



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:49 PM GMT

PDB ID : 4AM6
Title : C-TERMINAL DOMAIN OF ACTIN-RELATED PROTEIN ARP8 FROM S.
CEREVISIAE
Authors : Wuerges, J.; Saravanan, M.; Bose, D.; Cook, N.J.; Zhang, X.; Wigley, D.B.
Deposited on : 2012-03-07
Resolution : 2.70 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

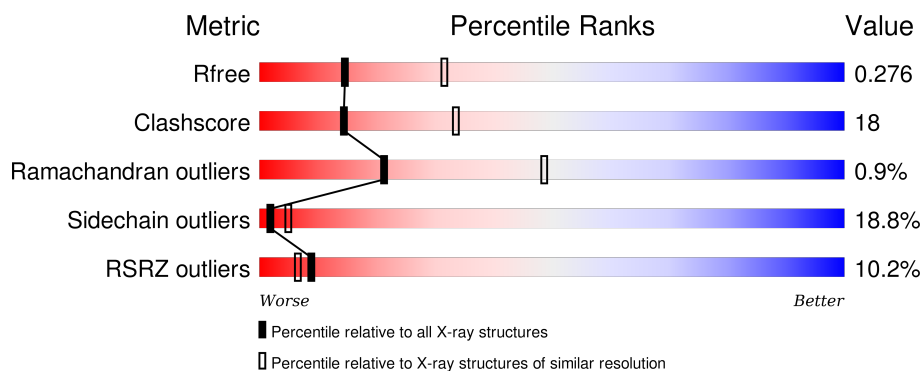
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (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. 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 $\leq 5\%$. 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	655	<div> <div>9%</div> <div>55%</div> <div>30%</div> <div>9%</div> <div>5%</div> </div>
1	B	655	<div> <div>11%</div> <div>56%</div> <div>30%</div> <div>9%</div> <div>5%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10114 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 ACTIN-LIKE PROTEIN ARP8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	623	Total	C	N	O	S	0	0	0
			5036	3219	840	963	14			
1	B	623	Total	C	N	O	S	0	0	0
			5036	3219	840	963	14			

There are 42 discrepancies between the modelled and reference sequences:

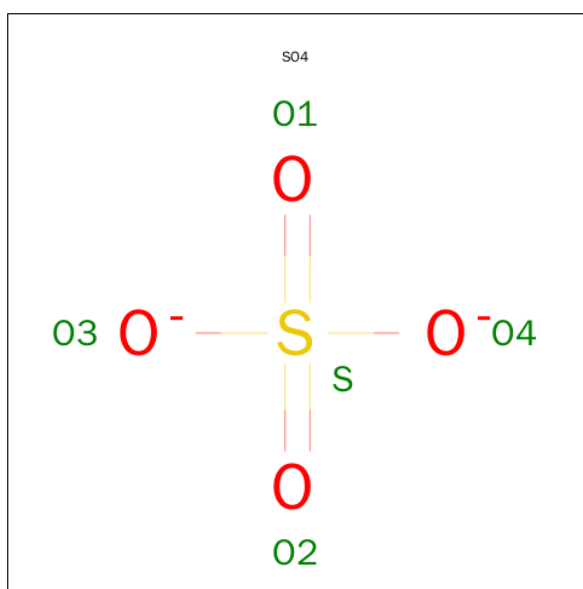
Chain	Residue	Modelled	Actual	Comment	Reference
A	227	MET	-	EXPRESSION TAG	UNP Q12386
A	228	GLY	-	EXPRESSION TAG	UNP Q12386
A	229	SER	-	EXPRESSION TAG	UNP Q12386
A	230	SER	-	EXPRESSION TAG	UNP Q12386
A	231	HIS	-	EXPRESSION TAG	UNP Q12386
A	232	HIS	-	EXPRESSION TAG	UNP Q12386
A	233	HIS	-	EXPRESSION TAG	UNP Q12386
A	234	HIS	-	EXPRESSION TAG	UNP Q12386
A	235	HIS	-	EXPRESSION TAG	UNP Q12386
A	236	HIS	-	EXPRESSION TAG	UNP Q12386
A	237	SER	-	EXPRESSION TAG	UNP Q12386
A	238	SER	-	EXPRESSION TAG	UNP Q12386
A	239	GLY	-	EXPRESSION TAG	UNP Q12386
A	240	LEU	-	EXPRESSION TAG	UNP Q12386
A	241	VAL	-	EXPRESSION TAG	UNP Q12386
A	242	PRO	-	EXPRESSION TAG	UNP Q12386
A	243	ARG	-	EXPRESSION TAG	UNP Q12386
A	244	GLY	-	EXPRESSION TAG	UNP Q12386
A	245	SER	-	EXPRESSION TAG	UNP Q12386
A	246	HIS	-	EXPRESSION TAG	UNP Q12386
A	247	MET	-	EXPRESSION TAG	UNP Q12386
B	227	MET	-	EXPRESSION TAG	UNP Q12386
B	228	GLY	-	EXPRESSION TAG	UNP Q12386
B	229	SER	-	EXPRESSION TAG	UNP Q12386
B	230	SER	-	EXPRESSION TAG	UNP Q12386

