



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 04:05 am GMT

PDB ID : 5AE2  
Title : Ether Lipid-Generating Enzyme AGPS in complex with inhibitor 1e  
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 : 2.00 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

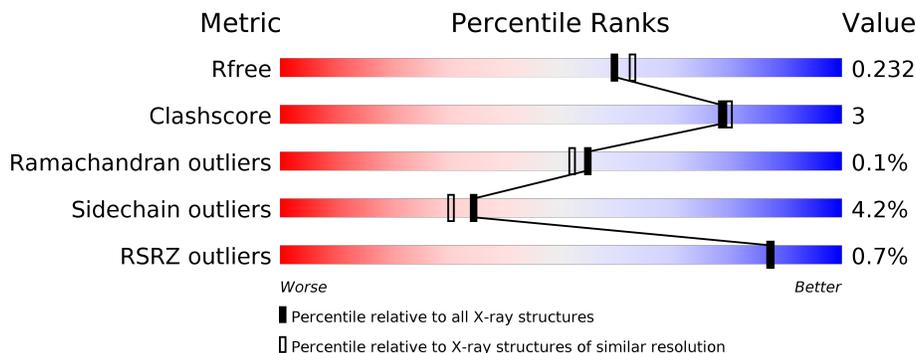
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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	
1	B	658	
1	C	658	
1	D	658	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	FYC	A	888	-	-	-	X
2	FYC	B	888	-	-	-	X
2	FYC	C	888	-	-	-	X
2	FYC	D	888	-	-	-	X
4	SO4	A	1659	-	-	-	X
4	SO4	B	1659	-	-	-	X
4	SO4	C	1659	-	-	-	X
4	SO4	D	1659	-	-	-	X

## 2 Entry composition [i](#)

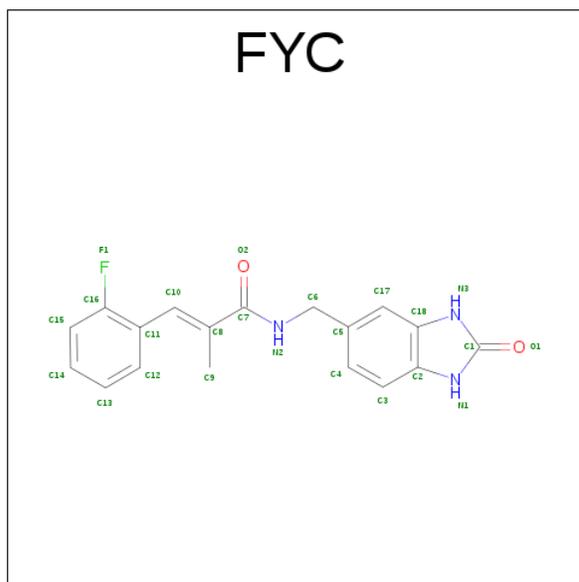
There are 5 unique types of molecules in this entry. The entry contains 18801 atoms, of which 0 are hydrogens and 0 are deuteriums.

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
			Total	C	N	O	S			
1	A	561	Total 4431	C 2813	N 769	O 825	S 24	0	0	0
1	B	542	Total 4292	C 2727	N 745	O 796	S 24	0	0	0
1	C	557	Total 4402	C 2789	N 766	O 823	S 24	0	1	0
1	D	550	Total 4355	C 2761	N 757	O 812	S 25	0	1	0

- Molecule 2 is 3-(2-FLUOROPHENYL)-2-METHYL-N-((2-OXO-2,3-DIHYDRO-1H-BENZO[D]IMIDAZOL-5-YL)METHYL)ACRYLAMIDE (three-letter code: FYC) (formula: C<sub>18</sub>H<sub>16</sub>FN<sub>3</sub>O<sub>2</sub>).



