wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 09:18 AM GMT

PDB ID	:	3HX6
Title	:	Crystal structure of Pseudomonas aeruginosa PilY1 C-terminal domain
Authors	:	Redinbo, M.R.; Orans, J.
Deposited on	:	2009-06-19
Resolution	:	2.10 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
http://wwpdb.org/validation/2016/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:

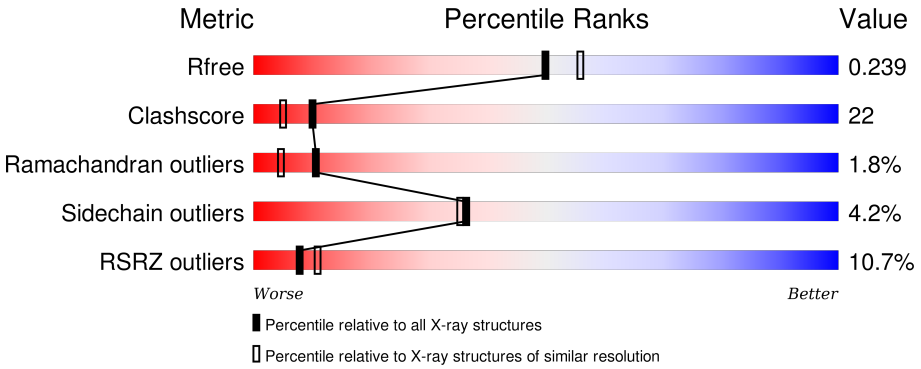
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.10 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	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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 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	570	8% 58% 26% • 14%
1	B	570	10% 56% 26% • 15%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8125 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 Type 4 fimbrial biogenesis protein PilY1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	491	Total	C	N	O	S	0	0	0
			3700	2320	641	730	9			
1	B	486	Total	C	N	O	S	0	0	0
			3663	2299	633	723	8			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	594	MET	-	EXPRESSION TAG	UNP Q9HVM8
A	595	GLY	-	EXPRESSION TAG	UNP Q9HVM8
A	596	SER	-	EXPRESSION TAG	UNP Q9HVM8
A	597	SER	-	EXPRESSION TAG	UNP Q9HVM8
A	598	HIS	-	EXPRESSION TAG	UNP Q9HVM8
A	599	HIS	-	EXPRESSION TAG	UNP Q9HVM8
A	600	HIS	-	EXPRESSION TAG	UNP Q9HVM8
A	601	HIS	-	EXPRESSION TAG	UNP Q9HVM8
A	602	HIS	-	EXPRESSION TAG	UNP Q9HVM8
A	603	HIS	-	EXPRESSION TAG	UNP Q9HVM8
A	604	SER	-	EXPRESSION TAG	UNP Q9HVM8
A	605	SER	-	EXPRESSION TAG	UNP Q9HVM8
A	606	GLY	-	EXPRESSION TAG	UNP Q9HVM8
A	607	LEU	-	EXPRESSION TAG	UNP Q9HVM8
A	608	VAL	-	EXPRESSION TAG	UNP Q9HVM8
A	609	PRO	-	EXPRESSION TAG	UNP Q9HVM8
A	610	ARG	-	EXPRESSION TAG	UNP Q9HVM8
A	611	GLY	-	EXPRESSION TAG	UNP Q9HVM8
A	612	SER	-	EXPRESSION TAG	UNP Q9HVM8
A	613	HIS	-	EXPRESSION TAG	UNP Q9HVM8
A	614	MET	-	EXPRESSION TAG	UNP Q9HVM8
A	712	MET	LEU	ENGINEERED	UNP Q9HVM8
A	812	MET	LEU	ENGINEERED	UNP Q9HVM8
A	823	MET	LEU	ENGINEERED	UNP Q9HVM8
B	594	MET	-	EXPRESSION TAG	UNP Q9HVM8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	595	GLY	-	EXPRESSION TAG	UNP Q9HVM8
B	596	SER	-	EXPRESSION TAG	UNP Q9HVM8
B	597	SER	-	EXPRESSION TAG	UNP Q9HVM8
B	598	HIS	-	EXPRESSION TAG	UNP Q9HVM8
B	599	HIS	-	EXPRESSION TAG	UNP Q9HVM8
B	600	HIS	-	EXPRESSION TAG	UNP Q9HVM8
B	601	HIS	-	EXPRESSION TAG	UNP Q9HVM8
B	602	HIS	-	EXPRESSION TAG	UNP Q9HVM8
B	603	HIS	-	EXPRESSION TAG	UNP Q9HVM8
B	604	SER	-	EXPRESSION TAG	UNP Q9HVM8
B	605	SER	-	EXPRESSION TAG	UNP Q9HVM8
B	606	GLY	-	EXPRESSION TAG	UNP Q9HVM8
B	607	LEU	-	EXPRESSION TAG	UNP Q9HVM8
B	608	VAL	-	EXPRESSION TAG	UNP Q9HVM8
B	609	PRO	-	EXPRESSION TAG	UNP Q9HVM8
B	610	ARG	-	EXPRESSION TAG	UNP Q9HVM8
B	611	GLY	-	EXPRESSION TAG	UNP Q9HVM8
B	612	SER	-	EXPRESSION TAG	UNP Q9HVM8
B	613	HIS	-	EXPRESSION TAG	UNP Q9HVM8
B	614	MET	-	EXPRESSION TAG	UNP Q9HVM8
B	712	MET	LEU	ENGINEERED	UNP Q9HVM8
B	812	MET	LEU	ENGINEERED	UNP Q9HVM8
B	823	MET	LEU	ENGINEERED	UNP Q9HVM8

