



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:37 PM GMT

PDB ID : 4A0F
Title : Structure of selenomethionine substituted bifunctional DAPA aminotransferase-dethiobiotin synthetase from *Arabidopsis thaliana* in its apo form.
Authors : Cobessi, D.; Dumas, R.; Pautre, V.; Meinguet, C.; Ferrer, J.L.; Alban, C.
Deposited on : 2011-09-09
Resolution : 2.71 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

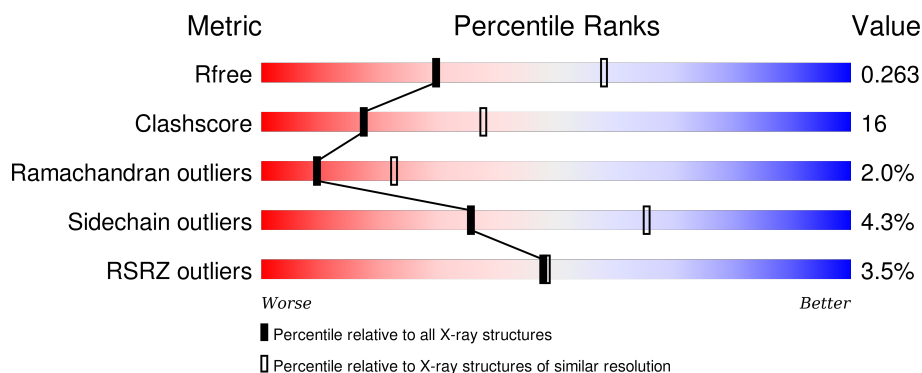
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


The reported resolution of this entry is 2.71 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	2439 (2.74-2.70)
Clashscore	102246	2771 (2.74-2.70)
Ramachandran outliers	100387	2726 (2.74-2.70)
Sidechain outliers	100360	2727 (2.74-2.70)
RSRZ outliers	91569	2443 (2.74-2.70)

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	831	 4% 64% 23% • 10%
1	B	831	 2% 60% 28% • 10%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SO4	A	1809	-	-	-	X
3	SO4	A	1810	-	-	-	X
3	SO4	B	1810	-	-	X	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 11317 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 ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	746	Total	C	N	O	S	Se	0	1	0
			5568	3562	937	1037	16	16			
1	B	750	Total	C	N	O	S	Se	0	0	0
			5619	3598	944	1046	16	15			

There are 42 discrepancies between the modelled and reference sequences:

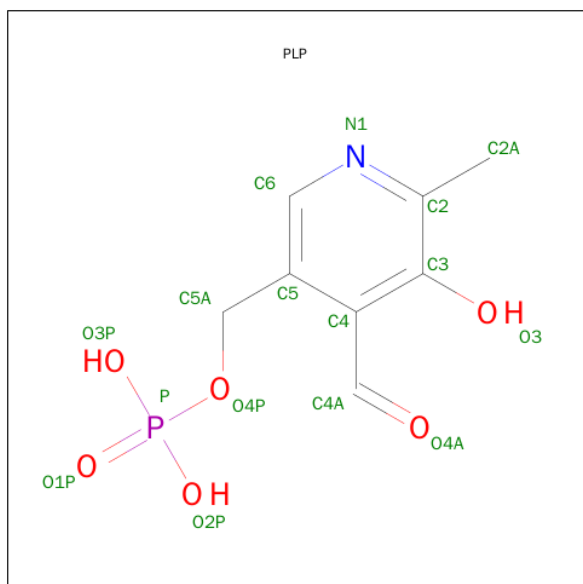
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	GLY	-	EXPRESSION TAG	UNP B0F481
A	-18	SER	-	EXPRESSION TAG	UNP B0F481
A	-17	SER	-	EXPRESSION TAG	UNP B0F481
A	-16	HIS	-	EXPRESSION TAG	UNP B0F481
A	-15	HIS	-	EXPRESSION TAG	UNP B0F481
A	-14	HIS	-	EXPRESSION TAG	UNP B0F481
A	-13	HIS	-	EXPRESSION TAG	UNP B0F481
A	-12	HIS	-	EXPRESSION TAG	UNP B0F481
A	-11	HIS	-	EXPRESSION TAG	UNP B0F481
A	-10	SER	-	EXPRESSION TAG	UNP B0F481
A	-9	SER	-	EXPRESSION TAG	UNP B0F481
A	-8	GLY	-	EXPRESSION TAG	UNP B0F481
A	-7	LEU	-	EXPRESSION TAG	UNP B0F481
A	-6	VAL	-	EXPRESSION TAG	UNP B0F481
A	-5	PRO	-	EXPRESSION TAG	UNP B0F481
A	-4	ARG	-	EXPRESSION TAG	UNP B0F481
A	-3	GLY	-	EXPRESSION TAG	UNP B0F481
A	-2	SER	-	EXPRESSION TAG	UNP B0F481
A	-1	HIS	-	EXPRESSION TAG	UNP B0F481
A	0	MSE	-	EXPRESSION TAG	UNP B0F481
A	326	TYR	PHE	ENGINEERED MUTATION	UNP B0F481
B	-19	GLY	-	EXPRESSION TAG	UNP B0F481
B	-18	SER	-	EXPRESSION TAG	UNP B0F481
B	-17	SER	-	EXPRESSION TAG	UNP B0F481

