Full wwPDB X-ray Structure Validation Report (i)

Sep 6, 2023 – 04:05 pm BST

PDB ID	:	8P5P
Title	:	Structure of TECPR1 N-terminal DysF domain
Authors	:	Boyle, K.B.; Elliott, P.R.; Randow, F.
Deposited on	:	2023-05-24
Resolution	:	1.90 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 (i) 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 (1)) were used in the production of this report:

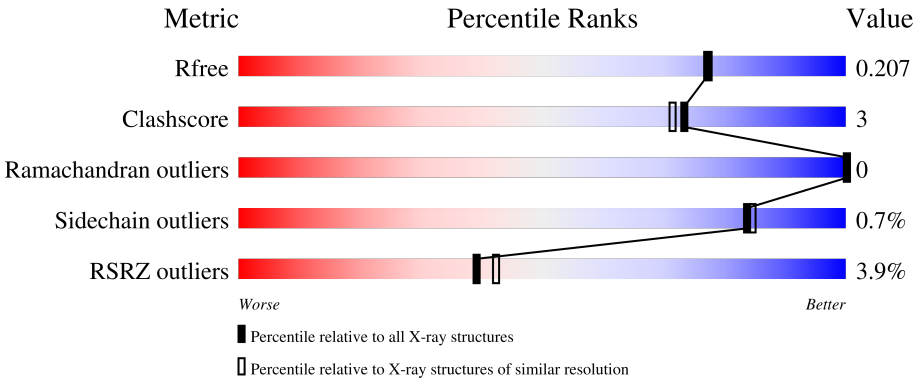
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	116	2% 85% 7% 8%
1	B	116	3% 80% 9% • 9%
1	C	116	6% 80% 10% • 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5736 atoms, of which 2633 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.

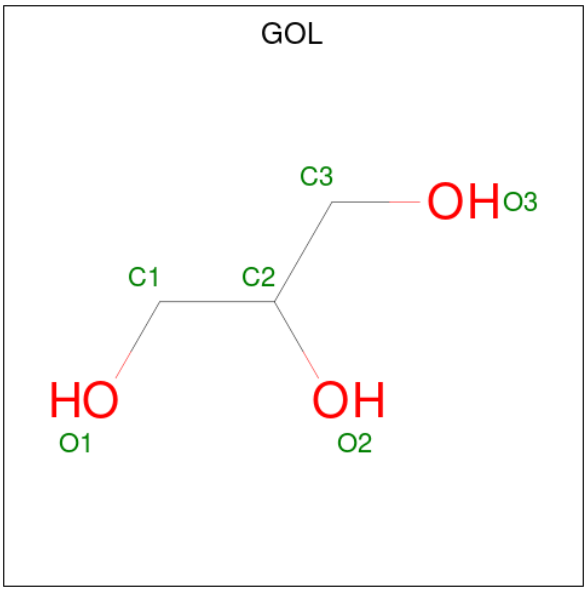
- Molecule 1 is a protein called Tectonin beta-propeller repeat-containing protein 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	107	Total	C	H	N	O	S	0	0	0
			1821	601	884	172	161	3			
1	B	105	Total	C	H	N	O	S	0	0	0
			1786	590	864	170	159	3			
1	C	106	Total	C	H	N	O	S	0	0	0
			1798	595	870	170	160	3			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	58	GLY	-	expression tag	UNP Q7Z6L1
A	59	ALA	-	expression tag	UNP Q7Z6L1
B	58	GLY	-	expression tag	UNP Q7Z6L1
B	59	ALA	-	expression tag	UNP Q7Z6L1
C	58	GLY	-	expression tag	UNP Q7Z6L1
C	59	ALA	-	expression tag	UNP Q7Z6L1

