wwPDB X-ray Structure Validation Summary Report (i)

Nov 13, 2023 – 06:10 pm GMT

PDB ID	:	7Z55
Title	:	Crystal Structure of the Ring Nuclease 0455 from Sulfolobus islandicus (Sis0455) in complex with its substrate
Authors	:	Molina, R.; Martin-Garcia, R.; Lopez-Mendez, B.; Jensen, A.L.G.; Marchena-Hurtado, J.; Stella, S.; Montoya, G.
Deposited on	:	2022-03-08
Resolution	:	1.66 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 (i)) were used in the production of this report:

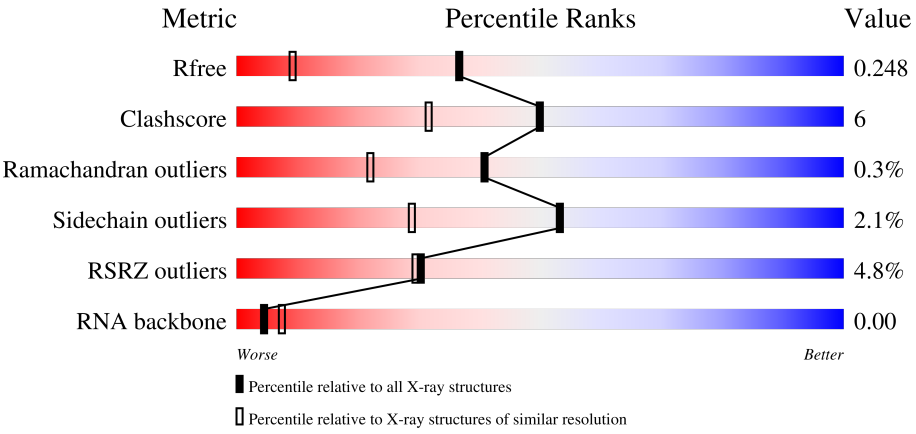
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.66 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	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)
RNA backbone	3102	1011 (2.36-0.96)

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	AAA	207	4% 81% 10% 9%
1	BBB	207	5% 75% 14% • 9%
2	CCC	4	50% 50%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3140 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 CRISPR-associated protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	188	Total	C	N	O	S	Se	0	0	0
			1495	955	247	286	4	3			
1	BBB	189	Total	C	N	O	S	Se	0	1	0
			1512	964	251	290	4	3			

There are 58 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	179	GLY	-	expression tag	UNP F0NGX6
AAA	180	SER	-	expression tag	UNP F0NGX6
AAA	181	GLU	-	expression tag	UNP F0NGX6
AAA	182	PHE	-	expression tag	UNP F0NGX6
AAA	183	GLU	-	expression tag	UNP F0NGX6
AAA	184	LEU	-	expression tag	UNP F0NGX6
AAA	185	GLU	-	expression tag	UNP F0NGX6
AAA	186	ASN	-	expression tag	UNP F0NGX6
AAA	187	LEU	-	expression tag	UNP F0NGX6
AAA	188	TYR	-	expression tag	UNP F0NGX6
AAA	189	PHE	-	expression tag	UNP F0NGX6
AAA	190	GLN	-	expression tag	UNP F0NGX6
AAA	191	GLY	-	expression tag	UNP F0NGX6
AAA	192	GLU	-	expression tag	UNP F0NGX6
AAA	193	LEU	-	expression tag	UNP F0NGX6
AAA	194	ARG	-	expression tag	UNP F0NGX6
AAA	195	ARG	-	expression tag	UNP F0NGX6
AAA	196	GLN	-	expression tag	UNP F0NGX6
AAA	197	ALA	-	expression tag	UNP F0NGX6
AAA	198	SER	-	expression tag	UNP F0NGX6
AAA	199	ALA	-	expression tag	UNP F0NGX6
AAA	200	LEU	-	expression tag	UNP F0NGX6
AAA	201	GLU	-	expression tag	UNP F0NGX6
AAA	202	HIS	-	expression tag	UNP F0NGX6
AAA	203	HIS	-	expression tag	UNP F0NGX6

Continued on next page...

