wwPDB X-ray Structure Validation Summary Report (i)

Aug 27, 2023 – 04:13 AM EDT

PDB ID	:	3GNC
Title	:	Crystal structure of Glutaryl-COA dehydrogenase from Burkholderia Pseudomallei with fragment 6421
Authors	:	Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on	:	2009-03-16
Resolution	:	2.15 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

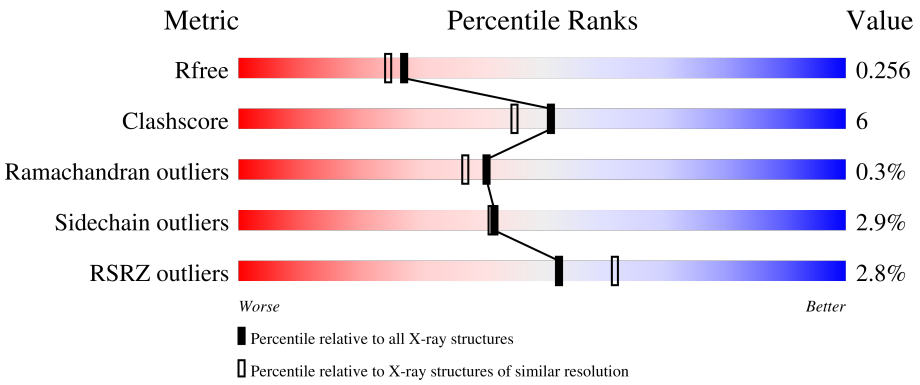
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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}	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=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	399	3% 85% 10% • 5%
1	B	399	3% 79% 14% • 7%
1	C	399	3% 85% 10% • •
1	D	399	3% 83% 11% • 5%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 11870 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.

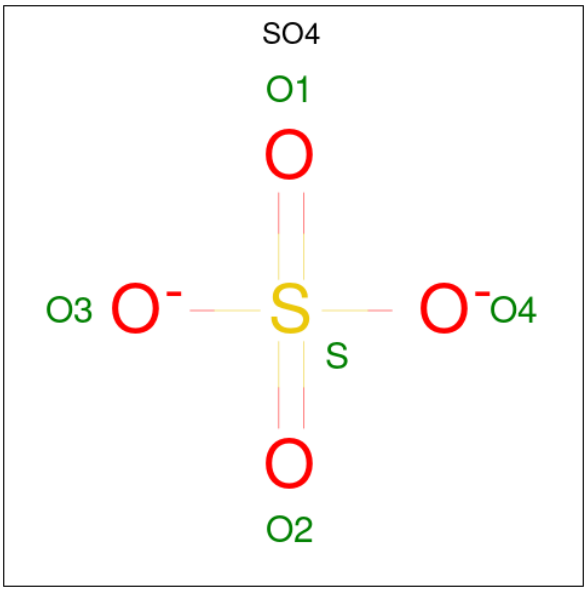
- Molecule 1 is a protein called Glutaryl-CoA dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	380	2896	1831	511	538	16	0	0	0
1	B	372	2845	1799	502	528	16	0	0	0
1	C	383	2904	1834	515	539	16	0	0	0
1	D	380	2901	1835	514	536	16	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