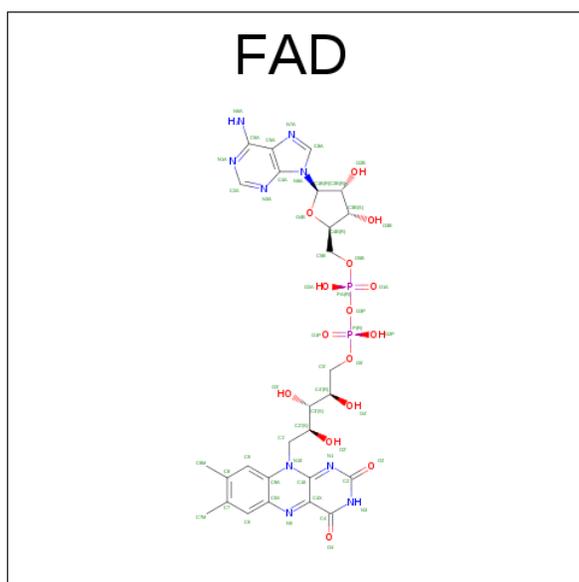
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
2	A	1	Total 24	C 18	F 1	N 3	O 2	0	0

*Continued on next page...*

Continued from previous page...

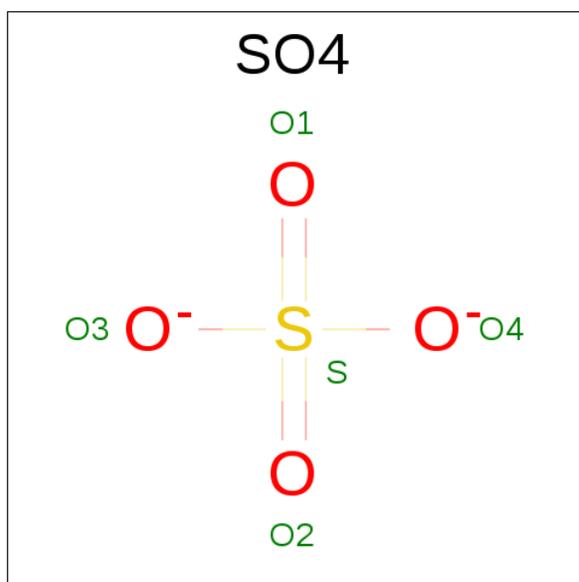
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	F	N	O	0	0
			24	18	1	3	2		
2	C	1	Total	C	F	N	O	0	0
			24	18	1	3	2		
2	D	1	Total	C	F	N	O	0	0
			24	18	1	3	2		

- 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 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0

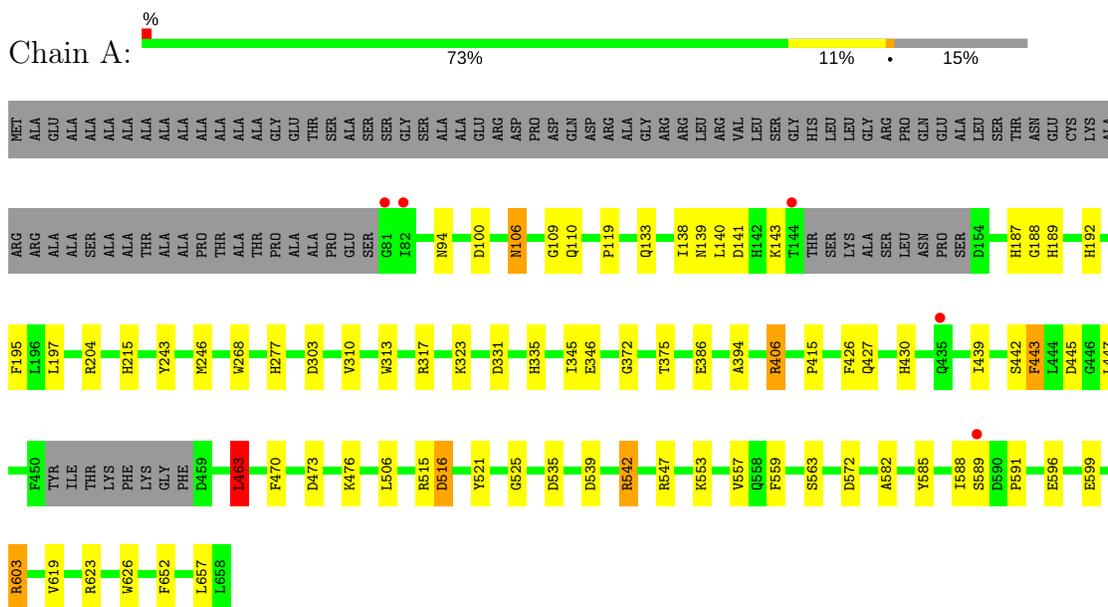
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	282	Total O 282 282	0	0
5	B	223	Total O 223 223	0	0
5	C	268	Total O 268 268	0	0
5	D	220	Total O 220 220	0	0

