



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 06:43 PM GMT

PDB ID : 4K6J
Title : Human cohesin inhibitor WapL
Authors : Tomchick, D.R.; Yu, H.; Ouyang, Z.
Deposited on : 2013-04-16
Resolution : 2.62 Å(reported)

This is a full wwPDB validation 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>

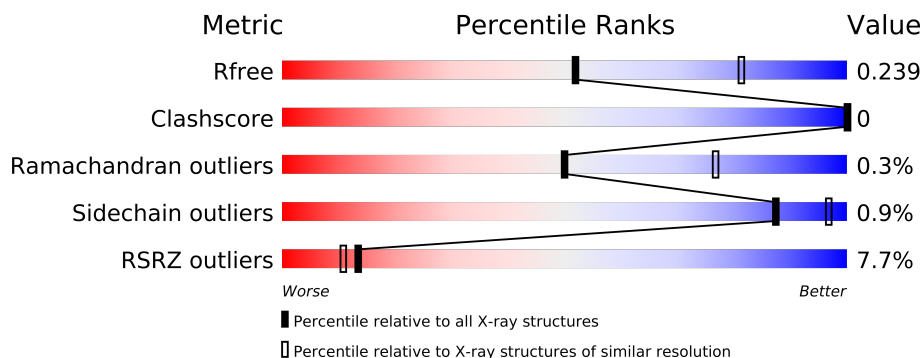
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.62 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}	66092	1940 (2.64-2.60)
Clashscore	79885	2404 (2.64-2.60)
Ramachandran outliers	78287	2360 (2.64-2.60)
Sidechain outliers	78261	2360 (2.64-2.60)
RSRZ outliers	66119	1939 (2.64-2.60)

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	568	
1	B	568	

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	ACT	A	1208	-	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 15557 atoms, of which 7730 are hydrogens 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 Wings apart-like protein homolog.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	497	Total	C	H	N	O	S	0	0	0
			7887	2451	3963	689	752	32			
1	B	471	Total	C	H	N	O	S	0	0	0
			7485	2330	3761	654	709	31			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	GLY	-	EXPRESSION TAG	UNP Q7Z5K2
A	-6	PRO	-	EXPRESSION TAG	UNP Q7Z5K2
A	-5	LEU	-	EXPRESSION TAG	UNP Q7Z5K2
A	-4	GLY	-	EXPRESSION TAG	UNP Q7Z5K2
A	-3	SER	-	EXPRESSION TAG	UNP Q7Z5K2
A	-2	GLY	-	EXPRESSION TAG	UNP Q7Z5K2
A	-1	ARG	-	EXPRESSION TAG	UNP Q7Z5K2
A	0	PRO	-	EXPRESSION TAG	UNP Q7Z5K2
B	-7	GLY	-	EXPRESSION TAG	UNP Q7Z5K2
B	-6	PRO	-	EXPRESSION TAG	UNP Q7Z5K2
B	-5	LEU	-	EXPRESSION TAG	UNP Q7Z5K2
B	-4	GLY	-	EXPRESSION TAG	UNP Q7Z5K2
B	-3	SER	-	EXPRESSION TAG	UNP Q7Z5K2
B	-2	GLY	-	EXPRESSION TAG	UNP Q7Z5K2
B	-1	ARG	-	EXPRESSION TAG	UNP Q7Z5K2
B	0	PRO	-	EXPRESSION TAG	UNP Q7Z5K2

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			7	2	3	2		
3	A	1	Total	C	H	O	0	0
			7	2	3	2		

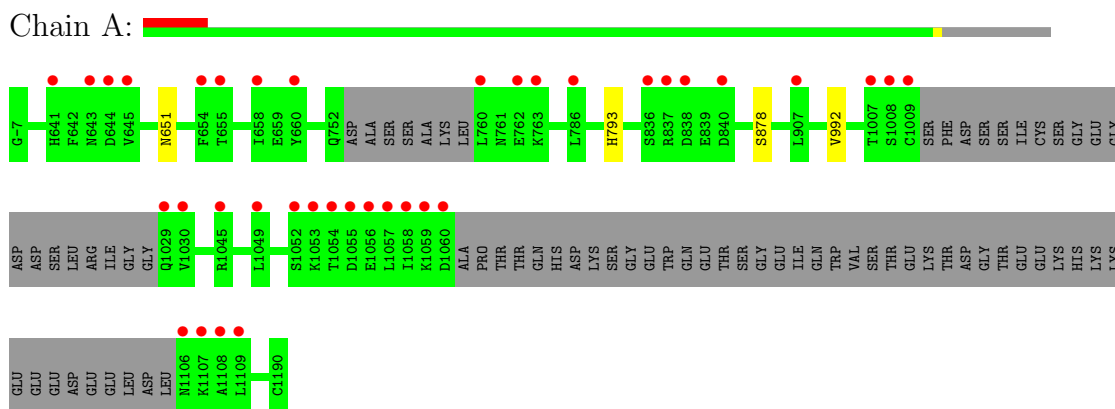
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	58	Total	O	0	0
			58	58		
4	B	48	Total	O	0	0
			48	48		

