W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 27, 2014 – 02:01 PM GMT

This is a wwPDB 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/ValidationPDFNotes.html

---

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.70 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	Percentile Ranks	Value
Rfree		0.316
Clashscore		24
Ramachandran outliers		2.9%
Sidechain outliers		10.6%
RSRZ outliers		2.6%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
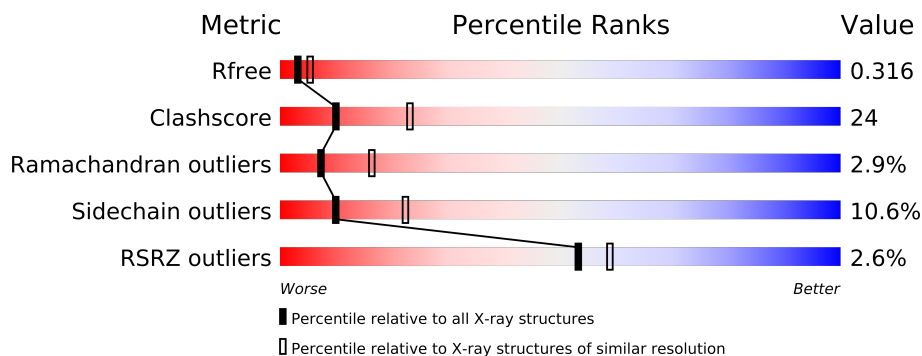
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	66092	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	297	
1	B	297	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	SO4	B	278	-	X

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 4069 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Dihydropteroate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	264	Total	C	N	O	S	0	0	0
			2000	1257	345	382	16			
1	B	262	Total	C	N	O	S	0	0	0
			1996	1256	344	380	16			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP B1UXN2
A	-18	GLY	-	EXPRESSION TAG	UNP B1UXN2
A	-17	SER	-	EXPRESSION TAG	UNP B1UXN2
A	-16	SER	-	EXPRESSION TAG	UNP B1UXN2
A	-15	HIS	-	EXPRESSION TAG	UNP B1UXN2
A	-14	HIS	-	EXPRESSION TAG	UNP B1UXN2
A	-13	HIS	-	EXPRESSION TAG	UNP B1UXN2
A	-12	HIS	-	EXPRESSION TAG	UNP B1UXN2
A	-11	HIS	-	EXPRESSION TAG	UNP B1UXN2
A	-10	HIS	-	EXPRESSION TAG	UNP B1UXN2
A	-9	SER	-	EXPRESSION TAG	UNP B1UXN2
A	-8	SER	-	EXPRESSION TAG	UNP B1UXN2
A	-7	GLY	-	EXPRESSION TAG	UNP B1UXN2
A	-6	LEU	-	EXPRESSION TAG	UNP B1UXN2
A	-5	VAL	-	EXPRESSION TAG	UNP B1UXN2
A	-4	PRO	-	EXPRESSION TAG	UNP B1UXN2
A	-3	ARG	-	EXPRESSION TAG	UNP B1UXN2
A	-2	GLY	-	EXPRESSION TAG	UNP B1UXN2
A	-1	SER	-	EXPRESSION TAG	UNP B1UXN2
A	0	HIS	-	EXPRESSION TAG	UNP B1UXN2
B	-19	MET	-	EXPRESSION TAG	UNP B1UXN2
B	-18	GLY	-	EXPRESSION TAG	UNP B1UXN2
B	-17	SER	-	EXPRESSION TAG	UNP B1UXN2
B	-16	SER	-	EXPRESSION TAG	UNP B1UXN2
B	-15	HIS	-	EXPRESSION TAG	UNP B1UXN2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	HIS	-	EXPRESSION TAG	UNP B1UXN2
B	-13	HIS	-	EXPRESSION TAG	UNP B1UXN2
B	-12	HIS	-	EXPRESSION TAG	UNP B1UXN2
B	-11	HIS	-	EXPRESSION TAG	UNP B1UXN2
B	-10	HIS	-	EXPRESSION TAG	UNP B1UXN2
B	-9	SER	-	EXPRESSION TAG	UNP B1UXN2
B	-8	SER	-	EXPRESSION TAG	UNP B1UXN2
B	-7	GLY	-	EXPRESSION TAG	UNP B1UXN2
B	-6	LEU	-	EXPRESSION TAG	UNP B1UXN2
B	-5	VAL	-	EXPRESSION TAG	UNP B1UXN2
B	-4	PRO	-	EXPRESSION TAG	UNP B1UXN2
B	-3	ARG	-	EXPRESSION TAG	UNP B1UXN2
B	-2	GLY	-	EXPRESSION TAG	UNP B1UXN2
B	-1	SER	-	EXPRESSION TAG	UNP B1UXN2
B	0	HIS	-	EXPRESSION TAG	UNP B1UXN2

