WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Jun 29, 2025 – 08:13 am BST

PDB ID	:	9FI0 / pdb_00009fi0
Title	:	Bacteroides ovatus polysaccharide lyase family 38 (BoPL38) mutant D108N
Authors	:	Tandrup, T.; Wilkens, C.
Deposited on	:	2024-05-28
Resolution	:	1.86 Å(reported)

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

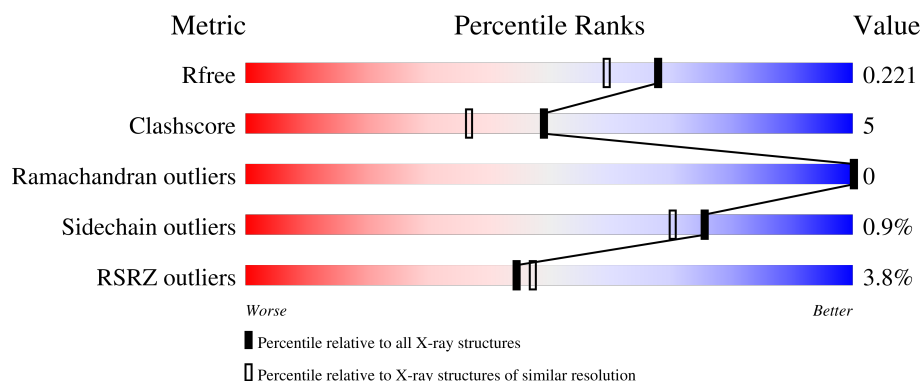|		
MolProbity	:	4-5-2 with Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.44

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
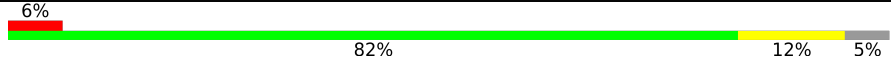*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.86 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	Percentile Ranks	Value
Rfree		0.221
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		0.9%
RSRZ outliers		3.8%

Worse ← → Better

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	3097 (1.86-1.86)
Clashscore	180529	3359 (1.86-1.86)
Ramachandran outliers	177936	3335 (1.86-1.86)
Sidechain outliers	177891	3335 (1.86-1.86)
RSRZ outliers	164620	3097 (1.86-1.86)

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	404	6% / 82% 12% 5%
1	B	404	.% / 82% 11% 6%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6677 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 Alginate lyase family protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	382	Total	C	N	O	S	0	13	0
			3135	2014	530	576	15			
1	B	380	Total	C	N	O	S	0	9	0
			3112	1997	528	573	14			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP A0A5M5BWR5
A	2	GLY	-	expression tag	UNP A0A5M5BWR5
A	3	SER	-	expression tag	UNP A0A5M5BWR5
A	4	SER	-	expression tag	UNP A0A5M5BWR5
A	5	HIS	-	expression tag	UNP A0A5M5BWR5
A	6	HIS	-	expression tag	UNP A0A5M5BWR5
A	7	HIS	-	expression tag	UNP A0A5M5BWR5
A	8	HIS	-	expression tag	UNP A0A5M5BWR5
A	9	HIS	-	expression tag	UNP A0A5M5BWR5
A	10	HIS	-	expression tag	UNP A0A5M5BWR5
A	11	SER	-	expression tag	UNP A0A5M5BWR5
A	12	SER	-	expression tag	UNP A0A5M5BWR5
A	13	GLY	-	expression tag	UNP A0A5M5BWR5
A	14	LEU	-	expression tag	UNP A0A5M5BWR5
A	15	VAL	-	expression tag	UNP A0A5M5BWR5
A	16	PRO	-	expression tag	UNP A0A5M5BWR5
A	17	ARG	-	expression tag	UNP A0A5M5BWR5
A	18	GLY	-	expression tag	UNP A0A5M5BWR5
A	19	SER	-	expression tag	UNP A0A5M5BWR5
A	20	HIS	-	expression tag	UNP A0A5M5BWR5
A	21	MET	-	expression tag	UNP A0A5M5BWR5
A	22	ALA	-	expression tag	UNP A0A5M5BWR5
A	23	SER	-	expression tag	UNP A0A5M5BWR5
A	108	ASN	ASP	engineered mutation	UNP A0A5M5BWR5
B	1	MET	-	initiating methionine	UNP A0A5M5BWR5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2	GLY	-	expression tag	UNP A0A5M5BWR5
B	3	SER	-	expression tag	UNP A0A5M5BWR5
B	4	SER	-	expression tag	UNP A0A5M5BWR5
B	5	HIS	-	expression tag	UNP A0A5M5BWR5
B	6	HIS	-	expression tag	UNP A0A5M5BWR5
B	7	HIS	-	expression tag	UNP A0A5M5BWR5
B	8	HIS	-	expression tag	UNP A0A5M5BWR5
B	9	HIS	-	expression tag	UNP A0A5M5BWR5
B	10	HIS	-	expression tag	UNP A0A5M5BWR5
B	11	SER	-	expression tag	UNP A0A5M5BWR5
B	12	SER	-	expression tag	UNP A0A5M5BWR5
B	13	GLY	-	expression tag	UNP A0A5M5BWR5
B	14	LEU	-	expression tag	UNP A0A5M5BWR5
B	15	VAL	-	expression tag	UNP A0A5M5BWR5
B	16	PRO	-	expression tag	UNP A0A5M5BWR5
B	17	ARG	-	expression tag	UNP A0A5M5BWR5
B	18	GLY	-	expression tag	UNP A0A5M5BWR5
B	19	SER	-	expression tag	UNP A0A5M5BWR5
B	20	HIS	-	expression tag	UNP A0A5M5BWR5
B	21	MET	-	expression tag	UNP A0A5M5BWR5
B	22	ALA	-	expression tag	UNP A0A5M5BWR5
B	23	SER	-	expression tag	UNP A0A5M5BWR5
B	108	ASN	ASP	engineered mutation	UNP A0A5M5BWR5