Continued on next page...

Continued from previous page...

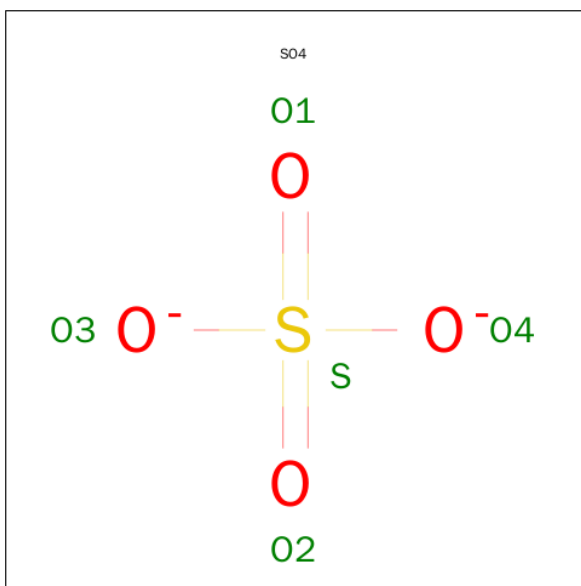
Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	HIS	-	EXPRESSION TAG	UNP B0F481
B	-15	HIS	-	EXPRESSION TAG	UNP B0F481
B	-14	HIS	-	EXPRESSION TAG	UNP B0F481
B	-13	HIS	-	EXPRESSION TAG	UNP B0F481
B	-12	HIS	-	EXPRESSION TAG	UNP B0F481
B	-11	HIS	-	EXPRESSION TAG	UNP B0F481
B	-10	SER	-	EXPRESSION TAG	UNP B0F481
B	-9	SER	-	EXPRESSION TAG	UNP B0F481
B	-8	GLY	-	EXPRESSION TAG	UNP B0F481
B	-7	LEU	-	EXPRESSION TAG	UNP B0F481
B	-6	VAL	-	EXPRESSION TAG	UNP B0F481
B	-5	PRO	-	EXPRESSION TAG	UNP B0F481
B	-4	ARG	-	EXPRESSION TAG	UNP B0F481
B	-3	GLY	-	EXPRESSION TAG	UNP B0F481
B	-2	SER	-	EXPRESSION TAG	UNP B0F481
B	-1	HIS	-	EXPRESSION TAG	UNP B0F481
B	0	MSE	-	EXPRESSION TAG	UNP B0F481
B	326	TYR	PHE	ENGINEERED MUTATION	UNP B0F481

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: C₈H₁₀NO₆P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	B	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	50	Total	O	0	0
			50	50		
4	B	30	Total	O	0	0
			30	30		