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	231	HIS	-	EXPRESSION TAG	UNP Q12386
B	232	HIS	-	EXPRESSION TAG	UNP Q12386
B	233	HIS	-	EXPRESSION TAG	UNP Q12386
B	234	HIS	-	EXPRESSION TAG	UNP Q12386
B	235	HIS	-	EXPRESSION TAG	UNP Q12386
B	236	HIS	-	EXPRESSION TAG	UNP Q12386
B	237	SER	-	EXPRESSION TAG	UNP Q12386
B	238	SER	-	EXPRESSION TAG	UNP Q12386
B	239	GLY	-	EXPRESSION TAG	UNP Q12386
B	240	LEU	-	EXPRESSION TAG	UNP Q12386
B	241	VAL	-	EXPRESSION TAG	UNP Q12386
B	242	PRO	-	EXPRESSION TAG	UNP Q12386
B	243	ARG	-	EXPRESSION TAG	UNP Q12386
B	244	GLY	-	EXPRESSION TAG	UNP Q12386
B	245	SER	-	EXPRESSION TAG	UNP Q12386
B	246	HIS	-	EXPRESSION TAG	UNP Q12386
B	247	MET	-	EXPRESSION TAG	UNP Q12386

- 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		

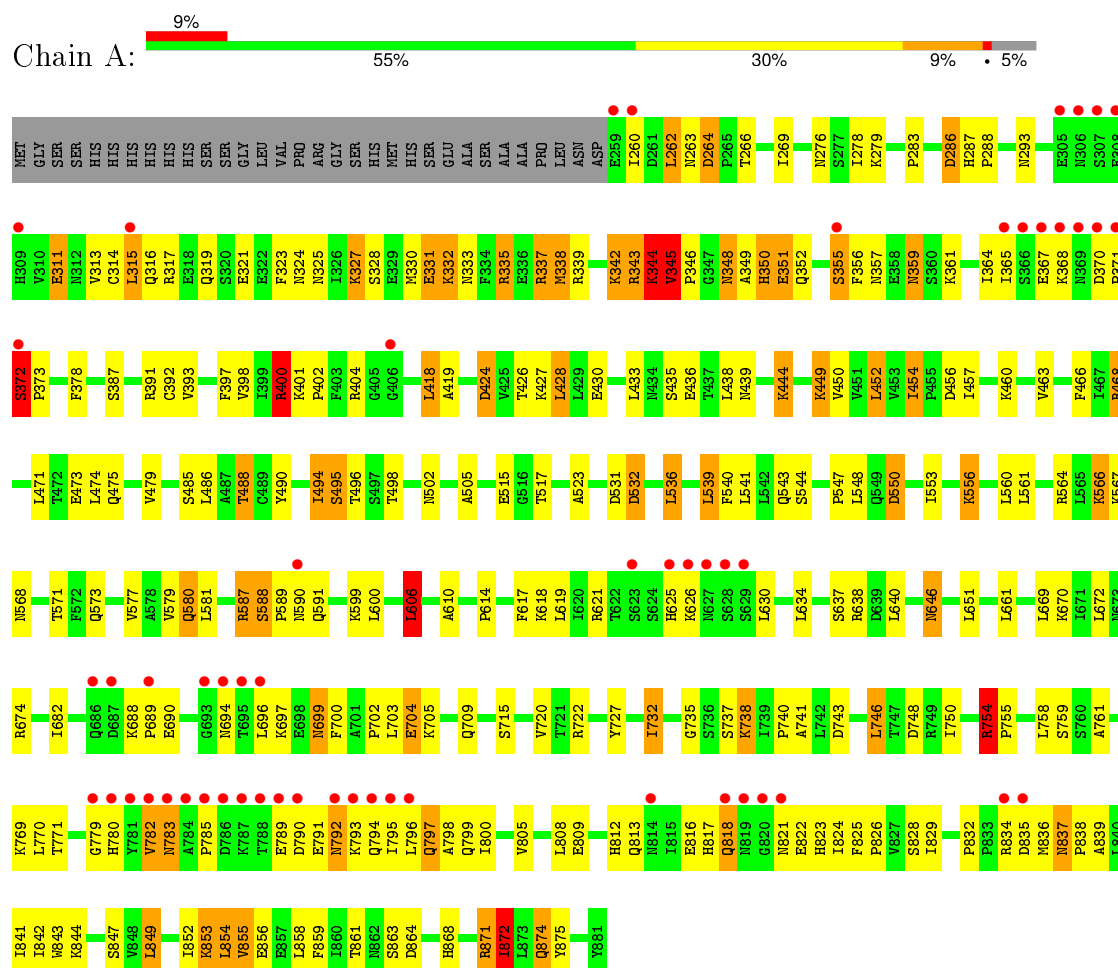
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	18	Total 18	O 18	0	0
3	B	14	Total 14	O 14	0	0

