wwPDB X-ray Structure Validation Summary Report (i)

Feb 12, 2024 – 02:33 PM EST

PDB ID	:	3I1A
Title	:	Crystal Structure of apo Spectinomycin Phosphotransferase, APH(9)-Ia
Authors	:	Berghuis, A.M.; Fong, D.H.; Lemke, C.T.; Hwang, J.; Xiong, B.
Deposited on	:	2009-06-25
Resolution	:	1.70 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:

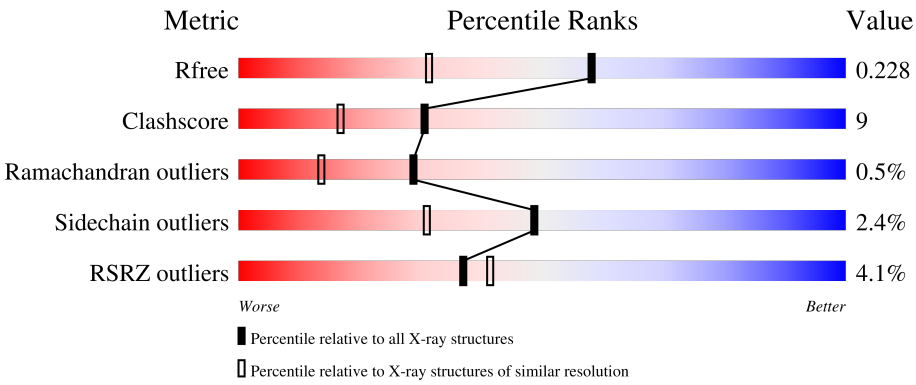
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	339	4% 82% 14% • •
1	B	339	4% 81% 13% • •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6505 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.

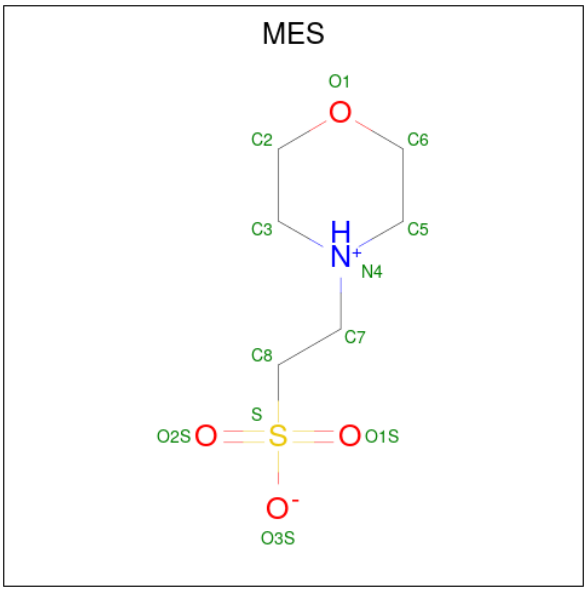
- Molecule 1 is a protein called Spectinomycin phosphotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	328	2801	1795	474	528	4	0	17	0
1	B	329	2787	1796	472	515	4	0	12	0

There are 16 discrepancies between the modelled and reference sequences:

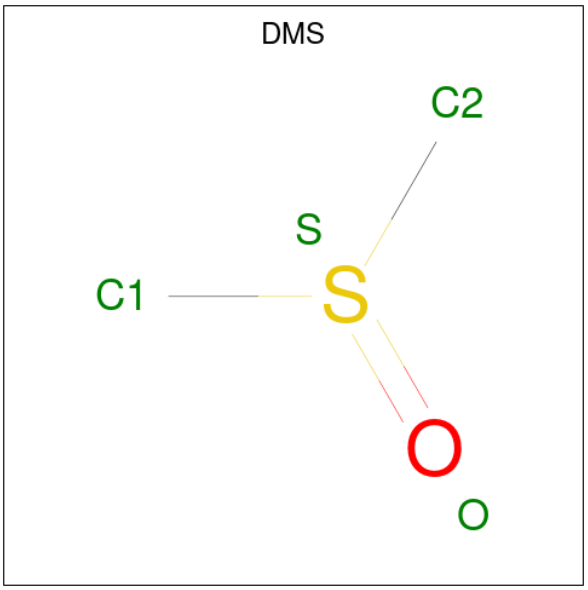
Chain	Residue	Modelled	Actual	Comment	Reference
A	332	LEU	-	expression tag	UNP O06916
A	333	GLU	-	expression tag	UNP O06916
A	334	HIS	-	expression tag	UNP O06916
A	335	HIS	-	expression tag	UNP O06916
A	336	HIS	-	expression tag	UNP O06916
A	337	HIS	-	expression tag	UNP O06916
A	338	HIS	-	expression tag	UNP O06916
A	339	HIS	-	expression tag	UNP O06916
B	1332	LEU	-	expression tag	UNP O06916
B	1333	GLU	-	expression tag	UNP O06916
B	1334	HIS	-	expression tag	UNP O06916
B	1335	HIS	-	expression tag	UNP O06916
B	1336	HIS	-	expression tag	UNP O06916
B	1337	HIS	-	expression tag	UNP O06916
B	1338	HIS	-	expression tag	UNP O06916
B	1339	HIS	-	expression tag	UNP O06916