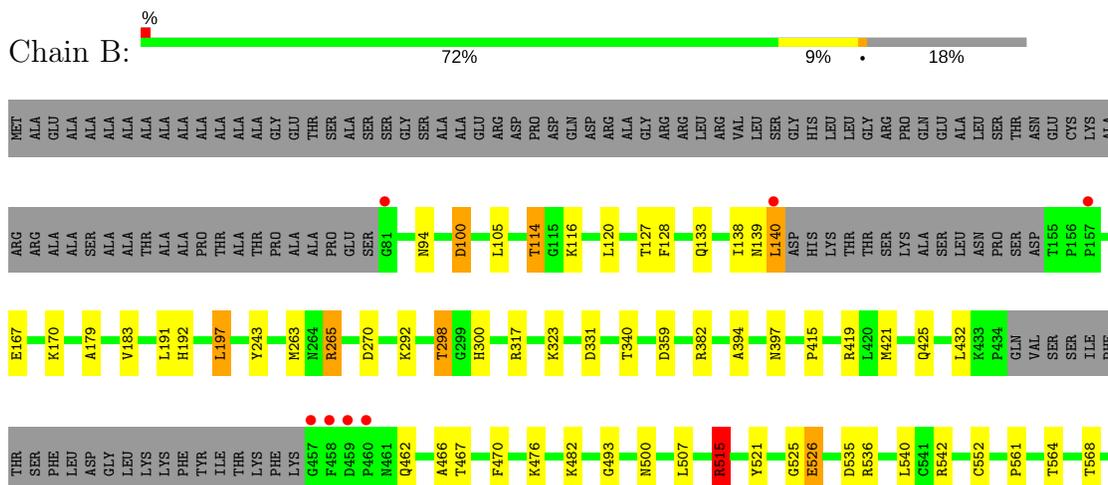
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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: ALKYLDIHYDROXYACETONEPHOSPHATE SYNTHASE, PEROXISOMAL



- Molecule 1: ALKYLDIHYDROXYACETONEPHOSPHATE SYNTHASE, PEROXISOMAL





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.53Å 98.57Å 107.07Å 90.57° 90.14° 95.67°	Depositor
Resolution (Å)	107.06 – 2.00 72.69 – 2.00	Depositor EDS
% Data completeness (in resolution range)	85.3 (107.06-2.00) 85.1 (72.69-2.00)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0124	Depositor
R, $R_{free}$	0.175 , 0.229 0.183 , 0.232	Depositor DCC
$R_{free}$ test set	1516 reflections (1.08%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtrriage
Anisotropy	0.019	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 49.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.017 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	18801	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.68% 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: FYC, SO4, FAD

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	0.84	2/4530 (0.0%)	0.99	17/6123 (0.3%)
1	B	0.85	1/4389 (0.0%)	0.98	17/5932 (0.3%)
1	C	0.84	0/4504	1.00	23/6090 (0.4%)
1	D	0.84	1/4455 (0.0%)	0.98	14/6022 (0.2%)
All	All	0.84	4/17878 (0.0%)	0.99	71/24167 (0.3%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	516	ASP	CB-CG	8.27	1.69	1.51
1	D	641	SER	CB-OG	6.17	1.50	1.42
1	B	603	ARG	CD-NE	-5.96	1.36	1.46
1	A	626	TRP	CE3-CZ3	5.48	1.47	1.38