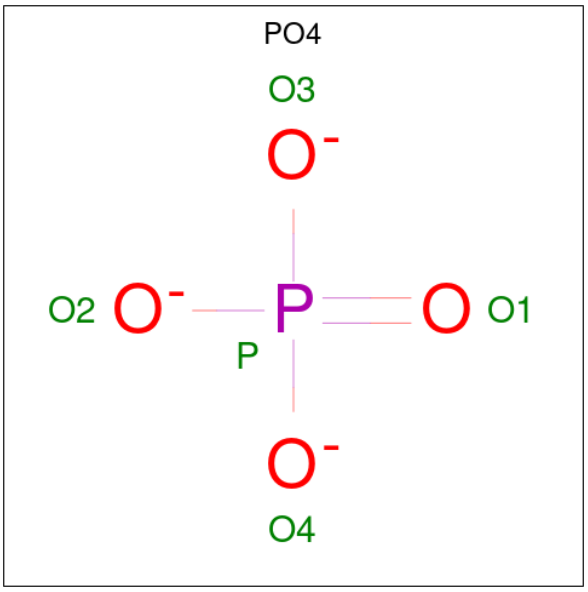
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	204	HIS	-	expression tag	UNP F0NGX6
AAA	205	HIS	-	expression tag	UNP F0NGX6
AAA	206	HIS	-	expression tag	UNP F0NGX6
AAA	207	HIS	-	expression tag	UNP F0NGX6
BBB	179	GLY	-	expression tag	UNP F0NGX6
BBB	180	SER	-	expression tag	UNP F0NGX6
BBB	181	GLU	-	expression tag	UNP F0NGX6
BBB	182	PHE	-	expression tag	UNP F0NGX6
BBB	183	GLU	-	expression tag	UNP F0NGX6
BBB	184	LEU	-	expression tag	UNP F0NGX6
BBB	185	GLU	-	expression tag	UNP F0NGX6
BBB	186	ASN	-	expression tag	UNP F0NGX6
BBB	187	LEU	-	expression tag	UNP F0NGX6
BBB	188	TYR	-	expression tag	UNP F0NGX6
BBB	189	PHE	-	expression tag	UNP F0NGX6
BBB	190	GLN	-	expression tag	UNP F0NGX6
BBB	191	GLY	-	expression tag	UNP F0NGX6
BBB	192	GLU	-	expression tag	UNP F0NGX6
BBB	193	LEU	-	expression tag	UNP F0NGX6
BBB	194	ARG	-	expression tag	UNP F0NGX6
BBB	195	ARG	-	expression tag	UNP F0NGX6
BBB	196	GLN	-	expression tag	UNP F0NGX6
BBB	197	ALA	-	expression tag	UNP F0NGX6
BBB	198	SER	-	expression tag	UNP F0NGX6
BBB	199	ALA	-	expression tag	UNP F0NGX6
BBB	200	LEU	-	expression tag	UNP F0NGX6
BBB	201	GLU	-	expression tag	UNP F0NGX6
BBB	202	HIS	-	expression tag	UNP F0NGX6
BBB	203	HIS	-	expression tag	UNP F0NGX6
BBB	204	HIS	-	expression tag	UNP F0NGX6
BBB	205	HIS	-	expression tag	UNP F0NGX6
BBB	206	HIS	-	expression tag	UNP F0NGX6
BBB	207	HIS	-	expression tag	UNP F0NGX6

- Molecule 2 is a RNA chain called Cyclic RNA cA4.

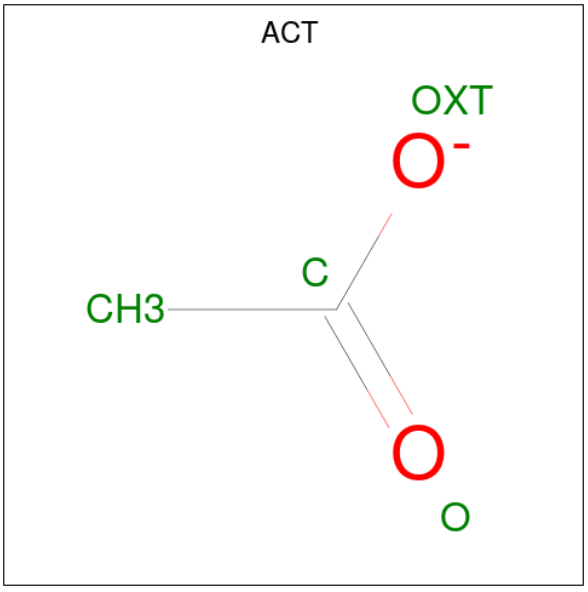
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	CCC	4	Total	C	N	O	P	0	0	0
			88	40	20	24	4			

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total O P 5 4 1	0	0

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



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

- Molecule 5 is water.

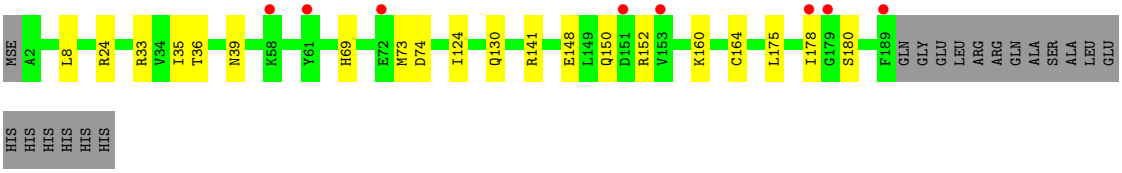
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	14	Total O 14 14	0	0
5	BBB	16	Total O 16 16	0	0
5	CCC	6	Total O 6 6	0	0

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.

