



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 09:34 PM GMT

PDB ID : 5AE1  
Title : Ether Lipid-Generating Enzyme AGPS in complex with inhibitor ZINC69435460  
Authors : Piano, V.; Benjamin, D.I.; Valente, S.; Nenci, S.; Marrocco, B.; Mai, A.; Aliverti, A.; Nomura, D.K.; Mattevi, A.  
Deposited on : 2015-08-25  
Resolution : 2.10 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

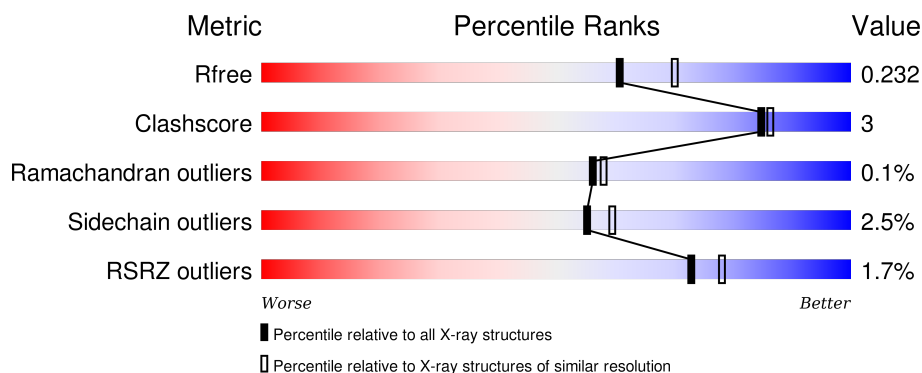
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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	658	<div> <div>2%</div> <div>78%</div> <div>7%</div> <div>15%</div> </div>
1	B	658	<div> <div>%</div> <div>75%</div> <div>7%</div> <div>18%</div> </div>
1	C	658	<div> <div>%</div> <div>76%</div> <div>8%</div> <div>15%</div> </div>
1	D	658	<div> <div>2%</div> <div>79%</div> <div>6%</div> <div>14%</div> </div>

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
4	SO4	B	1659	-	-	-	X
4	SO4	D	1659	-	-	-	X

## 2 Entry composition [i](#)

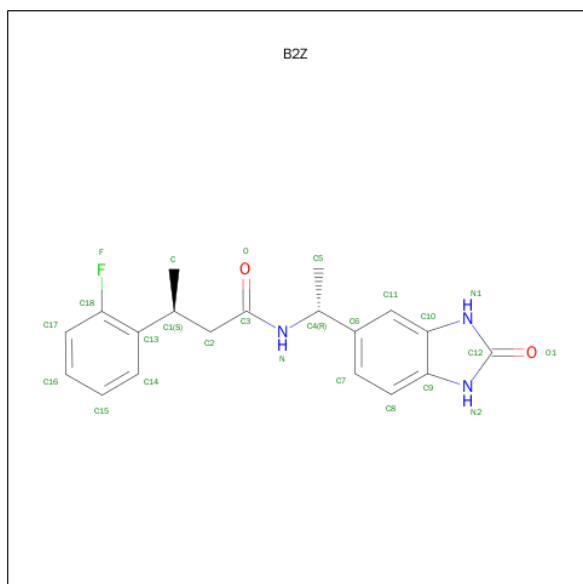
There are 5 unique types of molecules in this entry. The entry contains 18358 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 ALKYLDIHYDROXYACETONEPHOSPHATE SYNTHASE, PEROXISOMAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	560	Total	C	N	O	S	0	1	0
			4410	2810	761	815	24			
1	B	539	Total	C	N	O	S	0	0	0
			4254	2703	740	787	24			
1	C	559	Total	C	N	O	S	0	2	0
			4407	2800	761	822	24			
1	D	564	Total	C	N	O	S	0	0	0
			4395	2786	765	820	24			

- Molecule 2 is (3-(2-FLUOROPHENYL)-N-(1-(2-OXO-2,3-DIHYDRO-1H-BENZO[D]IMIDAZOL-5-YL)ETHYL)BUTANAMIDE) (three-letter code: B2Z) (formula: C<sub>19</sub>H<sub>20</sub>FN<sub>3</sub>O<sub>2</sub>).



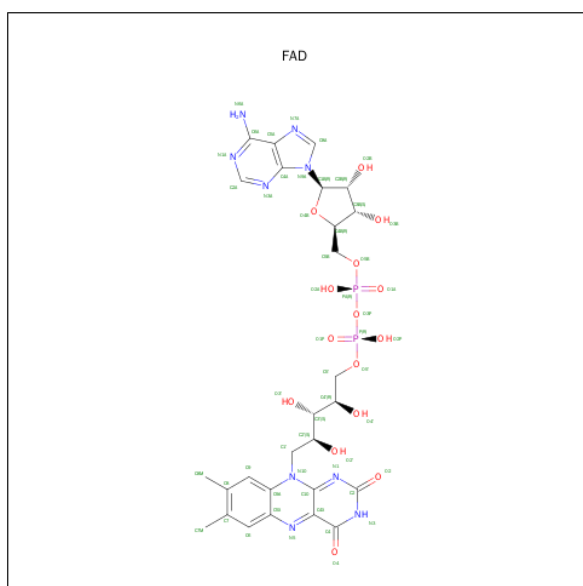
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			25	19	1	3	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	F	N	O	0	0
			25	19	1	3	2		
2	C	1	Total	C	F	N	O	0	0
			25	19	1	3	2		
2	D	1	Total	C	F	N	O	0	0
			25	19	1	3	2		