- Molecule 2 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	A	1	12	6	1	4	1	0	0
			Total	C	N	O	S		
2	B	1	12	6	1	4	1	0	0
			Total	C	N	O	S		
2	B	1	12	6	1	4	1	0	0

- Molecule 3 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



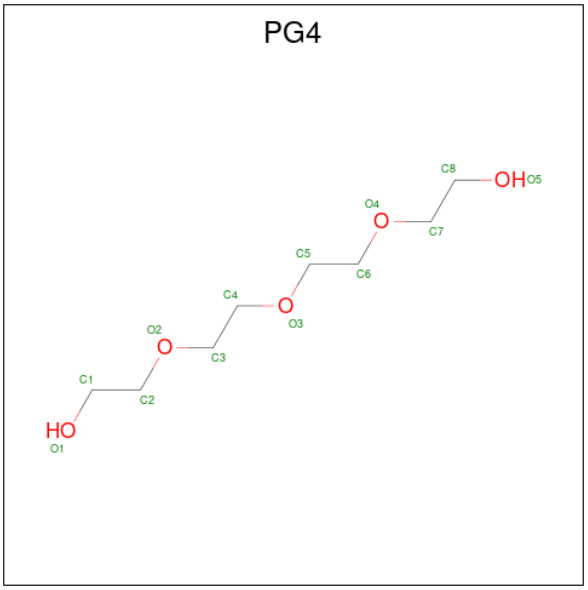
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
3	A	1	4	2	1	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
3	B	1	4	2	1	1	0	0

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	B	1	13	8	5	0	0

- Molecule 5 is water.

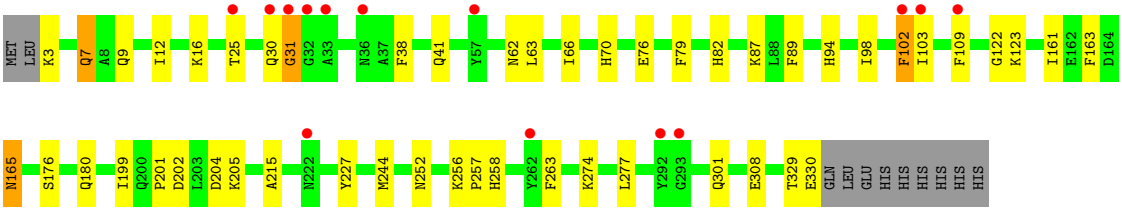
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	412	412	412	0	0
5	B	448	448	448	0	0

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.

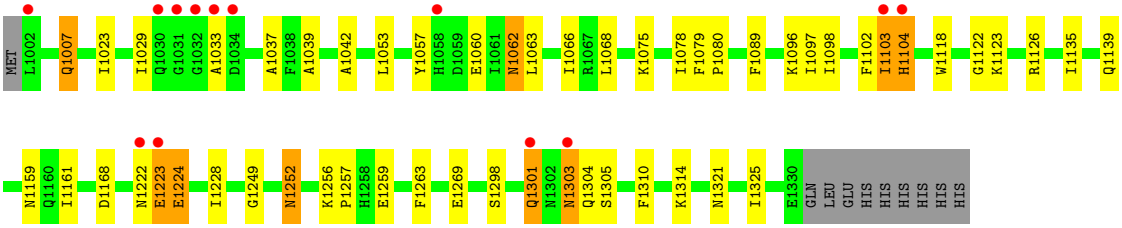
- Molecule 1: Spectinomycin phosphotransferase

Chain A:



- Molecule 1: Spectinomycin phosphotransferase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	47.41Å 70.69Å 99.31Å 90.00° 95.93° 90.00°	Depositor
Resolution (Å)	44.38 – 1.70 49.39 – 1.69	Depositor EDS
% Data completeness (in resolution range)	89.3 (44.38-1.70) 95.2 (49.39-1.69)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	2.85 (at 1.70Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.182 , 0.228 0.180 , 0.228	Depositor DCC
R_{free} test set	6473 reflections (9.05%)	wwPDB-VP
Wilson B-factor (Å^2)	18.3	Xtriage
Anisotropy	0.208	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 57.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6505	wwPDB-VP
Average B, all atoms (Å^2)	24.0	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 48.74 % 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 8.2425e-05. 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: DMS, PG4, MES

