



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 01:31 am 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](mailto: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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

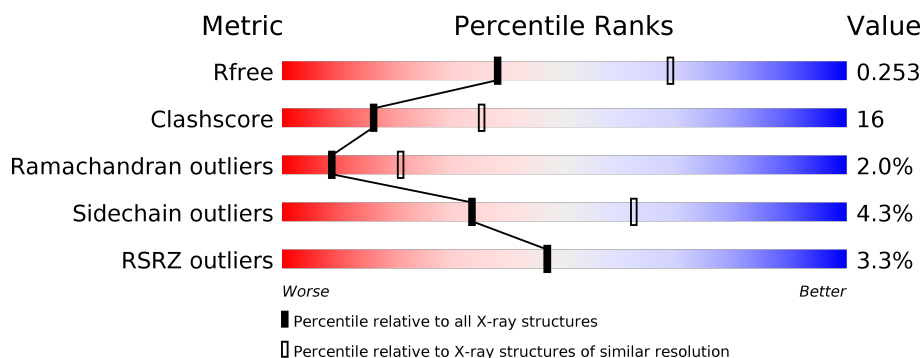
# 1 Overall quality at a glance

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))
$R_{free}$	100719	2649 (2.74-2.70)
Clashscore	112137	2993 (2.74-2.70)
Ramachandran outliers	110173	2946 (2.74-2.70)
Sidechain outliers	110143	2947 (2.74-2.70)
RSRZ outliers	101464	2665 (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  $\geq 3$ , 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	
1	B	831	

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 [i](#)

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:

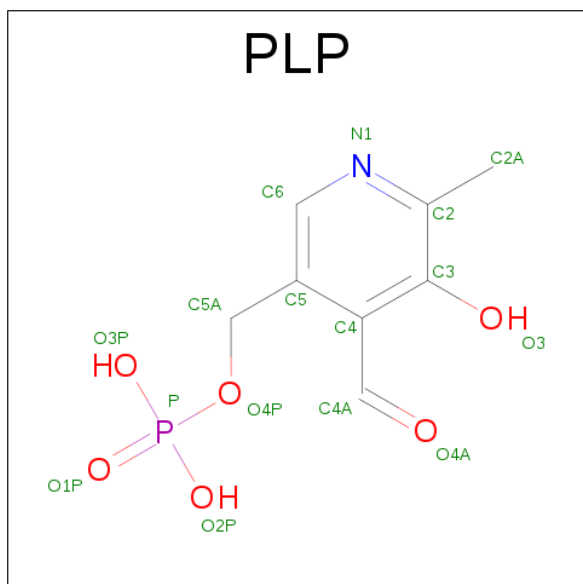
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	GLY	-	EXPRESSION TAG	UNP B0F481
A	-18	SER	-	EXPRESSION TAG	UNP B0F481
A	-17	SER	-	EXPRESSION TAG	UNP B0F481
A	-16	HIS	-	EXPRESSION TAG	UNP B0F481
A	-15	HIS	-	EXPRESSION TAG	UNP B0F481
A	-14	HIS	-	EXPRESSION TAG	UNP B0F481
A	-13	HIS	-	EXPRESSION TAG	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...

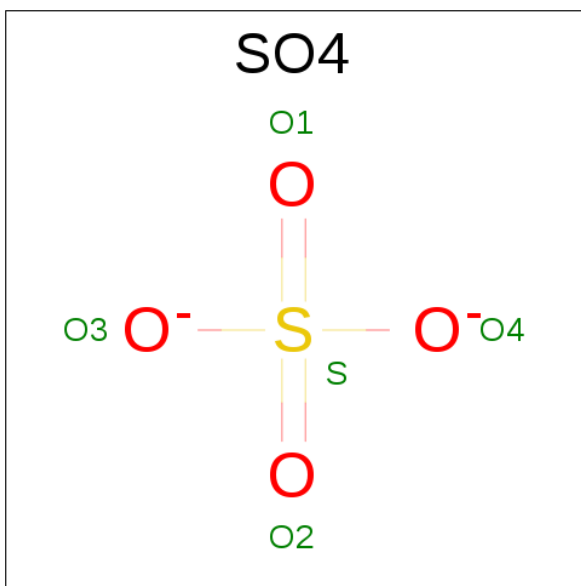
Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	HIS	-	EXPRESSION TAG	UNP B0F481
B	-15	HIS	-	EXPRESSION TAG	UNP B0F481
B	-14	HIS	-	EXPRESSION TAG	UNP B0F481
B	-13	HIS	-	EXPRESSION TAG	UNP B0F481
B	-12	HIS	-	EXPRESSION TAG	UNP B0F481
B	-11	HIS	-	EXPRESSION TAG	UNP B0F481
B	-10	SER	-	EXPRESSION TAG	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:  $C_8H_{10}NO_6P$ ).



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: O<sub>4</sub>S).