T49	T49	H38	L262	G373	V464	H562	L645	T79
K50	K50	E139	V266	F374	K465	LEU	M650	ALA
L51	L51	I142	L382	L382	V466	GLN	V651	S757
Q58	Q58	H145	L269	A383	I467	HIS	P652	Y759
T59	T59	H146	D277	R384	R470	SER	T659	L763
G60	G60	L146	P278	E385	H474	GLY	D660	L764
F61	F61	R150	S279	K386	L478	VAL	A661	L765
P62	P62	L168	D280	G387	L478	ARG	V662	M766
S67	S67	L168	D281	Y388	G479	GLN	F666	E769
F71	F71	M172	L282	R392	A480	SER	S667	T774
S72	S72	GLU	V287	R395	K481	H574	G668	N779
K73	K73	CYS	F293	V396	E482	V575	D669	L783
L74	L74	GLY	F398	R397	A485	A576	S670	S790
L79	L79	VAL	L296	F399	P486	L578	K671	L797
R80	R80	LVS	K297	P399	S487	L579	L672	L798
R81	R81	SER	M300	V402	P488	I580	L675	L801
I85	I85	GLU	V301	Y403	G491	V583	S680	R804
S86	S86	LVS	L302	E411	H497	I584	Y681	L805
I87	I87	D182	A315	L414	R501	M590	A685	G806
S88	S88	L183	K316	V417	P508	H591	M686	F808
N89	N89	C185	L317	S428	T509	H592	G687	ASN
S90	S90	L186	V321	D429	V510	D594	G688	ARG
V91	V91	V187	P325	M430	H513	P595	T708	THR
L92	L92	P196	T327	T433	G515	F597	S709	
H93	H93	L206	K330	I437	I519	H598	Q710	
S94	S94	T207	E333	A438	S520	R599	L714	
P11	P11	R208	K336	M441	F525	L610	L717	
L12	L12	P209	V337	A442	SER	F611	W718	
M13	M13	P97	I340	F443	GLU	V612	D719	
H14	H14	S101	R221	R444	ILE	L613	V723	
L18	L18	F210	L222	K445	ALA	F614	Q724	
S21	S21	R211	G223	P446	PRO	D615	H729	
T24	T24	L104	G224	C447	GLU	E616	S730	
S25	S25	ASN	I225	V448	TTR	V617	A731	
L26	L26	VAL	I229	H450	GLY	M622	V737	
G27	G27	SER	E233	M451	T534	R623	I738	
K28	K28	E110	R236	CYS	F535	L624	L741	
T29	T29	M113	E129	ALA	T536	T628	L746	
L30	L30	C114	L131	THR	S537	T630	LYS	
V31	V31	R119	L252	GLU	R538	K636	ALA	
G34	G34	A127	P257	GLU	R546	T639	ASP	
T35	T35	P128	T134	LYS	D647	L639	ALA	
A36	A36	E129	Y136	HIS	S556	A643	SER	
L41	L41	L130	A137	V463	I559	K644	ASN	
Q42	Q42	L131			S560		GLY	
Q43	Q43	PRO						
PRO	PRO	SER						
SER	SER	SER						
S47	S47							
A48	A48							

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	233.67Å 75.97Å 88.63Å 90.00° 109.20° 90.00°	Depositor
Resolution (Å)	41.85 – 2.71 41.85 – 2.71	Depositor EDS
% Data completeness (in resolution range)	99.1 (41.85-2.71) 99.2 (41.85-2.71)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.82 (at 2.73Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.200 , 0.261 0.200 , 0.263	Depositor DCC
R_{free} test set	1990 reflections (5.30%)	DCC
Wilson B-factor (Å ²)	48.0	Xtriage
Anisotropy	0.463	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 46.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 39567 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	11317	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.71% 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, PLP

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	0.44	0/5681	0.60	1/7722 (0.0%)
1	B	0.45	0/5731	0.60	0/7787
All	All	0.45	0/11412	0.60	1/15509 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	44	PRO	N-CA-CB	6.07	110.59	103.30

There are no chirality outliers.

There are no planarity outliers.

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	5568	0	5367	172	0
1	B	5619	0	5425	206	0
2	A	15	0	6	2	0
2	B	15	0	6	2	0
3	A	10	0	0	1	0
3	B	10	0	0	3	0
4	A	50	0	0	1	0
4	B	30	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	11317	0	10804	355	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 16.