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	230	Total 230	O 230	0	0
2	B	200	Total 200	O 200	0	0

3 Residue-property plots ⓘ

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: Alginate lyase family protein

Chain A:

• Molecule 1: Alginate lyase family protein

Chain B:
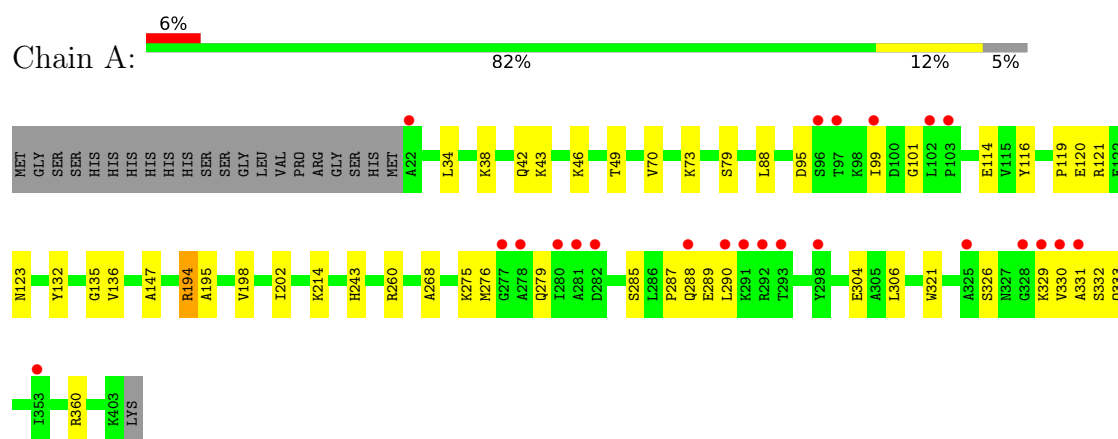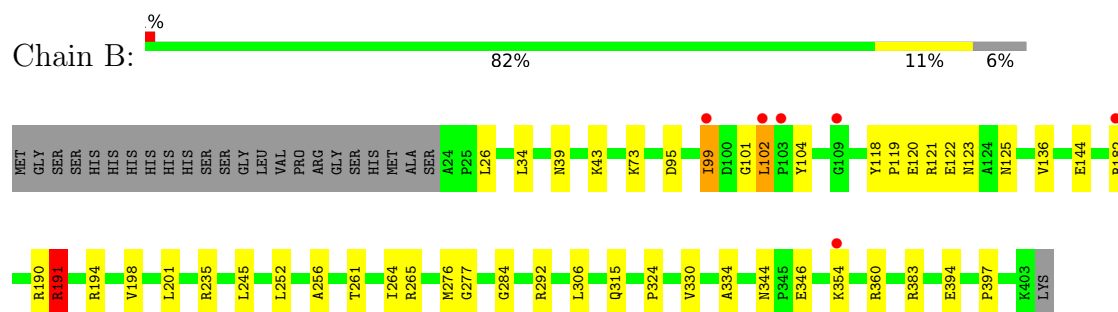
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	54.74Å 104.49Å 151.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.49 – 1.86 48.49 – 1.86	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.49-1.86) 99.9 (48.49-1.86)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.22 (at 1.86Å)	Xtriage
Refinement program	PHENIX 1.21_5207	Depositor
R, R_{free}	0.194 , 0.221 0.194 , 0.221	Depositor DCC
R_{free} test set	3609 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	32.1	Xtriage
Anisotropy	0.064	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 34.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6677	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.35 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.7342e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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	0.32	0/3252	0.51	0/4402
1	B	0.29	0/3208	0.48	1/4344 (0.0%)
All	All	0.30	0/6460	0.50	1/8746 (0.0%)