The worst 5 of 71 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	623	ARG	NE-CZ-NH2	-17.53	111.53	120.30
1	D	623	ARG	NE-CZ-NH2	-16.11	112.25	120.30
1	A	623	ARG	NE-CZ-NH2	-13.31	113.65	120.30
1	C	623	ARG	NE-CZ-NH1	12.33	126.46	120.30
1	A	623	ARG	NE-CZ-NH1	10.92	125.76	120.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-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 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	4431	0	4360	39	0
1	B	4292	0	4234	30	0
1	C	4402	0	4340	28	0
1	D	4355	0	4300	19	0
2	A	24	0	16	4	0
2	B	24	0	16	5	0
2	C	24	0	16	1	0
2	D	24	0	16	2	0
3	A	53	0	31	1	0
3	B	53	0	31	2	0
3	C	53	0	31	1	0
3	D	53	0	31	2	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	282	0	0	3	0
5	B	223	0	0	2	0
5	C	268	0	0	6	0
5	D	220	0	0	1	0
All	All	18801	0	17422	115	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 115 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:516:ASP:OD2	5:C:2012:HOH:O	1.87	0.92
2:B:888:FYC:H62C	3:B:999:FAD:HM72	1.53	0.88
1:C:419:ARG:HD3	5:C:2197:HOH:O	1.80	0.79
2:D:888:FYC:H62C	3:D:999:FAD:HM72	1.66	0.78
1:B:192:HIS:HB3	1:B:243:TYR:OH	1.83	0.77

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%)	38	33
1	B	536/658 (82%)	526 (98%)	10 (2%)	0	100	100
1	C	554/658 (84%)	540 (98%)	13 (2%)	1 (0%)	51	48
1	D	547/658 (83%)	535 (98%)	12 (2%)	0	100	100
All	All	2192/2632 (83%)	2144 (98%)	45 (2%)	3 (0%)	55	52

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	141	ASP
1	C	106	ASN
1	A	443	PHE

### 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	478/545 (88%)	462 (97%)	16 (3%)	43	41
1	B	463/545 (85%)	442 (96%)	21 (4%)	32	27
1	C	477/545 (88%)	459 (96%)	18 (4%)	38	35
1	D	472/545 (87%)	448 (95%)	24 (5%)	28	22
All	All	1890/2180 (87%)	1811 (96%)	79 (4%)	34	30

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	613	LEU
1	C	224	LEU
1	D	563	SER
1	B	646	VAL
1	C	141	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	335	HIS
1	A	362	HIS
1	C	423	ASN
1	A	328	ASN
1	D	285	GLN

### 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	0.60	0	6,6,6	0.53	0
2	FYC	A	888	-	23,26,26	1.20	3 (13%)	27,36,36	2.98	12 (44%)
3	FAD	A	999	-	51,58,58	1.33	5 (9%)	54,89,89	2.54	13 (24%)
4	SO4	B	1659	-	4,4,4	0.69	0	6,6,6	0.86	0
2	FYC	B	888	-	23,26,26	1.27	4 (17%)	27,36,36	3.18	13 (48%)
3	FAD	B	999	-	51,58,58	1.57	10 (19%)	54,89,89	2.81	17 (31%)
4	SO4	C	1659	-	4,4,4	0.49	0	6,6,6	0.60	0
2	FYC	C	888	-	23,26,26	1.05	1 (4%)	27,36,36	2.46	10 (37%)
3	FAD	C	999	-	51,58,58	1.54	9 (17%)	54,89,89	2.32	16 (29%)
4	SO4	D	1659	-	4,4,4	0.49	0	6,6,6	0.48	0
2	FYC	D	888	-	23,26,26	0.98	0	27,36,36	2.49	13 (48%)
3	FAD	D	999	-	51,58,58	1.33	4 (7%)	54,89,89	2.45	13 (24%)

