,22,23	1.03	1 (4%)	18,23,25	1.89	2 (11%)
1	P1L	A	360[A]	-	21,22,23	0.72	0	18,23,25	0.80	1 (5%)

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
1	P1L	A	360[B]	-	-	10/20/22/24	-
1	P1L	A	360[A]	-	-	9/20/22/24	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	360[B]	P1L	CB-SG	-3.72	1.72	1.81

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	360[B]	P1L	CB-SG-C7	-6.03	92.40	100.84
1	A	360[B]	P1L	O7-C7-C8	-3.87	119.42	123.99
1	A	360[A]	P1L	CB-SG-C7	-2.28	97.66	100.84

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	360[B]	P1L	C8-C7-SG-CB

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	A	360[B]	P1L	C7-C8-C9-C10
1	A	360[A]	P1L	N-CA-CB-SG
1	A	360[A]	P1L	C-CA-CB-SG
1	A	360[A]	P1L	C8-C7-SG-CB

There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	360[B]	P1L	7	0
1	A	360[A]	P1L	5	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

6 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	$\#\|Z\| > 2$
3	GOL	A	502	-	5,5,5	0.80	0	5,5,5	1.88	2 (40%)
3	GOL	A	501	-	5,5,5	0.38	0	5,5,5	0.26	0
3	GOL	A	505	-	5,5,5	0.20	0	5,5,5	0.66	0
3	GOL	A	504	-	5,5,5	0.40	0	5,5,5	1.15	0
3	GOL	A	503	-	5,5,5	0.51	0	5,5,5	0.49	0
4	ACE	B	101	2	1,2,2	1.07	0	1,1,1	0.06	0

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
3	GOL	A	503	-	-	2/4/4/4	-
3	GOL	A	505	-	-	2/4/4/4	-
3	GOL	A	502	-	-	1/4/4/4	-
3	GOL	A	504	-	-	2/4/4/4	-
3	GOL	A	501	-	-	2/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	502	GOL	O2-C2-C3	3.05	122.54	109.12
3	A	502	GOL	O2-C2-C1	-2.32	98.90	109.12

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	501	GOL	O1-C1-C2-C3
3	A	504	GOL	O1-C1-C2-C3
3	A	503	GOL	O1-C1-C2-C3
3	A	505	GOL	O1-C1-C2-O2
3	A	505	GOL	O1-C1-C2-C3

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	GOL	1	0
3	A	505	GOL	1	0
3	A	504	GOL	1	0

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	211/221 (95%)	0.07	11 (5%) 27 24	17, 25, 63, 104	0
2	B	20/20 (100%)	0.28	2 (10%) 7 5	20, 34, 74, 88	0
All	All	231/241 (95%)	0.09	13 (5%) 24 21	17, 25, 64, 104	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	231	ASP	7.6
1	A	232	PRO	6.3
1	A	248	PRO	6.0
1	A	250	TYR	5.2
2	B	27	ALA	5.0

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

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
1	P1L	A	360[B]	23/24	0.91	0.15	21,58,97,104	20
1	P1L	A	360[A]	23/24	0.91	0.15	21,58,97,104	20

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
3	GOL	A	501	6/6	-0.18	0.38	116,139,154,155	0
3	GOL	A	505	6/6	0.25	0.29	78,97,123,123	0
3	GOL	A	504	6/6	0.28	0.21	74,90,108,108	0
3	GOL	A	502	6/6	0.71	0.17	34,66,88,88	0
3	GOL	A	503	6/6	0.81	0.13	53,64,78,94	0
4	ACE	B	101	3/3	0.88	0.10	40,40,42,58	0

6.5 Other polymers

There are no such residues in this entry.