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

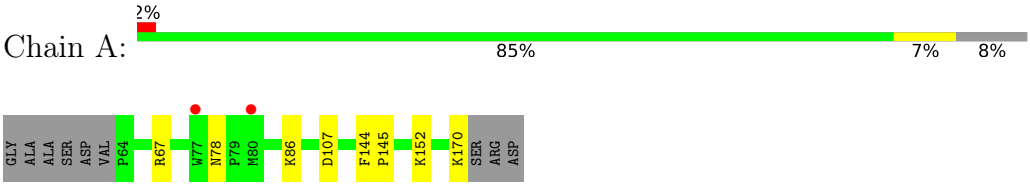
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	96	Total O 96 96	0	0
3	B	106	Total O 106 106	0	0
3	C	102	Total O 102 102	0	0

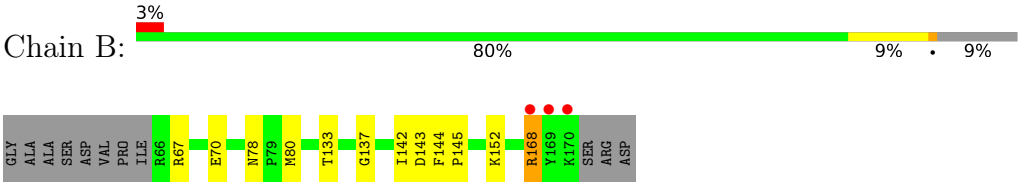
3 Residue-property plots (i)

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.

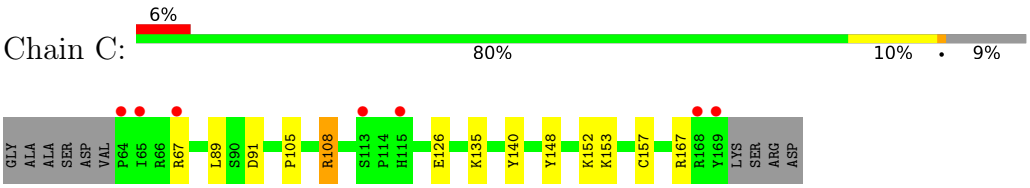
• Molecule 1: Tectonin beta-propeller repeat-containing protein 1



• Molecule 1: Tectonin beta-propeller repeat-containing protein 1



• Molecule 1: Tectonin beta-propeller repeat-containing protein 1



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	75.72Å 75.72Å 122.86Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.83 – 1.90 44.83 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (44.83-1.90) 99.8 (44.83-1.90)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.82 (at 1.89Å)	Xtriage
Refinement program	PHENIX (1.19.2_4158: ???)	Depositor
R, R_{free}	0.166 , 0.208 0.165 , 0.207	Depositor DCC
R_{free} test set	1601 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	24.8	Xtriage
Anisotropy	0.260	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 51.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.067 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5736	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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: MLY, GOL

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	0.57	0/928	0.71	0/1262
1	B	0.53	0/912	0.70	0/1240
1	C	0.55	0/919	0.74	0/1251
All	All	0.55	0/2759	0.72	0/3753

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	B	0	1
1	C	0	2
All	All	0	5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	67	ARG	Sidechain
1	A	86	MLY	Mainchain
1	B	168	ARG	Sidechain
1	C	108	ARG	Sidechain
1	C	153	MLY	Mainchain