- Molecule 2 is 4-{[2-(2-AMINO-4-OXO-3,4,7,8-TETRAHYDROPTERIDIN-6-YL)ETHYL]AMINO}BENZOICACID (three-letter code: B63) (formula: $C_{15}H_{16}N_6O_3$).
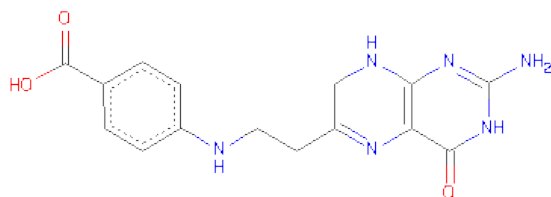
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 24	C 15	N O 6 3	0	0
2	B	1	Total 24	C 15	N O 6 3	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 5	O 4	S 1	0	0
3	A	1	Total 5	O 4	S 1	0	0
3	B	1	Total 5	O 4	S 1	0	0
3	B	1	Total 5	O 4	S 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total 2	O 2	0	0
4	B	3	Total 3	O 3	0	0

3 Residue-property plots ⓘ

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.
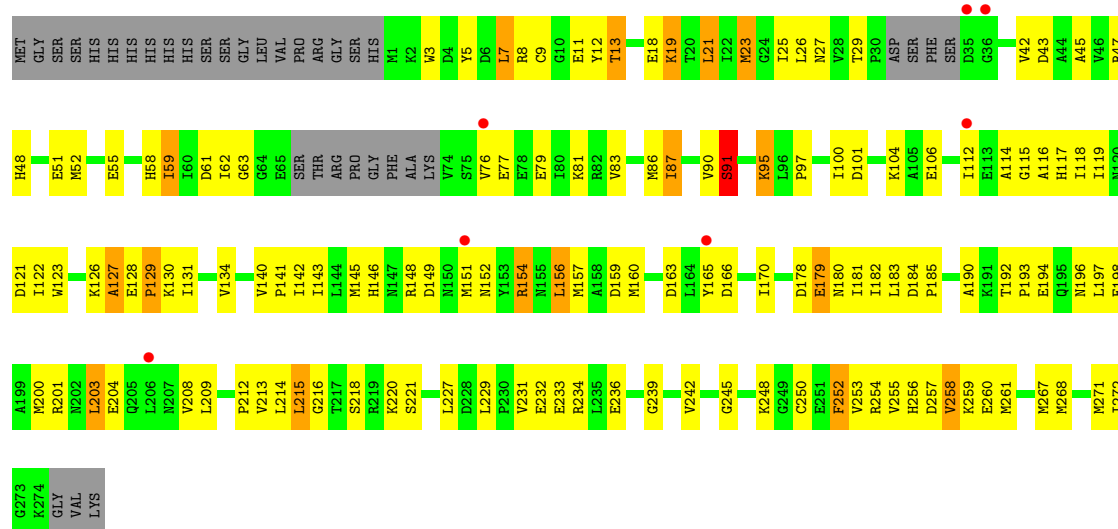
• Molecule 1: Dihydropteroate synthase