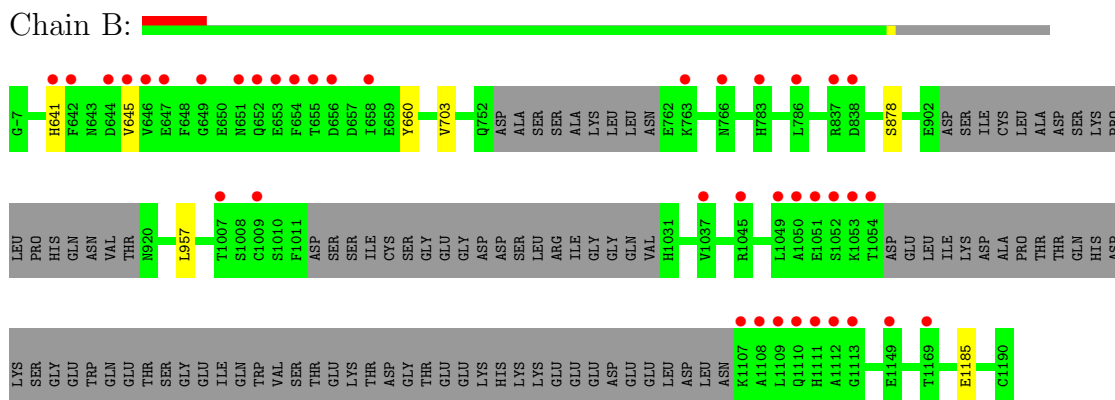
3 Residue-property plots i

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: Wings apart-like protein homolog



- Molecule 1: Wings apart-like protein homolog



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	107.54Å 107.54Å 300.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.68 – 2.62 41.84 – 2.62	Depositor EDS
% Data completeness (in resolution range)	94.8 (29.68-2.62) 94.8 (41.84-2.62)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.42 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.184 , 0.237 0.187 , 0.239	Depositor DCC
R_{free} test set	2584 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	34.7	Xtriage
Anisotropy	0.261	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.42 , 38.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 51187 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	15557	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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: SO4, 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	A	0.22	0/3979	0.39	0/5369
1	B	0.22	0/3776	0.40	0/5089
All	All	0.22	0/7755	0.39	0/10458

There are no bond length outliers.

There are no bond angle outliers.

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	3924	3963	0	0	0
1	B	3724	3761	0	0	0
2	A	35	0	0	0	0
2	B	30	0	0	0	0
3	A	8	6	0	0	0
4	A	58	0	0	0	0
4	B	48	0	0	0	0
All	All	7827	7730	0	0	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 0.

There are no clashes within the asymmetric unit.

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	489/568 (86%)	471 (96%)	17 (4%)	1 (0%)	56	81
1	B	461/568 (81%)	444 (96%)	15 (3%)	2 (0%)	43	71
All	All	950/1136 (84%)	915 (96%)	32 (3%)	3 (0%)	50	77

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	645	VAL
1	A	651	ASN
1	B	641	HIS

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	443/504 (88%)	440 (99%)	3 (1%)	91	98
1	B	418/504 (83%)	413 (99%)	5 (1%)	82	95
All	All	861/1008 (85%)	853 (99%)	8 (1%)	87	97

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	793	HIS
1	A	878	SER
1	A	992	VAL
1	B	660	TYR
1	B	703	VAL
1	B	878	SER
1	B	957	LEU
1	B	1185	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	879	GLN
1	A	920	ASN
1	A	974	GLN
1	A	976	GLN
1	A	998	ASN
1	A	1029	GLN
1	A	1189	HIS
1	B	712	HIS
1	B	714	GLN
1	B	715	ASN
1	B	879	GLN
1	B	890	HIS
1	B	896	GLN
1	B	939	ASN
1	B	974	GLN
1	B	998	ASN
1	B	1034	GLN
1	B	1038	GLN
1	B	1139	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 ⓘ

15 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	$\# Z > 2$
2	SO4	A	1201	-	4,4,4	0.26	0	6,6,6	0.11	0
2	SO4	A	1202	-	4,4,4	0.22	0	6,6,6	0.12	0
2	SO4	A	1203	-	4,4,4	0.18	0	6,6,6	0.12	0
2	SO4	A	1204	-	4,4,4	0.10	0	6,6,6	0.06	0
2	SO4	A	1205	-	4,4,4	0.12	0	6,6,6	0.08	0
2	SO4	A	1206	-	4,4,4	0.16	0	6,6,6	0.07	0
2	SO4	A	1207	-	4,4,4	0.15	0	6,6,6	0.06	0
3	ACT	A	1208	-	1,3,3	1.39	0	0,3,3	0.00	-
3	ACT	A	1209	-	1,3,3	1.29	0	0,3,3	0.00	-
2	SO4	B	1201	-	4,4,4	0.22	0	6,6,6	0.09	0
2	SO4	B	1202	-	4,4,4	0.28	0	6,6,6	0.15	0
2	SO4	B	1203	-	4,4,4	0.17	0	6,6,6	0.11	0
2	SO4	B	1204	-	4,4,4	0.17	0	6,6,6	0.23	0
2	SO4	B	1205	-	4,4,4	0.18	0	6,6,6	0.08	0
2	SO4	B	1206	-	4,4,4	0.11	0	6,6,6	0.07	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	SO4	A	1201	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1202	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1203	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1204	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1205	-	-	0/0/0/0	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SO4	A	1206	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1207	-	-	0/0/0/0	0/0/0/0
3	ACT	A	1208	-	-	0/0/0/0	0/0/0/0
3	ACT	A	1209	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1201	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1202	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1203	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1204	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1205	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1206	-	-	0/0/0/0	0/0/0/0

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.