- Molecule 3 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula:  $C_{27}H_{33}N_9O_{15}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
3	B	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
3	C	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
3	D	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	150	Total	O	0	0
			150	150		
5	B	123	Total	O	0	0
			123	123		
5	C	171	Total	O	0	0
			171	171		
5	D	116	Total	O	0	0
			116	116		



[illegible]



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.79Å 99.02Å 107.02Å 90.64° 89.96° 95.49°	Depositor
Resolution (Å)	107.01 – 2.10 40.14 – 2.10	Depositor EDS
% Data completeness (in resolution range)	93.6 (107.01-2.10) 93.3 (40.14-2.10)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.73 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0124	Depositor
R, $R_{free}$	0.186 , 0.229 0.193 , 0.232	Depositor DCC
$R_{free}$ test set	1484 reflections (1.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	26.9	Xtriage
Anisotropy	0.026	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 45.1	EDS
Estimated twinning fraction	0.013 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 135722 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	18358	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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, 0.375 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: B2Z, FAD, 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	0.80	2/4514 (0.0%)	0.92	11/6103 (0.2%)
1	B	0.80	0/4350	0.87	7/5882 (0.1%)
1	C	0.83	0/4515	0.93	14/6110 (0.2%)
1	D	0.80	0/4493	0.90	7/6084 (0.1%)
All	All	0.81	2/17872 (0.0%)	0.90	39/24179 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	315	SER	CB-OG	-5.66	1.34	1.42
1	A	174	SER	CB-OG	-5.39	1.35	1.42

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	515	ARG	NE-CZ-NH2	-13.12	113.74	120.30
1	A	317	ARG	NE-CZ-NH1	10.85	125.73	120.30
1	C	317	ARG	NE-CZ-NH1	10.27	125.44	120.30
1	D	317	ARG	NE-CZ-NH1	9.51	125.05	120.30
1	C	515	ARG	NE-CZ-NH1	8.80	124.70	120.30

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	4410	0	4328	21	0
1	B	4254	0	4191	21	0
1	C	4407	0	4318	36	0
1	D	4395	0	4270	18	0
2	A	25	0	20	3	0
2	B	25	0	20	3	0
2	C	25	0	20	3	0
2	D	25	0	20	3	0
3	A	53	0	31	2	0
3	B	53	0	31	2	0
3	C	53	0	31	2	0
3	D	53	0	31	3	0
4	A	5	0	0	0	0
4	B	5	0	0	0	0
4	C	5	0	0	0	0
4	D	5	0	0	0	0
5	A	150	0	0	2	0
5	B	123	0	0	2	0
5	C	171	0	0	9	0
5	D	116	0	0	1	0
All	All	18358	0	17311	101	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 3.

The worst 5 of 101 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:C:411:GLN:HG2	5:C:2127:HOH:O	1.65	0.95
2:A:888:B2Z:H51C	3:A:999:FAD:HM82	1.58	0.84
1:D:300:HIS:ND1	1:D:302:PRO:HD3	1.99	0.78
1:C:150:LEU:HD22	1:C:180:ASP:HA	1.74	0.70
1:C:386:GLU:CD	5:C:2120:HOH:O	2.31	0.68

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	555/658 (84%)	543 (98%)	10 (2%)	2 (0%)	39	37
1	B	533/658 (81%)	522 (98%)	11 (2%)	0	100	100
1	C	557/658 (85%)	547 (98%)	10 (2%)	0	100	100
1	D	558/658 (85%)	549 (98%)	9 (2%)	0	100	100
All	All	2203/2632 (84%)	2161 (98%)	40 (2%)	2 (0%)	56	58

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	198	ARG
1	A	586	ARG

### 5.3.2 Protein sidechains [i](#)

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	471/545 (86%)	462 (98%)	9 (2%)	65	70
1	B	457/545 (84%)	446 (98%)	11 (2%)	57	61
1	C	474/545 (87%)	460 (97%)	14 (3%)	48	51
1	D	465/545 (85%)	452 (97%)	13 (3%)	51	55
All	All	1867/2180 (86%)	1820 (98%)	47 (2%)	55	59

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	128	PHE
1	C	379	ILE
1	D	510	VAL
1	C	144	THR
1	C	406	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	430	HIS
1	D	388	GLN
1	C	425	GLN
1	A	424	GLN
1	C	423	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](#)