Chain A:
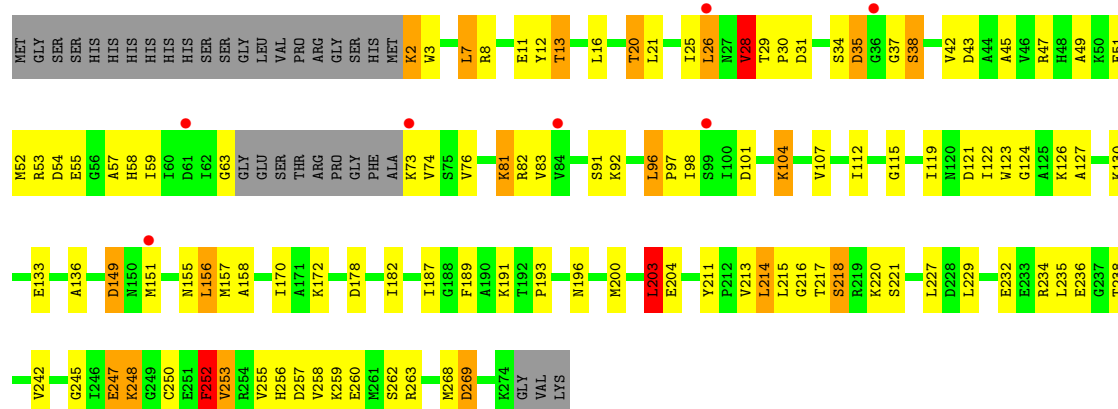
• Molecule 1: Dihydropteroate synthase

Chain B:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 62 2 2	Depositor
Cell constants a, b, c, α, β, γ	99.29Å 99.29Å 263.05Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.78 – 2.70 48.78 – 2.61	Depositor EDS
% Data completeness (in resolution range)	97.3 (48.78-2.70) 93.7 (48.78-2.61)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$<I/\sigma(I)>$ [1]	8.64 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.271 , 0.311 0.282 , 0.316	Depositor DCC
R_{free} test set	1144 reflections (5.08%)	DCC
Wilson B-factor (Å^2)	49.1	Xtriage
Anisotropy	0.930	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 24.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$<|L|> = 0.43, <L^2> = 0.26$	Xtriage
Outliers	4 of 22798 reflections (0.018%)	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	4069	wwPDB-VP
Average B, all atoms (Å^2)	49.0	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 46.76 % 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 1.0926e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: B63, SO4

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	1.10	9/2029 (0.4%)	1.06	7/2744 (0.3%)
1	B	0.93	1/2024 (0.0%)	1.03	3/2733 (0.1%)
All	All	1.02	10/4053 (0.2%)	1.04	10/5477 (0.2%)

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	81	LYS	CE-NZ	8.21	1.69	1.49
1	A	73	LYS	N-CA	8.10	1.62	1.46
1	A	248	LYS	CE-NZ	7.63	1.68	1.49
1	A	73	LYS	CA-CB	7.33	1.70	1.53
1	A	211	TYR	CE1-CZ	6.71	1.47	1.38

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	148	ARG	NE-CZ-NH1	8.38	124.49	120.30
1	A	203	LEU	CB-CG-CD1	-6.70	99.61	111.00
1	B	148	ARG	NE-CZ-NH2	-6.18	117.21	120.30
1	A	263	ARG	NE-CZ-NH2	-6.15	117.23	120.30
1	A	214	LEU	CB-CG-CD1	-5.58	101.52	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens

added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2000	0	1974	87	0
1	B	1996	0	1984	103	0
2	A	24	0	15	4	0
2	B	24	0	15	2	0
3	A	10	0	0	1	0
3	B	10	0	0	0	0
4	A	2	0	0	0	0
4	B	3	0	0	0	0
All	All	4069	0	3988	192	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 24.