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	FYC	A	888	-	-	0/13/13/13	0/3/3/3
3	FAD	A	999	-	-	0/28/50/50	0/6/6/6
4	SO4	B	1659	-	-	0/0/0/0	0/0/0/0
2	FYC	B	888	-	-	0/13/13/13	0/3/3/3
3	FAD	B	999	-	-	0/28/50/50	0/6/6/6
4	SO4	C	1659	-	-	0/0/0/0	0/0/0/0
2	FYC	C	888	-	-	0/13/13/13	0/3/3/3
3	FAD	C	999	-	-	0/28/50/50	0/6/6/6
4	SO4	D	1659	-	-	0/0/0/0	0/0/0/0
2	FYC	D	888	-	-	0/13/13/13	0/3/3/3
3	FAD	D	999	-	-	0/28/50/50	0/6/6/6

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	999	FAD	C1'-N10	-4.85	1.43	1.48
3	B	999	FAD	C2-N1	-3.64	1.31	1.38
2	B	888	FYC	C11-C10	-2.94	1.42	1.46
3	B	999	FAD	C1'-N10	-2.67	1.45	1.48
3	C	999	FAD	C2-N3	-2.60	1.33	1.38

The worst 5 of 107 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	999	FAD	N3A-C2A-N1A	-6.89	122.86	128.86
2	D	888	FYC	C6-N2-C7	-6.76	112.81	122.08
3	C	999	FAD	C4-C4X-C10	-6.58	114.64	119.96
2	A	888	FYC	C4-C3-C2	-6.46	112.92	120.84
3	D	999	FAD	C4-C4X-C10	-6.38	114.80	119.96

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	FYC	4	0
3	A	999	FAD	1	0
2	B	888	FYC	5	0
3	B	999	FAD	2	0
2	C	888	FYC	1	0
3	C	999	FAD	1	0
2	D	888	FYC	2	0
3	D	999	FAD	2	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	561/658 (85%)	-0.36	5 (0%) 84 83	11, 24, 54, 90	0
1	B	542/658 (82%)	-0.34	7 (1%) 77 77	12, 24, 49, 96	0
1	C	557/658 (84%)	-0.36	2 (0%) 92 92	10, 24, 46, 81	1 (0%)
1	D	550/658 (83%)	-0.35	1 (0%) 94 94	13, 25, 49, 75	0
All	All	2210/2632 (83%)	-0.35	15 (0%) 87 87	10, 24, 49, 96	1 (0%)

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	458	PHE	5.4
1	A	144	THR	4.8
1	A	81	GLY	4.1
1	B	459	ASP	3.9
1	C	435	GLN	3.6

### 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	B	1659	5/5	0.91	0.20	17.89	50,53,57,58	0
2	FYC	D	888	24/24	0.91	0.19	5.53	32,43,57,58	0
4	SO4	D	1659	5/5	0.96	0.18	5.26	38,46,48,48	0
2	FYC	C	888	24/24	0.91	0.16	4.13	27,36,52,57	0
2	FYC	A	888	24/24	0.94	0.15	4.04	25,37,63,66	0
4	SO4	C	1659	5/5	0.97	0.15	3.61	40,43,49,56	0
2	FYC	B	888	24/24	0.91	0.17	3.54	32,47,57,62	0
4	SO4	A	1659	5/5	0.97	0.12	2.61	36,39,41,44	0
3	FAD	C	999	53/53	0.99	0.10	0.32	9,10,13,13	0
3	FAD	D	999	53/53	0.99	0.09	-0.07	10,12,14,15	0
3	FAD	B	999	53/53	0.99	0.09	-0.08	11,13,16,18	0
3	FAD	A	999	53/53	0.99	0.09	-0.38	11,13,17,20	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.