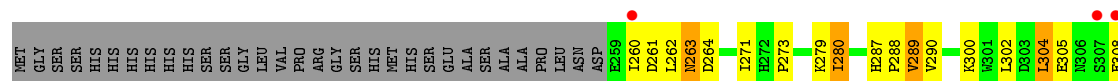
3 Residue-property plots [i](#)

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 ($\text{RSRZ} > 2$). 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: ACTIN-LIKE PROTEIN ARP8



• Molecule 1: ACTIN-LIKE PROTEIN ARP8



L840	L841	I842	W843	K844	S847	W848	L849	L852	K853	L854	W855	L858	F859	L860	W867	R871	L872	L873	Q874	F879	T880	Y881																																
L746	T747	D748	R749	I752	L753	R754	P755	K769	L770	I774	G779	H780	Y781	Y782	N783	A784	P785	D786	K787	T788	E789	D790	E791	N792	K793	Q794	I795	L796	Q797	A798	Q799	I800	K801	E806	E807	L808	N814	Q818	N821	P826	V827	S828	I829	I830	D835	K836	N837	P838	A839					
N660	L661	L665	K670	L671	L672	L673	L674	D683	Q684	L685	Q686	D687	K688	P689	E690	K691	Y692	G693	N694	T695	L696	K697	L703	E704	Q709	A712	L716	N717	L718	L719	N720	L721	R722	Y727	S728	N729	L730	L731	I732	V733	G734	G735	L736	K737	K738	L739	P740	L741	L742	I745				
F572	Q573	V577	A578	V579	Q580	L581	N586	N587	S588	P589	N590	Q591	P592	T593	E594	Y595	Y596	L600	E603	L606	A607	P608	L616	P617	P618	L619	L620	H621	T622	S623	S624	H625	K626	S629	S637	H638	D639	L640	F641	T642	N643	E644	L645	N646	S652	S656	E657	G558	N659	L560	R564	L565	K566	T571
Q475	F476	Q477	L486	A487	T488	C489	G493	T496	S497	T498	C499	V500	V501	N502	I503	A506	E507	T508	R509	I510	T517	I524	S423	T525	L526	D527	G530	D531	D532	I533	T534	R535	L536	L539	Q543	L548	K552	K556	H557	G558	N559	L560	R564	L565	K566	T571								
L383	A389	L390	R391	C392	V393	D394	E395	V398	I399	R400	R404	S407	V410	K411	S412	Y415	L418	L421	L422	I429	F430	L433	L438	N439	V440	K441	P442	K449	L452	Y453	L454	I457	K460	S461	F466	Y469	L474																	
H309	V310	V313	C314	L315	Q316	R317	E318	Q319	F323	N324	N325	S328	R339	Y340	Y341	K342	K343	K344	V345	P346	G347	N348	A349	H350	E351	V352	V353	V354	S355	F356	N357	E358	H359	S360	K361	P362	E363	I364	L365	S366	E367	K368	N369	D370	P371	S372	P373	I374	I377	F378	D379	D380	S381	K382