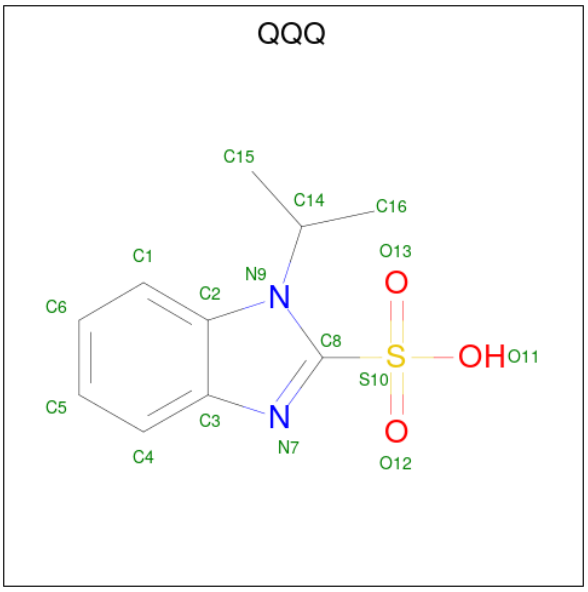
Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP Q3JP94
A	-2	PRO	-	expression tag	UNP Q3JP94
A	-1	GLY	-	expression tag	UNP Q3JP94
A	0	SER	-	expression tag	UNP Q3JP94
B	-3	GLY	-	expression tag	UNP Q3JP94
B	-2	PRO	-	expression tag	UNP Q3JP94
B	-1	GLY	-	expression tag	UNP Q3JP94
B	0	SER	-	expression tag	UNP Q3JP94
C	-3	GLY	-	expression tag	UNP Q3JP94
C	-2	PRO	-	expression tag	UNP Q3JP94
C	-1	GLY	-	expression tag	UNP Q3JP94
C	0	SER	-	expression tag	UNP Q3JP94
D	-3	GLY	-	expression tag	UNP Q3JP94
D	-2	PRO	-	expression tag	UNP Q3JP94
D	-1	GLY	-	expression tag	UNP Q3JP94
D	0	SER	-	expression tag	UNP Q3JP94

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is 1-(1-methylethyl)-1H-benzimidazole-2-sulfonic acid (three-letter code: QQQ) (formula: $C_{10}H_{12}N_2O_3S$).



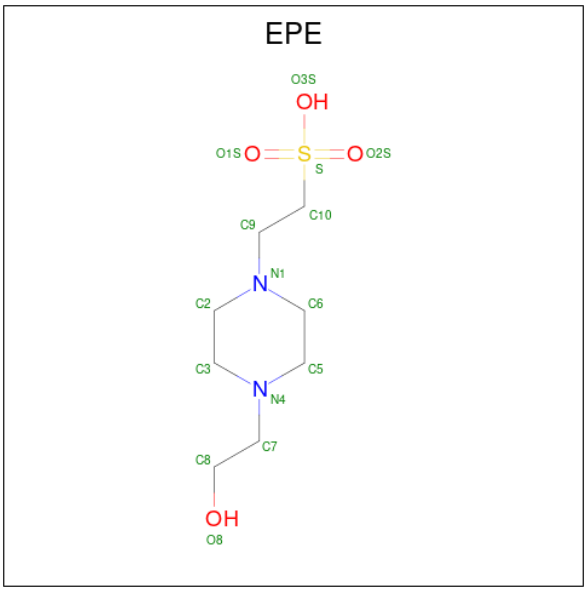
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	A	1	16	10	2	3	1	0	0
			Total	C	N	O	S		
3	B	1	16	10	2	3	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
3	C	1	16	10	2	3	1	0	0
			Total	C	N	O	S		
3	D	1	16	10	2	3	1	0	0

- Molecule 4 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
4	B	1	15	8	2	4	1	0	0
			Total	C	N	O	S		
4	C	1	15	8	2	4	1	0	0
			Total	C	N	O	S		
4	D	1	15	8	2	4	1	0	0

- Molecule 5 is water.