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, 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	497/568 (87%)	0.13	37 (7%)	14 12	17, 38, 99, 147	0
1	B	471/568 (82%)	0.22	39 (8%)	11 8	15, 44, 104, 148	0
All	All	968/1136 (85%)	0.17	76 (7%)	13 10	15, 41, 103, 148	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1058	ILE	9.0
1	A	1060	ASP	8.3
1	A	643	ASN	6.3
1	A	837	ARG	6.3
1	B	652	GLN	6.2
1	B	641	HIS	6.0
1	A	1057	LEU	5.9
1	B	1109	LEU	5.8
1	B	651	ASN	5.7
1	B	1108	ALA	5.7
1	A	1059	LYS	5.3
1	B	653	GLU	5.3
1	B	1052	SER	5.0
1	B	645	VAL	4.9
1	B	642	PHE	4.7
1	A	644	ASP	4.6
1	B	1054	THR	4.5
1	A	760	LEU	4.5
1	A	1056	GLU	4.4
1	B	1053	LYS	4.4
1	A	1055	ASP	4.2
1	A	1054	THR	4.2
1	A	1030	VAL	4.0
1	B	655	THR	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	1049	LEU	3.9
1	B	644	ASP	3.8
1	B	646	VAL	3.8
1	B	654	PHE	3.7
1	B	1111	HIS	3.6
1	A	838	ASP	3.5
1	A	1107	LYS	3.5
1	A	1053	LYS	3.3
1	B	1110	GLN	3.3
1	A	763	LYS	3.2
1	A	1029	GLN	3.2
1	A	1009	CYS	3.2
1	B	647	GLU	3.2
1	A	1052	SER	3.1
1	A	1106	ASN	3.1
1	A	641	HIS	3.1
1	A	645	VAL	3.1
1	B	649	GLY	3.1
1	A	836	SER	3.1
1	A	658	ILE	3.0
1	B	1107	LYS	3.0
1	B	1112	ALA	3.0
1	B	1007	THR	2.8
1	B	837	ARG	2.8
1	B	763	LYS	2.7
1	A	1049	LEU	2.7
1	A	786	LEU	2.7
1	B	838	ASP	2.7
1	B	1149	GLU	2.6
1	B	656	ASP	2.6
1	A	660	TYR	2.6
1	A	654	PHE	2.6
1	A	907	LEU	2.5
1	A	1045	ARG	2.5
1	B	1051	GLU	2.5
1	B	783	HIS	2.4
1	B	1169	THR	2.4
1	B	786	LEU	2.4
1	A	1109	LEU	2.4
1	A	840	ASP	2.3
1	B	1045	ARG	2.3
1	B	1050	ALA	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	762	GLU	2.2
1	A	655	THR	2.1
1	A	1108	ALA	2.1
1	A	1007	THR	2.1
1	A	1008	SER	2.1
1	B	1037	VAL	2.1
1	B	1009	CYS	2.1
1	B	1113	GLY	2.0
1	B	766	ASN	2.0
1	B	658	ILE	2.0

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	ACT	A	1208	4/4	0.20	2.65	34,40,53,63	0
2	SO4	A	1204	5/5	0.20	1.73	66,79,92,94	0
2	SO4	B	1205	5/5	0.15	1.26	47,53,57,73	0
2	SO4	A	1207	5/5	0.14	0.64	49,53,68,69	0
3	ACT	A	1209	4/4	0.17	0.54	25,30,50,55	0
2	SO4	A	1202	5/5	0.14	0.43	39,47,54,66	0
2	SO4	B	1203	5/5	0.13	0.38	42,53,54,56	0
2	SO4	A	1201	5/5	0.13	-0.42	34,38,44,44	0
2	SO4	B	1201	5/5	0.12	-0.69	34,40,63,69	0
2	SO4	B	1206	5/5	0.14	-0.73	45,59,64,71	0
2	SO4	A	1205	5/5	0.14	-0.77	63,66,78,86	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	SO4	B	1204	5/5	0.12	-0.79	34,45,64,75	0
2	SO4	B	1202	5/5	0.12	-0.97	27,35,39,49	0
2	SO4	A	1206	5/5	0.13	-1.03	50,57,59,59	0
2	SO4	A	1203	5/5	0.09	-1.06	46,50,59,66	0

6.5 Other polymers ⓘ

There are no such residues in this entry.