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	138.20 Å 87.86 Å 149.37 Å 90.00° 115.40° 90.00°	Depositor
Resolution (Å)	28.89 – 2.70 28.89 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.9 (28.89-2.70) 98.9 (28.89-2.70)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.83 (at 2.72 Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, R_{free}	0.223 , 0.286 0.221 , 0.276	Depositor DCC
R_{free} test set	2190 reflections (5.23%)	DCC
Wilson B-factor (Å ²)	70.3	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 71.4	EDS
Estimated twinning fraction	0.015 for h,-k,-h-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 44050 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10114	wwPDB-VP
Average B, all atoms (Å ²)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 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	0.90	4/5149 (0.1%)	0.94	14/6982 (0.2%)
1	B	0.75	5/5149 (0.1%)	0.86	11/6982 (0.2%)
All	All	0.83	9/10298 (0.1%)	0.90	25/13964 (0.2%)

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	2
1	B	0	1
All	All	0	3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	331	GLU	CD-OE1	15.34	1.42	1.25
1	A	331	GLU	CD-OE2	13.05	1.40	1.25
1	A	344	LYS	CE-NZ	10.53	1.75	1.49
1	A	327	LYS	CD-CE	6.42	1.67	1.51
1	B	424	ASP	CB-CG	6.02	1.64	1.51

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	424	ASP	CB-CG-OD1	10.67	127.91	118.30
1	A	424	ASP	CB-CG-OD1	-7.80	111.28	118.30
1	A	754	ARG	NE-CZ-NH1	7.54	124.07	120.30
1	A	342	LYS	N-CA-C	6.94	129.74	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	344	LYS	CD-CE-NZ	6.92	127.63	111.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	349	ALA	Peptide
1	A	587	ARG	Peptide
1	B	342	LYS	Peptide

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	5036	0	4985	206	0
1	B	5036	0	4985	159	0
2	A	10	0	0	1	0
3	A	18	0	0	2	0
3	B	14	0	0	0	0
All	All	10114	0	9970	361	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 18.

The worst 5 of 361 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:344:LYS:CE	1:A:344:LYS:NZ	1.75	1.46
1:A:785:PRO:HB2	1:A:793:LYS:HD2	1.21	1.19
1:A:488:THR:HG22	1:A:842:ILE:CD1	1.79	1.11
1:A:315:LEU:HD23	1:A:316:GLN:N	1.66	1.09
1:A:315:LEU:CG	1:A:316:GLN:H	1.71	1.03

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	621/655 (95%)	554 (89%)	64 (10%)	3 (0%)	34	63
1	B	621/655 (95%)	544 (88%)	69 (11%)	8 (1%)	15	37
All	All	1242/1310 (95%)	1098 (88%)	133 (11%)	11 (1%)	21	49

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	346	PRO
1	A	372	SER
1	B	373	PRO
1	B	688	LYS
1	B	371	PRO

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	569/595 (96%)	464 (82%)	105 (18%)	2	5
1	B	569/595 (96%)	460 (81%)	109 (19%)	2	4
All	All	1138/1190 (96%)	924 (81%)	214 (19%)	2	5

5 of 214 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	828	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	358	GLU
1	B	780	HIS
1	A	853	LYS
1	B	287	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	794	GLN
1	B	316	GLN
1	B	818	GLN
1	A	818	GLN
1	B	369	ASN

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 carbohydrates 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
2	SO4	A	1882	-	4,4,4	0.49	0	6,6,6	0.49	0
2	SO4	A	1883	-	4,4,4	0.42	0	6,6,6	0.43	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	1882	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1883	-	-	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1883	SO4	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	623/655 (95%)	0.43	56 (8%) 12 9	32, 64, 146, 240	0
1	B	623/655 (95%)	0.64	71 (11%) 7 5	41, 85, 174, 263	0
All	All	1246/1310 (95%)	0.54	127 (10%) 9 6	32, 74, 162, 263	0

The worst 5 of 127 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	786	ASP	15.3
1	B	373	PRO	14.0
1	B	689	PRO	12.4
1	B	372	SER	11.6
1	A	782	VAL	11.5

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 carbohydrates 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. 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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	SO4	A	1882	5/5	0.96	0.22	1.26	62,63,64,67	0
2	SO4	A	1883	5/5	0.95	0.11	-1.87	70,73,73,73	0

6.5 Other polymers [i](#)

There are no such residues in this entry.