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

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

- Molecule 3 is water.

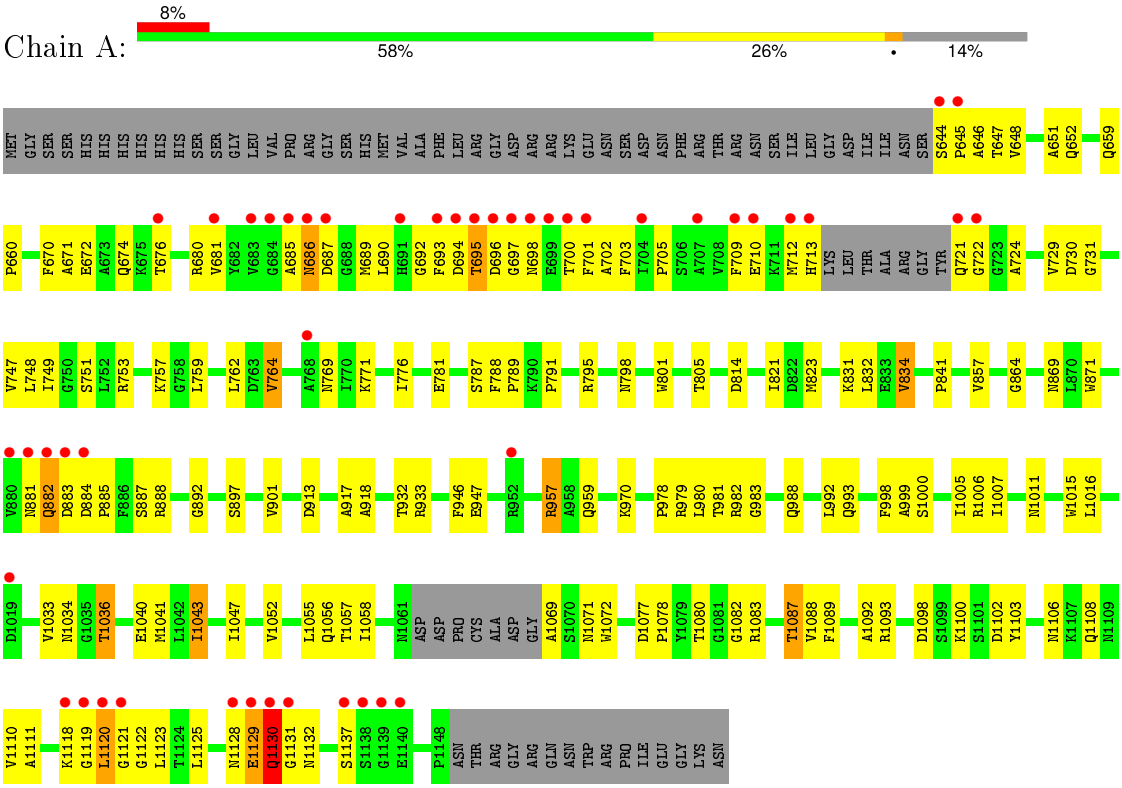
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	432	Total O 432 432	0	0
3	B	328	Total O 328 328	0	0

