wwPDB X-ray Structure Validation Summary Report (i)

Jan 6, 2024 – 10:54 pm GMT

PDB ID	:	5NXN
Title	:	Structure of a L5-deletion mutant of Providencia stuartii Omp-Pst1
Authors	:	El Khatib, M.; Colletier, J.P.
Deposited on	:	2017-05-10
Resolution	:	3.12 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 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:

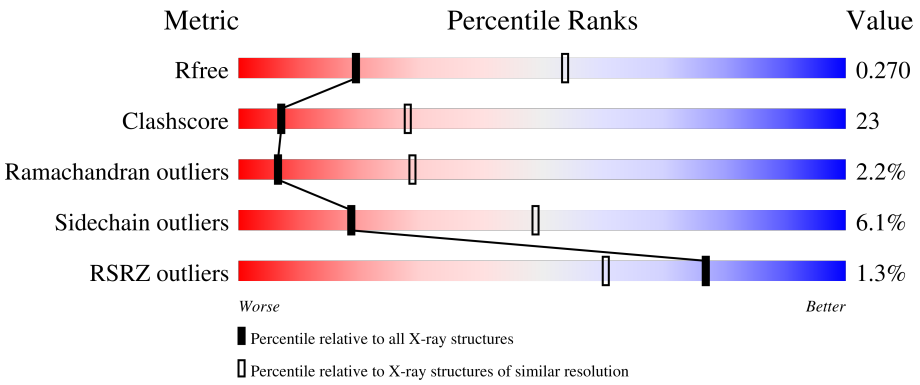
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.12 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	1292 (3.14-3.10)
Clashscore	141614	1389 (3.14-3.10)
Ramachandran outliers	138981	1337 (3.14-3.10)
Sidechain outliers	138945	1337 (3.14-3.10)
RSRZ outliers	127900	1260 (3.14-3.10)

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	342	.% 56% 39% 5%
1	B	342	2% 66% 30% . .
1	C	342	.% 63% 33% .

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8654 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 Porin 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	342	2711	1698	465	543	5	0	4	0
1	B	342	2705	1696	462	542	5	0	3	0
1	C	342	2717	1702	466	544	5	0	5	0

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	VAL	deletion	UNP E3U904
A	?	-	VAL	deletion	UNP E3U904
A	?	-	THR	deletion	UNP E3U904
A	?	-	SER	deletion	UNP E3U904
A	?	-	GLU	deletion	UNP E3U904
A	?	-	GLY	deletion	UNP E3U904
A	?	-	ASP	deletion	UNP E3U904
A	?	-	SER	deletion	UNP E3U904
A	?	-	TYR	deletion	UNP E3U904
A	?	-	TYR	deletion	UNP E3U904
A	293	GLY	ASN	conflict	UNP E3U904
B	?	-	VAL	deletion	UNP E3U904
B	?	-	VAL	deletion	UNP E3U904
B	?	-	THR	deletion	UNP E3U904
B	?	-	SER	deletion	UNP E3U904
B	?	-	GLU	deletion	UNP E3U904
B	?	-	GLY	deletion	UNP E3U904
B	?	-	ASP	deletion	UNP E3U904
B	?	-	SER	deletion	UNP E3U904
B	?	-	TYR	deletion	UNP E3U904
B	?	-	TYR	deletion	UNP E3U904
B	293	GLY	ASN	conflict	UNP E3U904
C	?	-	VAL	deletion	UNP E3U904

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	VAL	deletion	UNP E3U904
C	?	-	THR	deletion	UNP E3U904
C	?	-	SER	deletion	UNP E3U904
C	?	-	GLU	deletion	UNP E3U904
C	?	-	GLY	deletion	UNP E3U904
C	?	-	ASP	deletion	UNP E3U904
C	?	-	SER	deletion	UNP E3U904
C	?	-	TYR	deletion	UNP E3U904
C	?	-	TYR	deletion	UNP E3U904
C	293	GLY	ASN	conflict	UNP E3U904

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
2	A	1	1	1	0	0

- Molecule 3 is water.