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	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	191	ARG	CG-CD-NE	5.92	125.02	112.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	194	ARG	Sidechain
1	B	191	ARG	Sidechain

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	3135	0	3167	34	0
1	B	3112	0	3123	34	0
2	A	230	0	0	11	0
2	B	200	0	0	5	0
All	All	6677	0	6290	67	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 5.

All (67) 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:194:ARG:NH2	2:A:501:HOH:O	1.89	1.04
1:A:116:TYR:HA	1:A:121[B]:ARG:HH22	1.43	0.81
1:B:182:ARG:HH22	1:B:235:ARG:HH21	1.35	0.72
1:A:260[B]:ARG:NH2	2:A:506:HOH:O	2.24	0.68
1:A:288:GLN:NE2	2:A:507:HOH:O	2.26	0.68
1:A:268:ALA:O	2:A:502:HOH:O	2.11	0.68
1:A:135:GLY:HA2	1:A:202[B]:ILE:HD11	1.76	0.66
1:A:42[A]:GLN:NE2	2:A:510:HOH:O	2.33	0.61
1:A:38:LYS:NZ	1:A:42[A]:GLN:OE1	2.35	0.60
1:B:34[A]:LEU:HD21	1:B:397:PRO:HG3	1.83	0.60
1:B:292:ARG:NH2	2:B:510:HOH:O	2.35	0.59
1:A:79:SER:O	2:A:503:HOH:O	2.17	0.58
1:B:118:TYR:OH	2:B:501:HOH:O	2.17	0.58
1:A:214:LYS:NZ	2:A:512:HOH:O	2.37	0.57
1:B:125:ASN:HA	1:B:191:ARG:HG2	1.86	0.56
1:B:125:ASN:OD1	1:B:191:ARG:NH1	2.40	0.55
1:A:99:ILE:HG22	1:B:99:ILE:HD11	1.89	0.54
1:B:191:ARG:HD2	1:B:194[A]:ARG:NH1	2.23	0.53
1:A:321:TRP:O	1:A:330:VAL:HG12	2.09	0.53
1:B:198[A]:VAL:HG13	1:B:201:LEU:HD12	1.91	0.52
1:A:147:ALA:HB1	1:A:202[B]:ILE:HG23	1.91	0.52
1:A:275:LYS:O	2:A:504:HOH:O	2.19	0.51
1:B:144:GLU:OE1	2:B:503:HOH:O	2.19	0.51
1:A:243:HIS:HD2	2:A:697:HOH:O	1.94	0.50
1:B:245:LEU:HD22	1:B:306:LEU:HD21	1.94	0.50
1:B:26:LEU:HA	1:B:315:GLN:OE1	2.12	0.49
1:A:88:LEU:HD22	1:A:114:GLU:HB3	1.93	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:326:SER:O	2:A:505:HOH:O	2.20	0.49
1:A:304:GLU:OE2	1:A:360:ARG:HD2	2.12	0.49
1:B:344:ASN:HB3	1:B:346:GLU:OE2	2.14	0.48
1:B:26:LEU:HG	1:B:256:ALA:HB1	1.95	0.47
1:B:34[B]:LEU:HD11	1:B:136:VAL:HG13	1.98	0.46
1:A:132:TYR:CD1	1:A:198[A]:VAL:HG21	2.51	0.45
1:A:195:ALA:O	1:A:198[B]:VAL:HG12	2.16	0.45
1:B:120:GLU:HA	1:B:123:ASN:HB2	1.99	0.45
1:A:46:LYS:HA	1:A:49[B]:THR:HG22	1.99	0.45
1:A:43[B]:LYS:HE2	1:A:43[B]:LYS:HB3	1.66	0.44
1:A:120:GLU:HA	1:A:123:ASN:HB2	2.00	0.44
1:A:329:LYS:HA	1:A:333:GLN:OE1	2.18	0.44
1:B:261:THR:O	1:B:265[B]:ARG:HG3	2.18	0.44
1:B:360:ARG:NH1	1:B:394:GLU:OE1	2.50	0.44
1:B:383:ARG:NH2	2:B:504:HOH:O	2.22	0.44
1:A:330:VAL:C	1:A:332:SER:H	2.24	0.44
1:A:279:GLN:O	1:A:287:PRO:HD2	2.18	0.44
1:A:34:LEU:HD11	1:A:136:VAL:HG13	2.00	0.44
1:B:235:ARG:HD2	2:B:677:HOH:O	2.18	0.44
1:B:102:LEU:O	1:B:104:TYR:N	2.50	0.44
1:B:73:LYS:HB3	1:B:119:PRO:HD2	2.00	0.43
1:A:132:TYR:HB2	1:A:194:ARG:HH21	1.83	0.43
1:B:95:ASP:O	1:B:101:GLY:HA2	2.18	0.43
1:A:38:LYS:O	1:A:42[A]:GLN:HG3	2.19	0.43
1:B:284:GLY:O	1:B:334:ALA:HA	2.19	0.43
1:B:252:LEU:HD22	1:B:264:ILE:HG23	2.00	0.42
1:A:285:SER:HB2	1:A:290:LEU:HD11	2.00	0.42
1:B:276[B]:MET:HE1	1:B:330:VAL:HA	2.02	0.41
1:A:276:MET:HG2	1:A:331:ALA:HB2	2.02	0.41
1:B:324:PRO:HB3	1:B:330:VAL:HG22	2.02	0.41
1:B:34[A]:LEU:HD23	1:B:34[A]:LEU:HA	1.86	0.41
1:B:354:LYS:HD3	1:B:354:LYS:HA	1.78	0.41
1:A:279:GLN:HG3	2:A:504:HOH:O	2.21	0.41
1:A:73:LYS:HB3	1:A:119:PRO:HD2	2.03	0.40
1:B:190:ARG:O	1:B:194[A]:ARG:HG3	2.21	0.40
1:B:121:ARG:HB3	1:B:122:GLU:OE1	2.21	0.40
1:B:276[B]:MET:HG3	1:B:277:GLY:N	2.36	0.40
1:B:39:ASN:HD21	1:B:43:LYS:NZ	2.18	0.40
1:A:95:ASP:O	1:A:101:GLY:HA2	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	393/404 (97%)	386 (98%)	7 (2%)	0	100	100
1	B	387/404 (96%)	380 (98%)	7 (2%)	0	100	100
All	All	780/808 (96%)	766 (98%)	14 (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	338/344 (98%)	334 (99%)	4 (1%)	67	59
1	B	333/344 (97%)	330 (99%)	3 (1%)	75	70
All	All	671/688 (98%)	664 (99%)	7 (1%)	75	67