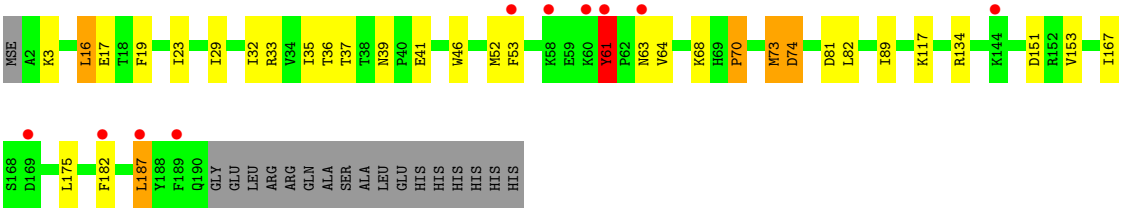
• Molecule 1: CRISPR-associated protein

Chain AAA:



• Molecule 1: CRISPR-associated protein

Chain BBB:



• Molecule 2: Cyclic RNA cA4

Chain CCC:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	42.98Å 79.53Å 95.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.92 – 1.66 47.87 – 1.66	Depositor EDS
% Data completeness (in resolution range)	98.9 (47.92-1.66) 98.2 (47.87-1.66)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	21.55 (at 1.66Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.209 , 0.252 0.214 , 0.248	Depositor DCC
R_{free} test set	1938 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	32.7	Xtriage
Anisotropy	0.152	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 45.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3140	wwPDB-VP
Average B, all atoms (Å^2)	39.0	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 22.96 % 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 5.1266e-03. 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

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, ACT

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	AAA	0.84	0/1513	1.06	2/2035 (0.1%)
1	BBB	0.86	0/1530	1.04	2/2058 (0.1%)
2	CCC	1.22	1/99 (1.0%)	1.60	0/152
All	All	0.86	1/3142 (0.0%)	1.07	4/4245 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	CCC	4	A	P-O5'	-5.24	1.54	1.59

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	AAA	33	ARG	CG-CD-NE	-6.10	99.00	111.80
1	BBB	73	MSE	CG-SE-CE	5.81	111.68	98.90
1	AAA	33	ARG	NE-CZ-NH2	-5.56	117.52	120.30
1	BBB	70	PRO	N-CA-CB	-5.09	97.00	102.60

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	AAA	1495	0	1532	12	0
1	BBB	1512	0	1545	28	0
2	CCC	88	0	44	2	0
3	AAA	5	0	0	0	0
4	BBB	4	0	3	1	0
5	AAA	14	0	0	0	0
5	BBB	16	0	0	1	0
5	CCC	6	0	0	0	0
All	All	3140	0	3124	36	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 6.

The worst 5 of 36 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:BBB:175:LEU:HD22	1:BBB:187:LEU:HD23	1.68	0.74
1:BBB:41:GLU:HG3	5:BBB:410:HOH:O	1.88	0.73
1:BBB:175:LEU:CD2	1:BBB:187:LEU:HD23	2.20	0.70
1:BBB:19:PHE:CD1	1:BBB:32:ILE:HD11	2.35	0.62
1:AAA:178:ILE:HD11	1:BBB:82:LEU:HD23	1.83	0.60

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	AAA	186/207 (90%)	183 (98%)	3 (2%)	0	100 100
1	BBB	188/207 (91%)	184 (98%)	3 (2%)	1 (0%)	29 11
All	All	374/414 (90%)	367 (98%)	6 (2%)	1 (0%)	41 22

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	BBB	61	TYR

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	AAA	167/179 (93%)	164 (98%)	3 (2%)	59 36
1	BBB	169/179 (94%)	165 (98%)	4 (2%)	49 23
All	All	336/358 (94%)	329 (98%)	7 (2%)	53 29

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	16	LEU
1	BBB	61	TYR
1	BBB	187	LEU
1	BBB	74	ASP
1	AAA	152	ARG

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

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	CCC	3/4 (75%)	3 (100%)	0

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	CCC	2	A
2	CCC	3	A
2	CCC	4	A

There are no RNA pucker outliers to report.

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)

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	$\#\|Z\| > 2$
4	ACT	BBB	301	-	3,3,3	0.99	0	3,3,3	0.73	0
3	PO4	AAA	301	-	4,4,4	1.09	0	6,6,6	0.42	0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	BBB	301	ACT	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	AAA	185/207 (89%)	0.29	8 (4%) 35 34	25, 36, 60, 76	0
1	BBB	186/207 (89%)	0.40	10 (5%) 25 24	27, 36, 58, 108	0
2	CCC	4/4 (100%)	0.01	0 100 100	33, 33, 34, 34	0
All	All	375/418 (89%)	0.34	18 (4%) 30 29	25, 36, 60, 108	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	61	TYR	6.9
1	BBB	144	LYS	4.8
1	BBB	58	LYS	3.5
1	AAA	178	ILE	3.4
1	AAA	58	LYS	3.1

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 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
4	ACT	BBB	301	4/4	0.90	0.15	29,35,41,48	0
3	PO4	AAA	301	5/5	0.96	0.08	38,42,44,47	0

6.5 Other polymers (i)

There are no such residues in this entry.