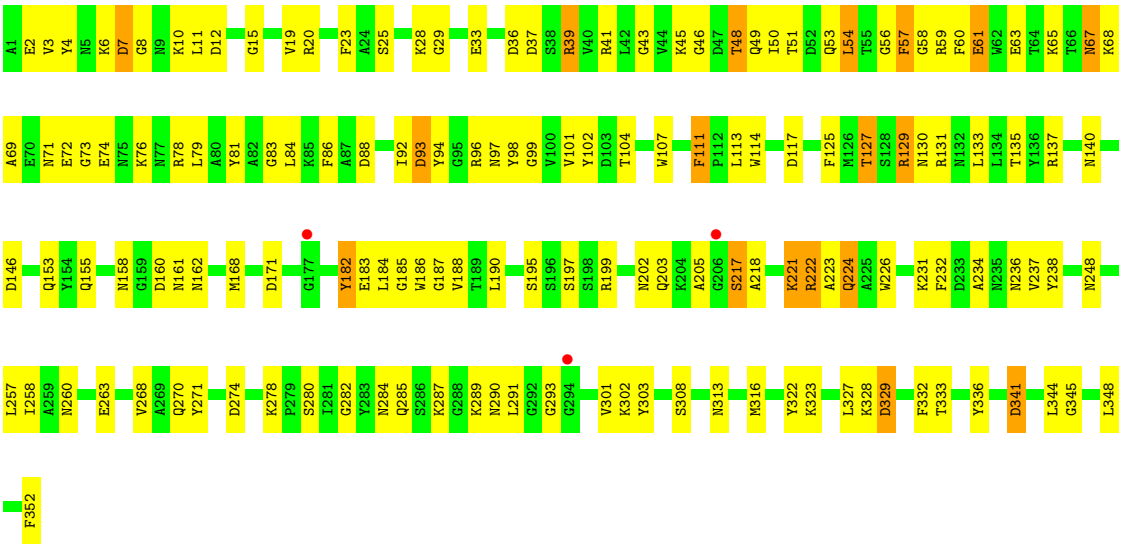
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	203	203	203	0	0
			Total	O		
3	B	168	168	168	0	0
			Total	O		
3	C	149	149	149	0	0

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.

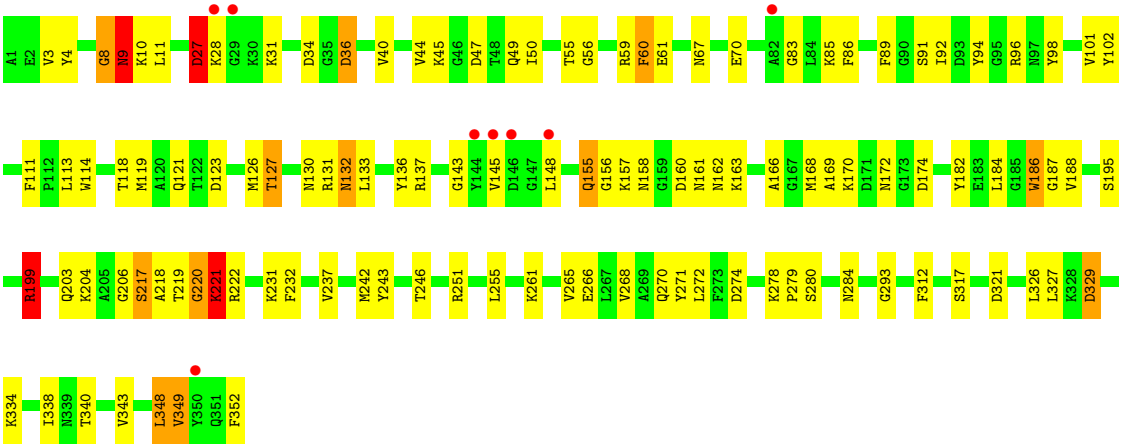
• Molecule 1: Porin 1

Chain A:



• Molecule 1: Porin 1

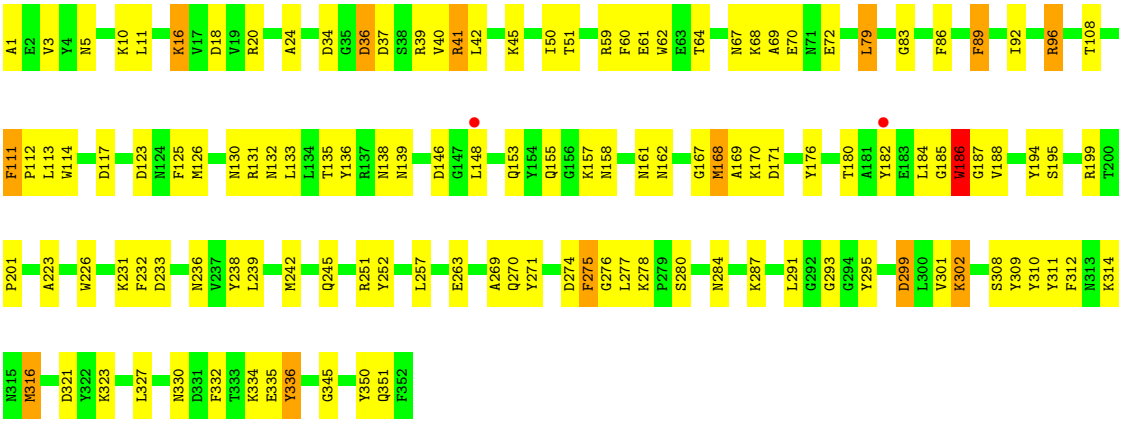
Chain B:



• Molecule 1: Porin 1

Chain C:

%
63%
33%


A1 E2 V3 Y4 N5 K10 L11 K16 V17 D18 V19 R20 A24 D34 G35 D36 D37 S38 R39 V40 R41 L42 K45 I50 T51 R59 F60 E61 W62 E63 T64 N67 K68 A69 E70 N71 E72 L79 G83 F86 F89 I92 R96 T108
F111 P112 L113 W114 D117 D123 N124 F125 M126 N130 R131 N132 L133 L134 T135 Y136 R137 N138 N139 D146 G147 L148 Q153 Y154 Q155 G156 K157 N158 N161 N162 G167 M168 A169 K170 D171 Y176 T180 A181 Y182 E183 L184 G185 W186 G187 V188 Y194 S195 R199 T200
P201 A223 W226 K231 F232 D233 N236 V237 Y238 L239 M242 Q245 R251 Y252 L257 E263 A269 Q270 Y271 D274 F275 G276 L277 K278 P279 S280 N284 K287 L291 G292 G293 G294 Y295 D299 L300 V301 K302 S308 Y309 Y310 Y311 F312 N313 K314
N315 M316 D321 Y322 K323 L327 N330 D331 F332 T333 K334 E335 Y336 G345 Y350 Q351 F352

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	126.32Å 158.99Å 101.33Å 90.00° 124.65° 90.00°	Depositor
Resolution (Å)	50.00 – 3.12 49.55 – 3.12	Depositor EDS
% Data completeness (in resolution range)	98.7 (50.00-3.12) 98.7 (49.55-3.12)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.41 (at 3.12Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.246 , 0.268 0.250 , 0.270	Depositor DCC
R_{free} test set	1482 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	76.4	Xtriage
Anisotropy	0.715	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 44.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	8654	wwPDB-VP
Average B, all atoms (Å^2)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.53% 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 (i)