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.

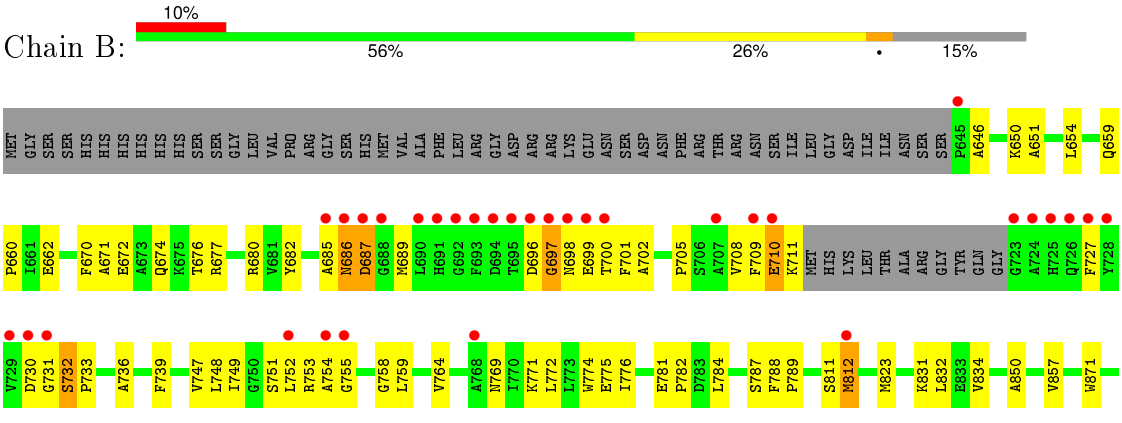
- Molecule 1: Type 4 fimbrial biogenesis protein PilY1

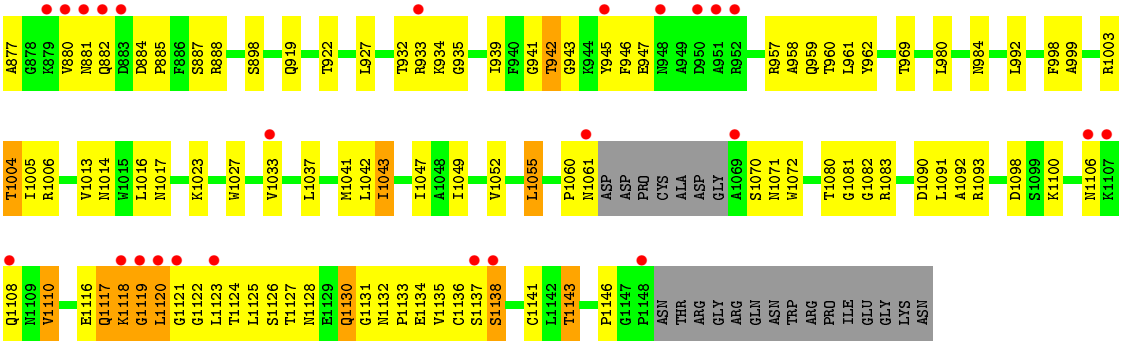
Chain A:



- Molecule 1: Type 4 fimbrial biogenesis protein PilY1

Chain B:




A877 G878 K879 V880 N881 Q882 D883 D884 P885 F886 S887 R888 S898 Q919 T922 L927 T932 R933 K934 G935 I939 F940 G941 T942 G943 K944 Y945 F946 E947 N948 A949 D950 A951 R952 R957 A958 Q959 T960 L961 Y962 T969 L980 N984 L992 F998 A999 R1003
T1004 I1005 R1006 V1013 N1014 W1015 L1016 N1017 K1023 W1027 V1033 L1037 M1041 L1042 I1043 I1047 A1048 I1049 V1052 L1055 P1060 N1061 ASP ASP PRO CYS ALA ASP GLY A1069 S1070 N1071 W1072 T1080 G1081 G1082 R1083 D1090 L1091 A1092 R1093 D1098 S1099 K1100 N1106 K1107
Q1108 N1109 V1110 E1116 Q1117 K1118 G1119 L1120 G1121 G1122 L1123 T1124 L1125 S1126 T1127 N1128 E1129 Q1130 G1131 N1132 P1133 E1134 V1135 C1136 S1137 S1138 C1141 L1142 T1143 P1146 G1147 P1148 ASN THR ARG GLY ARG GLN ASN TRP ARG PRO ILE GLU GLY LYS ASN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	64.00Å 108.00Å 159.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.10 47.58 – 2.09	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-2.10) 96.5 (47.58-2.09)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.07 (at 2.08Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.196 , 0.235 0.201 , 0.239	Depositor DCC
R_{free} test set	6250 reflections (9.85%)	DCC
Wilson B-factor (Å^2)	23.4	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 52.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Outliers	3 of 123850 reflections (0.002%)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8125	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% 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: 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	0.33	0/3781	0.65	0/5144
1	B	0.32	0/3743	0.62	0/5093
All	All	0.32	0/7524	0.64	0/10237

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	3700	0	3560	160	0
1	B	3663	0	3529	164	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	432	0	0	24	0
3	B	328	0	0	24	0
All	All	8125	0	7089	317	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 22.

The worst 5 of 317 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:857:VAL:HG21	1:A:888:ARG:HD3	1.36	1.08
1:A:680:ARG:HE	1:A:700:THR:HG21	1.23	1.01
1:A:1040:GLU:OE2	1:A:1057:THR:HG21	1.60	0.99
1:B:857:VAL:HG21	1:B:888:ARG:HD3	1.43	0.98
1:A:1080:THR:HG23	1:A:1082:GLY:H	1.31	0.94

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	485/570 (85%)	443 (91%)	34 (7%)	8 (2%)	12	6
1	B	480/570 (84%)	451 (94%)	20 (4%)	9 (2%)	10	4
All	All	965/1140 (85%)	894 (93%)	54 (6%)	17 (2%)	11	5

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	882	GLN
1	A	1129	GLU
1	A	1130	GLN
1	B	687	ASP
1	B	732	SER

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	386/453 (85%)	374 (97%)	12 (3%)	47 50
1	B	382/453 (84%)	362 (95%)	20 (5%)	29 25
All	All	768/906 (85%)	736 (96%)	32 (4%)	36 35

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	730	ASP
1	B	934	LYS
1	B	1118	LYS
1	B	881	ASN
1	B	942	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1132	ASN
1	B	698	ASN
1	B	1117	GLN
1	B	674	GLN
1	B	769	ASN

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 2 ligands modelled in this entry, 2 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	491/570 (86%)	0.54	47 (9%) 10 14	11, 23, 63, 87	0
1	B	486/570 (85%)	0.70	58 (11%) 6 8	13, 27, 62, 83	0
All	All	977/1140 (85%)	0.62	105 (10%) 8 11	11, 25, 63, 87	0

The worst 5 of 105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	694	ASP	11.5
1	A	693	PHE	10.9
1	A	1139	GLY	10.5
1	B	882	GLN	8.3
1	B	730	ASP	8.1

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 carbohydrates 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. 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
2	CA	B	1	1/1	0.99	0.11	0.32	18,18,18,18	0
2	CA	A	1	1/1	0.99	0.03	-6.19	20,20,20,20	0

6.5 Other polymers (i)

There are no such residues in this entry.