The worst 5 of 192 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:81:LYS:NZ	1:A:81:LYS:CE	1.69	1.52
1:A:248:LYS:NZ	1:A:248:LYS:CE	1.68	1.51
1:B:9:CYS:HB2	1:B:12:TYR:O	1.65	0.96
1:A:2:LYS:C	1:A:2:LYS:HD3	1.95	0.85
1:A:242:VAL:HG13	1:A:253:VAL:HG11	1.59	0.84

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	260/297 (88%)	221 (85%)	33 (13%)	6 (2%)	10	24
1	B	256/297 (86%)	222 (87%)	25 (10%)	9 (4%)	6	12
All	All	516/594 (87%)	443 (86%)	58 (11%)	15 (3%)	7	16

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	156	LEU
1	B	179	GLU
1	B	239	GLY
1	B	258	VAL
1	A	37	GLY

5.3.2 Protein sidechains ⓘ

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	208/247 (84%)	185 (89%)	23 (11%)	9	20
1	B	208/247 (84%)	187 (90%)	21 (10%)	11	24
All	All	416/494 (84%)	372 (89%)	44 (11%)	10	22

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	218	SER
1	B	11	GLU
1	B	213	VAL
1	A	247	GLU
1	A	253	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	88	GLN
1	B	180	ASN
1	B	146	HIS
1	A	155	ASN
1	B	152	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

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	$\#|Z| > 2$
3	SO4	A	278	-	4,4,4	0.91	0	6,6,6	0.47	0
3	SO4	A	279	-	4,4,4	0.76	0	6,6,6	0.50	0
2	B63	A	501	-	26,26,26	1.90	5 (19%)	34,36,36	1.87	8 (23%)
3	SO4	B	278	-	4,4,4	0.68	0	6,6,6	1.11	1 (16%)
3	SO4	B	279	-	4,4,4	0.75	0	6,6,6	1.03	1 (16%)
2	B63	B	502	-	26,26,26	1.75	4 (15%)	34,36,36	1.90	8 (23%)

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
3	SO4	A	278	-	-	0/0/0/0	0/0/0/0
3	SO4	A	279	-	-	0/0/0/0	0/0/0/0
2	B63	A	501	-	-	0/10/19/19	0/1/3/3
3	SO4	B	278	-	-	0/0/0/0	0/0/0/0
3	SO4	B	279	-	-	0/0/0/0	0/0/0/0
2	B63	B	502	-	-	0/10/19/19	0/1/3/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	B63	C3-N1	5.76	1.35	1.28
2	B	502	B63	C3-N1	5.19	1.35	1.28
2	B	502	B63	C8-N4	4.19	1.38	1.32
2	A	501	B63	C8-N5	3.90	1.38	1.33
2	B	502	B63	C7-C6	3.81	1.49	1.42

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	B63	C7-C6-C5	5.66	118.43	114.55
2	B	502	B63	C8-N5-C5	4.92	124.60	117.61
2	B	502	B63	N4-C8-N3	4.48	122.79	117.86
2	A	501	B63	N2-C5-N5	4.12	121.86	115.82
2	A	501	B63	C5-C6-N1	-3.79	119.30	123.15

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers (i)

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	264/297 (88%)	0.44	7 (2%)	52	57	31, 48, 68, 76	0
1	B	262/297 (88%)	0.23	7 (2%)	52	57	34, 49, 60, 69	0
All	All	526/594 (88%)	0.33	14 (2%)	53	57	31, 48, 63, 76	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	36	GLY	4.0
1	B	76	VAL	2.8
1	A	84	VAL	2.7
1	A	73	LYS	2.6
1	B	112	ILE	2.5

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, 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	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	SO4	B	278	5/5	0.29	2.02	92,95,96,97	0
3	SO4	A	279	5/5	0.26	0.71	92,93,95,96	0
2	B63	B	502	24/24	0.19	0.13	36,42,44,44	0
3	SO4	B	279	5/5	0.17	0.10	89,89,90,90	0
2	B63	A	501	24/24	0.16	-1.44	39,41,42,43	0
3	SO4	A	278	5/5	0.14	-3.44	90,90,90,90	0

6.5   Other polymers (i)

There are no such residues in this entry.