5.1 Standard geometry (i)

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

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	0.95	3/2766 (0.1%)	1.08	13/3729 (0.3%)
1	B	1.00	3/2761 (0.1%)	1.14	14/3724 (0.4%)
1	C	0.95	0/2772	1.05	9/3736 (0.2%)
All	All	0.96	6/8299 (0.1%)	1.09	36/11189 (0.3%)

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	5
1	C	0	1
All	All	0	7

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	8	GLY	C-O	-7.00	1.12	1.23
1	A	61	GLU	CD-OE1	-6.16	1.18	1.25
1	B	9	ASN	N-CA	-5.53	1.35	1.46
1	A	56	GLY	C-O	-5.47	1.14	1.23
1	A	15	GLY	C-O	-5.33	1.15	1.23

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	199	ARG	NE-CZ-NH1	-13.92	113.34	120.30
1	B	199	ARG	NE-CZ-NH2	9.38	124.99	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	39	ARG	NE-CZ-NH2	-8.41	116.09	120.30
1	B	199	ARG	CD-NE-CZ	8.17	135.04	123.60
1	B	206	GLY	C-N-CA	7.91	141.46	121.70

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	41	ARG	Sidechain
1	B	59	ARG	Sidechain
1	B	60	PHE	Mainchain
1	B	8	GLY	Mainchain
1	B	9	ASN	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	2711	0	2556	149	0
1	B	2705	0	2544	108	0
1	C	2717	0	2565	139	0
2	A	1	0	0	0	0
3	A	203	0	0	9	0
3	B	168	0	0	3	1
3	C	149	0	0	5	0
All	All	8654	0	7665	363	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 23.

The worst 5 of 363 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:199:ARG:NH1	1:B:203:GLN:HB3	1.58	1.17
1:B:220:GLY:O	1:B:221:LYS:HG2	1.46	1.16

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:111:PHE:CE1	1:A:268:VAL:CG1	2.35	1.09
1:C:257:LEU:CD1	1:C:336:TYR:CE1	2.37	1.08
1:C:257:LEU:CD1	1:C:336:TYR:HE1	1.67	1.07

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)
3:B:528:HOH:O	3:B:528:HOH:O[2_655]	1.55	0.65

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	344/342 (101%)	303 (88%)	31 (9%)	10 (3%)	4	23
1	B	343/342 (100%)	307 (90%)	30 (9%)	6 (2%)	9	35
1	C	345/342 (101%)	303 (88%)	36 (10%)	6 (2%)	9	35
All	All	1032/1026 (101%)	913 (88%)	97 (9%)	22 (2%)	6	29

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	27	ASP
1	B	217	SER
1	A	28	LYS
1	A	160	ASP
1	C	168	MET

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	278/274 (102%)	263 (95%)	15 (5%)	22	53
1	B	277/274 (101%)	255 (92%)	22 (8%)	12	39
1	C	279/274 (102%)	264 (95%)	15 (5%)	22	53
All	All	834/822 (102%)	782 (94%)	52 (6%)	18	48

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	222	ARG
1	B	338	ILE
1	C	284	ASN
1	B	255	LEU
1	B	329[A]	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	290	ASN
1	C	49	GLN
1	C	290	ASN
1	C	172	ASN
1	C	9	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is 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 (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 (i)

6.1 Protein, DNA and RNA chains (i)

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	342/342 (100%)	-0.20	3 (0%)	84	71	74, 99, 143, 163	0
1	B	342/342 (100%)	-0.14	8 (2%)	60	39	66, 91, 129, 188	0
1	C	342/342 (100%)	-0.22	2 (0%)	89	79	65, 94, 135, 176	0
All	All	1026/1026 (100%)	-0.19	13 (1%)	77	60	65, 94, 138, 188	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	144	TYR	4.7
1	B	146	ASP	3.4
1	B	28	LYS	3.1
1	B	82	ALA	2.8
1	C	148	LEU	2.7

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 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, 95th 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	CA	A	401	1/1	0.92	0.24	79,79,79,79	0

6.5 Other polymers (i)

There are no such residues in this entry.