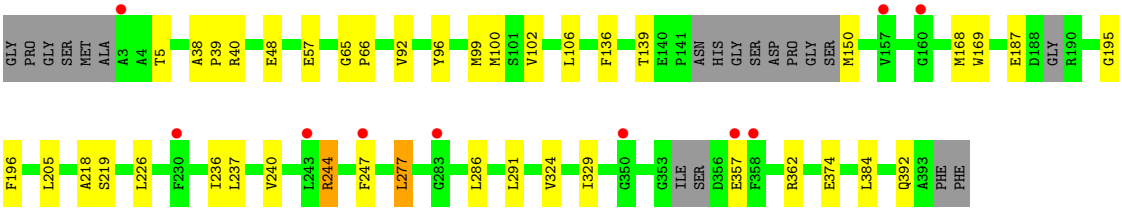
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	50	50	50	0	0
			Total	O		
5	B	51	51	51	0	0
			Total	O		
5	C	50	50	50	0	0
			Total	O		
5	D	59	59	59	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

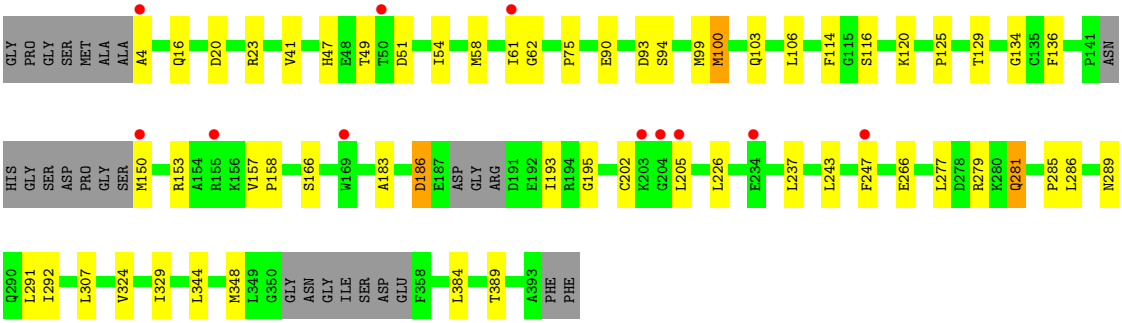
• Molecule 1: Glutaryl-CoA dehydrogenase

Chain A:



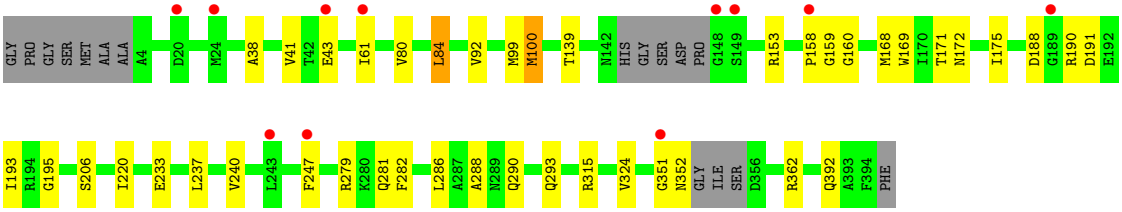
• Molecule 1: Glutaryl-CoA dehydrogenase

Chain B:



• Molecule 1: Glutaryl-CoA dehydrogenase

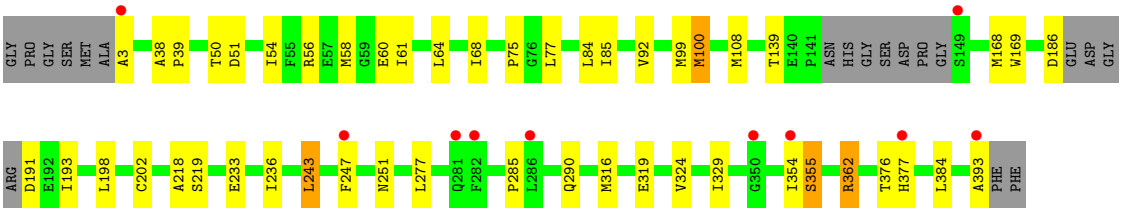
Chain C:



• Molecule 1: Glutaryl-CoA dehydrogenase

Chain D:

3%
83%
11%
5%


GLY PRO GLY SER MET ALA A3 A38 P39 T50 D51 I54 F55 R56 E57 M58 G59 E60 I61 L64 I68 P75 G76 L77 L84 I85 V92 M99 M100 M108 T139 E140 P141 ASN HIS GLY SER ASP PRO GLY S149 M168 W169 D186 GLU ASP GLY
ARG D191 E192 I193 L198 C202 A218 S219 E233 I236 L243 F247 N251 L277 Q281 F282 P285 L286 Q290 M316 E319 V324 I329 G350 I354 S355 R362 T376 H377 L384 A393 PHE PHE

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	98.24Å 106.10Å 144.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.68 – 2.15 44.41 – 2.15	Depositor EDS
% Data completeness (in resolution range)	(Not available) (46.68-2.15) 99.4 (44.41-2.15)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.81 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, R_{free}	0.206 , 0.255 0.211 , 0.256	Depositor DCC
R_{free} test set	4105 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	39.0	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 40.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11870	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.80% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: EPE, SO4, QQQ

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	0.46	0/2950	0.62	0/3987
1	B	0.48	0/2897	0.60	1/3916 (0.0%)
1	C	0.47	0/2958	0.61	0/3999
1	D	0.50	0/2956	0.66	2/3996 (0.1%)
All	All	0.48	0/11761	0.62	3/15898 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	362	ARG	NE-CZ-NH1	-6.63	116.99	120.30
1	D	362	ARG	NE-CZ-NH2	5.70	123.15	120.30
1	B	344	LEU	CA-CB-CG	5.39	127.70	115.30

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	2896	0	2867	22	0
1	B	2845	0	2840	45	0
1	C	2904	0	2878	31	0
1	D	2901	0	2890	43	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	5	0	0	0	0
3	A	16	0	11	0	0
3	B	16	0	11	0	0
3	C	16	0	11	0	0
3	D	16	0	11	1	0
4	B	15	0	17	0	0
4	C	15	0	17	0	0
4	D	15	0	17	0	0
5	A	50	0	0	1	0
5	B	51	0	0	6	0
5	C	50	0	0	7	0
5	D	59	0	0	3	0
All	All	11870	0	11570	129	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 6.

The worst 5 of 129 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:D:247:PHE:HZ	1:D:324:VAL:HG11	1.12	1.10
1:D:247:PHE:CZ	1:D:324:VAL:HG11	1.99	0.98
1:C:100:MET:HE2	5:C:415:HOH:O	1.66	0.92
1:D:50:THR:HG22	1:D:218:ALA:O	1.69	0.91
1:B:100:MET:HE3	5:B:433:HOH:O	1.71	0.91

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	372/399 (93%)	359 (96%)	12 (3%)	1 (0%)	41	37
1	B	364/399 (91%)	353 (97%)	11 (3%)	0	100	100
1	C	377/399 (94%)	360 (96%)	15 (4%)	2 (0%)	29	22
1	D	374/399 (94%)	365 (98%)	8 (2%)	1 (0%)	41	37
All	All	1487/1596 (93%)	1437 (97%)	46 (3%)	4 (0%)	41	37

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	282	PHE
1	D	355	SER
1	A	187	GLU
1	C	188	ASP

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	294/314 (94%)	284 (97%)	10 (3%)	37	35
1	B	292/314 (93%)	284 (97%)	8 (3%)	44	46
1	C	294/314 (94%)	284 (97%)	10 (3%)	37	35
1	D	296/314 (94%)	290 (98%)	6 (2%)	55	59
All	All	1176/1256 (94%)	1142 (97%)	34 (3%)	42	42

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	84	LEU
1	D	100	MET
1	D	243	LEU
1	B	100	MET
1	B	61	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such

sidechains are listed below:

Mol	Chain	Res	Type
1	C	392	GLN
1	D	377	HIS
1	D	392	GLN
1	B	235	ASN
1	A	392	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