12 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
4	SO4	A	1659	-	4,4,4	0.43	0	6,6,6	0.69	0
2	B2Z	A	888	-	24,27,27	1.22	4 (16%)	29,38,38	1.41	4 (13%)
3	FAD	A	999	-	52,58,58	1.54	6 (11%)	52,89,89	2.74	18 (34%)
4	SO4	B	1659	-	4,4,4	1.01	0	6,6,6	0.77	0
2	B2Z	B	888	-	24,27,27	1.23	4 (16%)	29,38,38	1.74	8 (27%)
3	FAD	B	999	-	52,58,58	1.42	9 (17%)	52,89,89	2.55	15 (28%)
4	SO4	C	1659	-	4,4,4	0.55	0	6,6,6	0.60	0
2	B2Z	C	888	-	24,27,27	1.03	2 (8%)	29,38,38	1.07	2 (6%)
3	FAD	C	999	-	52,58,58	1.45	8 (15%)	52,89,89	2.78	16 (30%)
4	SO4	D	1659	-	4,4,4	1.09	0	6,6,6	0.80	0
2	B2Z	D	888	-	24,27,27	1.25	3 (12%)	29,38,38	1.55	7 (24%)
3	FAD	D	999	-	52,58,58	1.09	5 (9%)	52,89,89	2.83	17 (32%)

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
4	SO4	A	1659	-	-	0/0/0/0	0/0/0/0
2	B2Z	A	888	-	-	0/16/16/16	0/3/3/3
3	FAD	A	999	-	-	0/30/50/50	0/6/6/6
4	SO4	B	1659	-	-	0/0/0/0	0/0/0/0
2	B2Z	B	888	-	-	0/16/16/16	0/3/3/3
3	FAD	B	999	-	-	0/30/50/50	0/6/6/6
4	SO4	C	1659	-	-	0/0/0/0	0/0/0/0
2	B2Z	C	888	-	-	0/16/16/16	0/3/3/3
3	FAD	C	999	-	-	0/30/50/50	0/6/6/6
4	SO4	D	1659	-	-	0/0/0/0	0/0/0/0
2	B2Z	D	888	-	-	0/16/16/16	0/3/3/3
3	FAD	D	999	-	-	0/30/50/50	0/6/6/6

The worst 5 of 41 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	999	FAD	C2-N1	-3.52	1.30	1.38
3	B	999	FAD	C2-N3	-2.95	1.32	1.38
3	C	999	FAD	PA-O2A	-2.77	1.43	1.55
2	B	888	B2Z	C13-C1	-2.75	1.47	1.52
3	A	999	FAD	C6-C5X	-2.57	1.37	1.41

The worst 5 of 87 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	999	FAD	C4-C4X-C10	-7.36	115.23	119.94
3	D	999	FAD	C4-C4X-C10	-7.27	115.29	119.94
3	B	999	FAD	C4-C4X-C10	-6.67	115.67	119.94
3	A	999	FAD	N3A-C2A-N1A	-6.54	123.73	128.87
3	D	999	FAD	N3A-C2A-N1A	-6.44	123.82	128.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	888	B2Z	3	0
3	A	999	FAD	2	0
2	B	888	B2Z	3	0
3	B	999	FAD	2	0
2	C	888	B2Z	3	0
3	C	999	FAD	2	0
2	D	888	B2Z	3	0
3	D	999	FAD	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	560/658 (85%)	-0.12	11 (1%) 68 73	15, 31, 57, 74	0
1	B	539/658 (81%)	-0.18	7 (1%) 79 84	17, 30, 53, 72	0
1	C	559/658 (84%)	-0.16	8 (1%) 78 82	15, 28, 50, 77	1 (0%)
1	D	564/658 (85%)	-0.12	11 (1%) 68 73	16, 31, 57, 78	0
All	All	2222/2632 (84%)	-0.14	37 (1%) 73 78	15, 30, 55, 78	1 (0%)

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	159	ILE	4.8
1	D	453	THR	4.5
1	C	455	PHE	3.8
1	C	458	PHE	3.3
1	B	494	LEU	3.2

### 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, 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
4	SO4	D	1659	5/5	0.93	0.26	10.58	42,50,58,76	0
4	SO4	B	1659	5/5	0.92	0.21	2.53	51,52,58,63	0
2	B2Z	A	888	25/25	0.94	0.13	0.61	30,34,57,65	0
4	SO4	C	1659	5/5	0.97	0.12	0.53	35,45,53,58	0
2	B2Z	D	888	25/25	0.94	0.14	0.43	31,38,51,56	0
2	B2Z	C	888	25/25	0.93	0.13	0.26	23,32,39,53	0
2	B2Z	B	888	25/25	0.94	0.13	0.25	33,38,58,70	0
3	FAD	A	999	53/53	0.99	0.14	0.24	15,19,21,22	0
3	FAD	D	999	53/53	0.98	0.13	0.08	15,17,20,21	0
3	FAD	B	999	53/53	0.98	0.13	0.06	15,18,20,24	0
4	SO4	A	1659	5/5	0.96	0.10	-0.20	52,54,60,60	0
3	FAD	C	999	53/53	0.99	0.12	-0.24	12,15,17,18	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.