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	70[A]	VAL
1	A	70[B]	VAL
1	A	289	GLU
1	A	306	LEU
1	B	99	ILE
1	B	102	LEU
1	B	191	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	108	ASN
1	A	232	GLN
1	A	270	GLN
1	A	288	GLN
1	B	39	ASN
1	B	238	HIS
1	B	270	GLN

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 oligosaccharides 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, 95th 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	382/404 (94%)	0.16	23 (6%)	29	31	13, 30, 59, 89	13 (3%)
1	B	380/404 (94%)	0.11	6 (1%)	70	73	12, 33, 60, 102	9 (2%)
All	All	762/808 (94%)	0.14	29 (3%)	44	47	12, 32, 60, 102	22 (2%)

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	102	LEU	5.5
1	B	102	LEU	5.2
1	A	281	ALA	4.7
1	A	330	VAL	4.6
1	B	99	ILE	4.3
1	A	277	GLY	4.2
1	A	97	THR	4.1
1	A	22	ALA	3.6
1	A	328	GLY	3.5
1	A	280	ILE	3.3
1	B	103	PRO	3.2
1	A	99	ILE	3.2
1	A	353	ILE	3.2
1	A	331	ALA	3.1
1	A	292	ARG	3.0
1	A	329	LYS	2.9
1	A	282	ASP	2.8
1	B	109	GLY	2.7
1	A	103	PRO	2.5
1	B	182	ARG	2.5
1	A	288	GLN	2.4
1	A	325	ALA	2.4
1	A	278	ALA	2.3
1	B	354	LYS	2.3

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	298	TYR	2.3
1	A	96	SER	2.2
1	A	293	THR	2.1
1	A	290	LEU	2.0
1	A	291	LYS	2.0

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 oligosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.