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	937	884	883	4	0
1	B	922	864	863	8	0
1	C	928	870	869	9	0
2	C	12	15	15	2	0
3	A	96	0	0	1	0
3	B	106	0	0	2	0
3	C	102	0	0	0	0
All	All	3103	2633	2630	18	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 (18) 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:B:80:MET:CE	1:C:89:LEU:HD11	2.35	0.57
1:B:78:ASN:HB2	3:B:204:HOH:O	2.06	0.55
1:C:126:GLU:OE2	1:C:135:LYS:HD2	2.08	0.52
1:C:91:ASP:O	2:C:202:GOL:H2	2.10	0.52
1:B:70:GLU:OE1	3:B:201:HOH:O	2.19	0.50
1:A:78:ASN:HB2	3:A:203:HOH:O	2.13	0.49
1:B:80:MET:HE2	1:C:89:LEU:HD11	1.96	0.48
1:B:142:ILE:HG13	1:B:143:ASP:N	2.28	0.47
1:B:152:MLY:HH21	1:C:157:CYS:SG	2.56	0.46
1:B:144:PHE:N	1:B:145:PRO:CD	2.81	0.44
1:C:105:PRO:HG2	1:C:108:ARG:HB2	1.99	0.44
1:A:170:LYS:HE3	1:A:170:LYS:HB2	1.88	0.43
1:C:67:ARG:NH1	1:C:167:ARG:HH22	2.17	0.43
1:A:107:ASP:OD1	1:A:107:ASP:N	2.51	0.42
1:A:144:PHE:N	1:A:145:PRO:CD	2.83	0.42
1:B:133:THR:OG1	1:B:137:GLY:HA2	2.20	0.41
1:C:152:MLY:HH12	2:C:201:GOL:O2	2.22	0.40
1:C:140:TYR:O	1:C:148:TYR:HA	2.20	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	101/116 (87%)	99 (98%)	2 (2%)	0	100	100
1	B	99/116 (85%)	98 (99%)	1 (1%)	0	100	100
1	C	100/116 (86%)	97 (97%)	3 (3%)	0	100	100
All	All	300/348 (86%)	294 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	91/97 (94%)	91 (100%)	0	100	100
1	B	89/97 (92%)	87 (98%)	2 (2%)	52	47
1	C	90/97 (93%)	90 (100%)	0	100	100
All	All	270/291 (93%)	268 (99%)	2 (1%)	84	84

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	67	ARG
1	B	168	ARG

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

Mol	Chain	Res	Type
1	C	103	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

12 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	MLY	B	152	1	9,10,11	0.69	0	6,11,13	0.84	0
1	MLY	A	86	1	9,10,11	0.53	0	6,11,13	1.04	0
1	MLY	A	153	1	9,10,11	0.54	0	6,11,13	0.95	0
1	MLY	A	152	1	9,10,11	0.88	1 (11%)	6,11,13	0.74	0
1	MLY	A	162	1	9,10,11	0.55	0	6,11,13	0.65	0
1	MLY	C	162	1	9,10,11	0.81	0	6,11,13	0.54	0
1	MLY	C	86	1	9,10,11	0.73	0	6,11,13	0.53	0
1	MLY	C	152	1	9,10,11	0.51	0	6,11,13	0.67	0
1	MLY	B	86	1	9,10,11	0.62	0	6,11,13	0.66	0
1	MLY	B	153	1	9,10,11	0.61	0	6,11,13	0.97	0
1	MLY	C	153	1	9,10,11	0.60	0	6,11,13	0.63	0
1	MLY	B	162	1	9,10,11	0.80	0	6,11,13	0.64	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
1	MLY	B	152	1	-	0/8/9/11	-
1	MLY	A	86	1	-	1/8/9/11	-
1	MLY	A	153	1	-	3/8/9/11	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	MLY	A	152	1	-	6/8/9/11	-
1	MLY	A	162	1	-	0/8/9/11	-
1	MLY	C	162	1	-	2/8/9/11	-
1	MLY	C	86	1	-	4/8/9/11	-
1	MLY	C	152	1	-	0/8/9/11	-
1	MLY	B	86	1	-	3/8/9/11	-
1	MLY	B	153	1	-	2/8/9/11	-
1	MLY	C	153	1	-	0/8/9/11	-
1	MLY	B	162	1	-	1/8/9/11	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	152	MLY	CE-NZ	2.26	1.54	1.46

There are no bond angle outliers.

There are no chirality outliers.