8 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	EPE	B	397	-	15,15,15	0.77	1 (6%)	18,20,20	1.07	1 (5%)
3	QQQ	C	396	-	14,17,17	1.76	2 (14%)	17,26,26	1.02	0
4	EPE	C	397	-	15,15,15	0.82	1 (6%)	18,20,20	1.75	5 (27%)
3	QQQ	B	396	-	14,17,17	1.80	2 (14%)	17,26,26	1.32	2 (11%)
3	QQQ	A	397	-	14,17,17	1.76	2 (14%)	17,26,26	1.08	2 (11%)
3	QQQ	D	396	-	14,17,17	1.76	2 (14%)	17,26,26	1.22	2 (11%)
2	SO4	A	396	-	4,4,4	0.11	0	6,6,6	0.16	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	EPE	D	397	-	15,15,15	1.03	1 (6%)	18,20,20	1.63	2 (11%)

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	EPE	B	397	-	-	1/9/19/19	0/1/1/1
3	QQQ	C	396	-	-	0/4/10/10	0/2/2/2
4	EPE	C	397	-	-	5/9/19/19	0/1/1/1
3	QQQ	B	396	-	-	0/4/10/10	0/2/2/2
3	QQQ	A	397	-	-	0/4/10/10	0/2/2/2
3	QQQ	D	396	-	-	0/4/10/10	0/2/2/2
4	EPE	D	397	-	-	4/9/19/19	0/1/1/1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	396	QQQ	C2-C3	4.47	1.49	1.40
3	D	396	QQQ	C2-C3	4.44	1.49	1.40
3	A	397	QQQ	C2-C3	4.44	1.49	1.40
3	C	396	QQQ	C2-C3	4.17	1.49	1.40
3	C	396	QQQ	O12-S10	3.82	1.62	1.43

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	397	EPE	O2S-S-C10	5.28	113.28	106.92
4	C	397	EPE	O3S-S-C10	4.26	112.66	105.77
3	B	396	QQQ	C16-C14-N9	-3.75	107.14	111.51
4	B	397	EPE	O1S-S-C10	3.23	110.81	106.92
3	D	396	QQQ	C16-C14-N9	-2.90	108.14	111.51

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	397	EPE	C8-C7-N4-C5
4	C	397	EPE	S-C10-C9-N1
4	D	397	EPE	C9-C10-S-O1S

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	D	397	EPE	C9-C10-S-O2S
4	D	397	EPE	C8-C7-N4-C5

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	396	QQQ	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	380/399 (95%)	0.12	10 (2%)	56	64	31, 45, 65, 70	0
1	B	372/399 (93%)	0.13	11 (2%)	50	59	28, 45, 65, 73	0
1	C	383/399 (95%)	0.13	11 (2%)	51	61	27, 48, 68, 77	0
1	D	380/399 (95%)	-0.04	10 (2%)	56	64	25, 40, 56, 69	0
All	All	1515/1596 (94%)	0.08	42 (2%)	53	62	25, 44, 65, 77	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	354	ILE	6.1
1	D	3	ALA	5.3
1	D	149	SER	5.2
1	B	50	THR	4.1
1	A	230	PHE	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 monosaccharides 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. 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	B-factors(Å^2)	Q<0.9
4	EPE	D	397	15/15	0.90	0.18	60,69,71,72	0
4	EPE	B	397	15/15	0.95	0.12	53,58,61,62	0
2	SO4	A	396	5/5	0.95	0.10	61,62,63,63	0
4	EPE	C	397	15/15	0.96	0.15	38,49,57,58	0
3	QQQ	A	397	16/16	0.98	0.10	42,44,45,45	0
3	QQQ	B	396	16/16	0.98	0.09	41,43,44,45	0
3	QQQ	D	396	16/16	0.98	0.09	32,36,38,39	0
3	QQQ	C	396	16/16	0.99	0.10	34,36,41,42	0

6.5 Other polymers (i)

There are no such residues in this entry.