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	0.46	0/2869	0.65	0/3879
1	B	0.51	0/2856	0.68	1/3860 (0.0%)
All	All	0.49	0/5725	0.67	1/7739 (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

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	1007	GLN	N-CA-C	-5.74	95.51	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	227	TYR	Sidechain

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	2801	0	2686	50	0
1	B	2787	0	2713	52	0
2	A	12	0	13	0	0
2	B	24	0	26	1	0
3	A	4	0	6	0	0
3	B	4	0	6	0	0
4	B	13	0	18	0	0
5	A	412	0	0	13	0
5	B	448	0	0	11	1
All	All	6505	0	5468	103	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 9.

The worst 5 of 103 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:1223:GLU:O	1:B:1224:GLU:HG2	1.78	0.83
1:B:1103[A]:ILE:H	1:B:1103[A]:ILE:HD13	1.41	0.83
1:B:1078:ILE:HD13	1:B:1103[B]:ILE:HD12	1.66	0.77
1:B:1222:ASN:O	1:B:1224:GLU:N	2.18	0.77
1:A:9[B]:GLN:NE2	1:A:12:ILE:HD12	2.03	0.73

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)
5:B:146:HOH:O	5:B:325:HOH:O[1_455]	2.17	0.03

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	343/339 (101%)	334 (97%)	8 (2%)	1 (0%)	41	24
1	B	339/339 (100%)	324 (96%)	13 (4%)	2 (1%)	25	11
All	All	682/678 (101%)	658 (96%)	21 (3%)	3 (0%)	29	18

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1224	GLU
1	B	1223	GLU
1	A	31	GLY

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	301/304 (99%)	294 (98%)	7 (2%)	50	33
1	B	300/304 (99%)	290 (97%)	10 (3%)	38	19
All	All	601/608 (99%)	584 (97%)	17 (3%)	49	25

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1252	ASN
1	B	1303	ASN
1	B	1057	TYR
1	B	1062	ASN
1	B	1103[A]	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1294	GLN
1	B	1320	ASN
1	A	301	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	320	ASN
1	B	1004	GLN

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)

6 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	PG4	B	1340	-	12,12,12	0.87	0	11,11,11	0.49	0
2	MES	A	340	-	12,12,12	1.31	2 (16%)	14,16,16	0.93	1 (7%)
3	DMS	B	1	-	3,3,3	0.24	0	3,3,3	0.65	0
2	MES	B	3	-	12,12,12	1.41	3 (25%)	14,16,16	0.63	0
2	MES	B	2	-	12,12,12	1.31	3 (25%)	14,16,16	0.89	1 (7%)
3	DMS	A	341	-	3,3,3	0.19	0	3,3,3	0.60	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	MES	B	2	-	-	0/6/14/14	0/1/1/1
2	MES	A	340	-	-	0/6/14/14	0/1/1/1
2	MES	B	3	-	-	0/6/14/14	0/1/1/1
4	PG4	B	1340	-	-	0/10/10/10	-

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	340	MES	C3-N4	2.69	1.54	1.46
2	B	3	MES	C3-N4	2.50	1.53	1.46
2	B	3	MES	C5-N4	2.41	1.53	1.46
2	B	2	MES	C3-N4	2.40	1.53	1.46
2	B	2	MES	C5-N4	2.27	1.53	1.46

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	2	MES	O2S-S-C8	-2.06	104.43	106.92
2	A	340	MES	O1S-S-C8	2.04	109.38	106.92

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
2	B	3	MES	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 (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	328/339 (96%)	0.17	14 (4%)	35	39	12, 22, 38, 60	0
1	B	329/339 (97%)	0.02	13 (3%)	38	42	9, 19, 38, 59	1 (0%)
All	All	657/678 (96%)	0.09	27 (4%)	37	41	9, 21, 38, 60	1 (0%)

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1032	GLY	10.7
1	A	33	ALA	10.5
1	B	1033	ALA	7.6
1	A	32	GLY	7.2
1	A	222	ASN	6.3

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	MES	B	3	12/12	0.80	0.25	54,55,61,61	0
2	MES	B	2	12/12	0.87	0.15	30,34,36,36	0
3	DMS	B	1	4/4	0.89	0.16	37,38,40,42	0
2	MES	A	340	12/12	0.92	0.15	34,37,38,38	0
3	DMS	A	341	4/4	0.93	0.13	55,56,56,57	0
4	PG4	B	1340	13/13	0.93	0.11	17,23,37,39	0

6.5 Other polymers (i)

There are no such residues in this entry.