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		



Tyr	L645	H562	V464	G373	L262	H138	T49
ALA	W650	LEU	K465	P374	V266	E139	K50
S757	V651	GLN	V466	L382	I142	I142	L51
L758	P652	GLY	I467	A383	H145	T59	Q58
Y759	T659	HIS	R470	R384	L146	G60	F61
L763	D660	GLY	H474	K385	D277	F61	P62
L764	A661	VAL	L478	K386	P278	R150	S67
I765	V662	ARG	G479	G387	D280	L168	F71
M766	F666	SER	A480	Y388	D281	M172	S72
E769	S667	GLN	M481	R392	L282	GLU	K73
T774	S668	A573	E482	H395	V287	CYS	L74
M779	D669	H574	A485	V396	F293	GLY	L79
L783	S670	G576	A486	K397	L296	VAL	R80
S790	K671	L578	P486	F398	K297	LYS	R81
L797	L672	I579	S487	P399	M300	SER	I85
L798	L675	I580	P488	V402	V301	GLU	I87
L801	L681	I584	G491	E411	L302	LYS	S88
R804	A685	M590	W497	L414	A315	D182	N89
L805	H686	H591	B501	V417	K316	L183	S90
G806	G687	M592	P508	S428	L317	L184	V91
E807	C688	V593	V510	M430	V321	C185	H93
F808	T708	P595	S513	D429	P325	L187	S94
ASN	S709	L596	H514	M433	Y326	L206	S95
ARG	Q710	F597	G515	T433	I327	Y207	L96
THR	L714	G598	G519	L437	K330	R208	P97
	L717	R599	I519	A438	F210	P209	S101
	W718	V600	S520	A441	R211	R211	L104
	D719	I610	F525	A442	G218	ASN	ASN
	V723	P611	SER	F443	D219	VAL	VAL
	Q724	V612	GLU	R444	G220	GLU	GLU
	H729	I613	ILE	K445	R221	VAL	VAL
S730	S730	F614	ALA	F446	L222	SER	SER
A731	A731	D615	PRO	C447	G223	E110	E110
		E616	GLU	V448	G224	M113	M113
		V617	TYR	D449	I225	C114	C114
		W622	GLY	H450	I229	R119	R119
		R623	T534	F452	D355	A127	A127
		L624	F535	N451	N356	P128	P128
		T628	S537	CYS	S357	L129	L129
		T629	R538	GLU	S360	L130	L130
		T630	R546	THR	Q361	L131	L131
		K636	D547	GLU	D241	T134	T134
		I639	S556	GLU	L252	L135	L135
		A643	L559	LYS	P257	Y136	Y136
		K644	S560	HIS		A137	A137
			K561	L462			
				V463			



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.82 (at 2.73Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.200 , 0.261 0.189 , 0.253	Depositor DCC
$R_{free}$ test set	1989 reflections (5.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	48.0	Xtriage
Anisotropy	0.463	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 46.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	11317	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	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.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: 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

### 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	733/831 (88%)	669 (91%)	50 (7%)	14 (2%)	9	22
1	B	734/831 (88%)	672 (92%)	46 (6%)	16 (2%)	8	19
All	All	1467/1662 (88%)	1341 (91%)	96 (6%)	30 (2%)	9	21

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%)	36	65
1	B	583/688 (85%)	556 (95%)	27 (5%)	31	60
All	All	1158/1376 (84%)	1108 (96%)	50 (4%)	33	62

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 [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](#)

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	2 (13%)	20,22,23	1.77	5 (25%)
3	SO4	A	1809	-	4,4,4	0.28	0	6,6,6	0.29	0
3	SO4	A	1810	-	4,4,4	0.18	0	6,6,6	0.41	0
2	PLP	B	1644	1	15,15,16	1.82	3 (20%)	20,22,23	1.68	4 (20%)
3	SO4	B	1809	-	4,4,4	0.30	0	6,6,6	0.17	0
3	SO4	B	1810	-	4,4,4	0.30	0	6,6,6	0.27	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 (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1644	PLP	O3-C3	-5.67	1.23	1.37
2	B	1644	PLP	O3-C3	-5.44	1.24	1.37
2	A	1644	PLP	C2-N1	2.07	1.38	1.33
2	B	1644	PLP	C6-N1	2.34	1.39	1.34
2	B	1644	PLP	C2-N1	2.54	1.39	1.33

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	-2.62	99.77	106.73
2	A	1644	PLP	O3P-P-O4P	-2.56	99.91	106.73
2	A	1644	PLP	C5-C6-N1	-2.32	119.95	123.87
2	B	1644	PLP	O3P-P-O2P	2.00	115.69	107.61
2	A	1644	PLP	O3P-P-O2P	2.05	115.88	107.61

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	731/831 (87%)	-0.03	30 (4%) 38 37	28, 54, 95, 119	0
1	B	735/831 (88%)	-0.06	18 (2%) 59 60	27, 56, 89, 114	0
All	All	1466/1662 (88%)	-0.04	48 (3%) 47 47	27, 55, 92, 119	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	551	LEU	5.8
1	A	449	ASP	4.6
1	A	574	HIS	4.2
1	A	762	SER	4.2
1	A	512	LEU	4.1

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	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.