The worst 5 of 355 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:21:SER:HB2	1:A:28:LYS:HD3	1.41	1.01
1:B:21:SER:HB2	1:B:28:LYS:HD3	1.41	1.00
1:A:510:VAL:HB	1:A:592:MSE:HE2	1.48	0.96
1:B:510:VAL:HB	1:B:592:MSE:HE2	1.51	0.91
1:B:766:MSE:HE2	1:B:804:ARG:HD2	1.57	0.86

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	A	733/831 (88%)	669 (91%)	50 (7%)	14 (2%)	10	24
1	B	734/831 (88%)	672 (92%)	46 (6%)	16 (2%)	8	20
All	All	1467/1662 (88%)	1341 (91%)	96 (6%)	30 (2%)	9	22

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	222	LEU
1	A	357	SER
1	B	222	LEU
1	B	357	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	730	SER

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	575/688 (84%)	552 (96%)	23 (4%)	38	68
1	B	583/688 (85%)	556 (95%)	27 (5%)	33	63
All	All	1158/1376 (84%)	1108 (96%)	50 (4%)	35	65

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	790	SER
1	B	131	LEU
1	B	680	SER
1	B	9	HIS
1	B	67	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	450	HIS
1	B	450	HIS
1	B	484	GLN
1	B	574	HIS

5.3.3 RNA ⓘ

There are no RNA molecules 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$
2	PLP	A	1644	1	15,15,16	1.78	1 (6%)	21,22,23	1.99	5 (23%)
3	SO4	A	1809	-	4,4,4	0.21	0	6,6,6	0.34	0
3	SO4	A	1810	-	4,4,4	0.18	0	6,6,6	0.36	0
2	PLP	B	1644	1	15,15,16	1.82	3 (20%)	21,22,23	1.89	4 (19%)
3	SO4	B	1809	-	4,4,4	0.28	0	6,6,6	0.23	0
3	SO4	B	1810	-	4,4,4	0.20	0	6,6,6	0.30	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	PLP	A	1644	1	-	0/6/6/8	0/1/1/1
3	SO4	A	1809	-	-	0/0/0/0	0/0/0/0
3	SO4	A	1810	-	-	0/0/0/0	0/0/0/0
2	PLP	B	1644	1	-	0/6/6/8	0/1/1/1
3	SO4	B	1809	-	-	0/0/0/0	0/0/0/0
3	SO4	B	1810	-	-	0/0/0/0	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1644	PLP	O3-C3	-5.59	1.23	1.37
2	B	1644	PLP	O3-C3	-5.37	1.24	1.37
2	B	1644	PLP	C6-N1	2.33	1.39	1.34
2	B	1644	PLP	C2-N1	2.40	1.39	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1644	PLP	O3P-P-O4P	-2.36	99.77	106.56
2	A	1644	PLP	O3P-P-O4P	-2.31	99.91	106.56
2	A	1644	PLP	C5-C6-N1	-2.25	119.95	123.86
2	B	1644	PLP	O3P-P-O2P	2.18	115.69	107.38
2	A	1644	PLP	O3P-P-O2P	2.23	115.88	107.38

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1644	PLP	2	0
3	A	1809	SO4	1	0
2	B	1644	PLP	2	0
3	B	1810	SO4	3	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	731/831 (87%)	0.01	32 (4%) 38 37	28, 54, 95, 119	0
1	B	735/831 (88%)	-0.03	19 (2%) 59 60	27, 56, 89, 114	0
All	All	1466/1662 (88%)	-0.01	51 (3%) 48 48	27, 55, 92, 119	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	551	LEU	5.8
1	A	449	ASP	4.8
1	A	574	HIS	4.3
1	A	762	SER	4.3
1	A	512	LEU	4.2

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
3	SO4	A	1809	5/5	0.93	0.22	3.36	48,57,85,89	5
3	SO4	B	1810	5/5	0.89	0.25	3.14	57,64,91,114	5
3	SO4	A	1810	5/5	0.96	0.21	2.76	31,34,44,57	5
3	SO4	B	1809	5/5	0.95	0.17	1.53	52,63,81,83	5
2	PLP	B	1644	15/16	0.97	0.21	0.37	34,47,56,60	0
2	PLP	A	1644	15/16	0.98	0.20	0.19	44,48,58,61	0

6.5 Other polymers [i](#)

There are no such residues in this entry.