All (22) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	86	MLY	O-C-CA-CB
1	B	86	MLY	O-C-CA-CB
1	C	86	MLY	O-C-CA-CB
1	A	153	MLY	CD-CE-NZ-CH1
1	A	153	MLY	CD-CE-NZ-CH2
1	C	162	MLY	CE-CD-CG-CB
1	C	86	MLY	CG-CD-CE-NZ
1	A	152	MLY	CD-CE-NZ-CH1
1	C	162	MLY	CD-CE-NZ-CH2
1	C	86	MLY	CA-CB-CG-CD
1	A	152	MLY	CE-CD-CG-CB
1	A	152	MLY	CA-CB-CG-CD
1	A	153	MLY	CG-CD-CE-NZ
1	A	152	MLY	N-CA-CB-CG
1	A	152	MLY	C-CA-CB-CG
1	B	153	MLY	CA-CB-CG-CD
1	B	86	MLY	CE-CD-CG-CB
1	B	86	MLY	CG-CD-CE-NZ
1	B	153	MLY	CE-CD-CG-CB
1	C	86	MLY	CE-CD-CG-CB

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	B	162	MLY	CA-CB-CG-CD
1	A	152	MLY	CG-CD-CE-NZ

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	152	MLY	1	0
1	C	152	MLY	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

2 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	GOL	C	201	-	5,5,5	0.69	0	5,5,5	0.94	0
2	GOL	C	202	-	5,5,5	1.10	1 (20%)	5,5,5	1.05	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
2	GOL	C	201	-	-	2/4/4/4	-
2	GOL	C	202	-	-	2/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	202	GOL	O1-C1	-2.06	1.33	1.42

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	201	GOL	O1-C1-C2-C3
2	C	202	GOL	O1-C1-C2-C3
2	C	201	GOL	O1-C1-C2-O2
2	C	202	GOL	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	201	GOL	1	0
2	C	202	GOL	1	0

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

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	103/116 (88%)	-0.25	2 (1%)	66	69	21, 31, 51, 67	0
1	B	101/116 (87%)	-0.20	3 (2%)	50	53	20, 31, 50, 78	0
1	C	102/116 (87%)	0.18	7 (6%)	16	19	14, 26, 71, 90	0
All	All	306/348 (87%)	-0.09	12 (3%)	39	42	14, 31, 60, 90	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	64	PRO	4.2
1	C	65	ILE	3.8
1	C	169	TYR	3.6
1	B	169	TYR	3.3
1	C	168	ARG	3.3
1	B	170	LYS	2.9
1	A	80	MET	2.5
1	C	113	SER	2.2
1	B	168	ARG	2.2
1	C	115	HIS	2.1
1	C	67	ARG	2.0
1	A	77	TRP	2.0

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	MLY	A	152	11/12	0.89	0.21	36,49,69,72	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	MLY	A	153	11/12	0.92	0.15	39,50,80,80	0
1	MLY	A	86	11/12	0.94	0.12	30,46,61,61	0
1	MLY	C	86	11/12	0.95	0.21	21,48,68,68	0
1	MLY	B	152	11/12	0.96	0.12	26,35,61,61	0
1	MLY	B	86	11/12	0.96	0.11	27,37,41,45	0
1	MLY	B	153	11/12	0.97	0.11	21,34,48,48	0
1	MLY	C	152	11/12	0.97	0.10	21,28,37,37	0
1	MLY	C	153	11/12	0.97	0.10	22,33,48,48	0
1	MLY	C	162	11/12	0.97	0.12	18,40,68,68	0
1	MLY	A	162	11/12	0.98	0.09	19,35,53,53	0
1	MLY	B	162	11/12	0.98	0.11	21,32,44,44	0

6.3 Carbohydrates (i)

There are no monosaccharides 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. 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	GOL	C	201	6/6	0.90	0.23	35,46,56,57	0
2	GOL	C	202	6/6	0.93	0.10	34,42,48,55	0

6.5 Other polymers (i)

There are no such residues in this entry.