



# wwPDB X-ray Structure Validation Summary Report

Feb 1, 2016 – 11:06 AM GMT

PDB ID : 3O3M  
Title : (R)-2-Hydroxyisocaproyl-CoA Dehydratase  
Authors : Knauer, S.H.; Buckel, W.; Dobbek, H.  
Deposited on : 2010-07-25  
Resolution : 1.82 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 (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

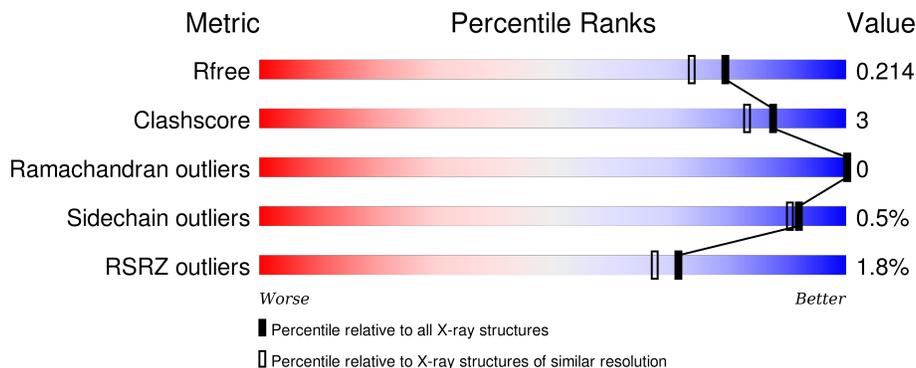
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.82 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	5422 (1.84-1.80)
Clashscore	102246	6347 (1.84-1.80)
Ramachandran outliers	100387	6276 (1.84-1.80)
Sidechain outliers	100360	6276 (1.84-1.80)
RSRZ outliers	91569	5439 (1.84-1.80)

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	408	 91% 7% •
1	C	408	 90% 8% •
2	B	385	 92% 5% •
2	D	385	 92% 5% •

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	A	410	-	-	-	X
4	SO4	C	410	-	-	-	X
5	H2S	B	387	-	-	-	X
5	H2S	D	387	-	-	-	X

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 26633 atoms, of which 12460 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 alpha subunit 2-hydroxyisocaproyl-CoA dehydratase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	399	6410	2050	3188	526	616	30	0	11	0
1	C	401	6448	2060	3209	529	619	31	0	11	0

- Molecule 2 is a protein called beta subunit 2-hydroxyacyl-CoA dehydratase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	373	5975	1886	3004	492	570	23	0	4	0
2	D	375	6063	1910	3059	496	575	23	11	8	0

There are 20 discrepancies between the modelled and reference sequences:

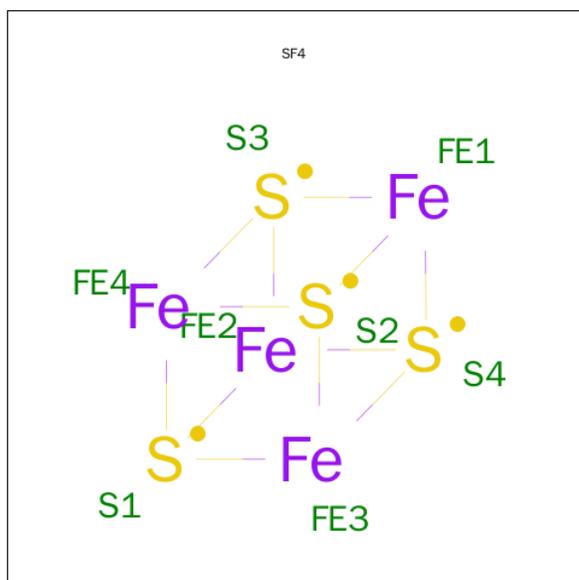
Chain	Residue	Modelled	Actual	Comment	Reference
B	376	SER	-	EXPRESSION TAG	UNP Q5U923
B	377	ALA	-	EXPRESSION TAG	UNP Q5U923
B	378	TRP	-	EXPRESSION TAG	UNP Q5U923
B	379	SER	-	EXPRESSION TAG	UNP Q5U923
B	380	HIS	-	EXPRESSION TAG	UNP Q5U923
B	381	PRO	-	EXPRESSION TAG	UNP Q5U923
B	382	GLN	-	EXPRESSION TAG	UNP Q5U923
B	383	PHE	-	EXPRESSION TAG	UNP Q5U923
B	384	GLU	-	EXPRESSION TAG	UNP Q5U923
B	385	LYS	-	EXPRESSION TAG	UNP Q5U923
D	376	SER	-	EXPRESSION TAG	UNP Q5U923
D	377	ALA	-	EXPRESSION TAG	UNP Q5U923
D	378	TRP	-	EXPRESSION TAG	UNP Q5U923
D	379	SER	-	EXPRESSION TAG	UNP Q5U923
D	380	HIS	-	EXPRESSION TAG	UNP Q5U923
D	381	PRO	-	EXPRESSION TAG	UNP Q5U923

*Continued on next page...*

Continued from previous page...

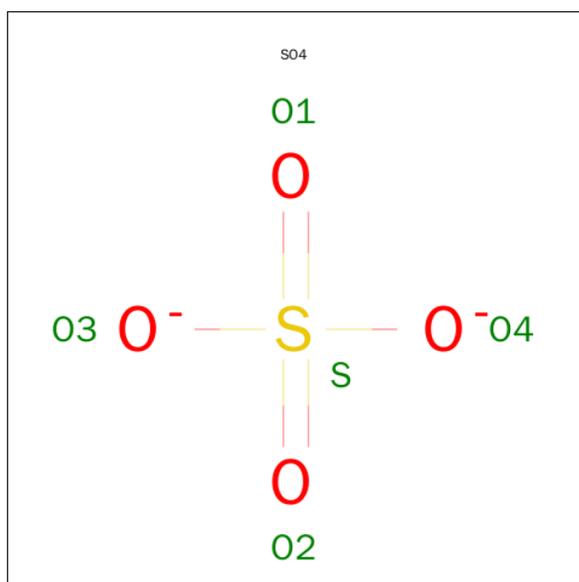
Chain	Residue	Modelled	Actual	Comment	Reference
D	382	GLN	-	EXPRESSION TAG	UNP Q5U923
D	383	PHE	-	EXPRESSION TAG	UNP Q5U923
D	384	GLU	-	EXPRESSION TAG	UNP Q5U923
D	385	LYS	-	EXPRESSION TAG	UNP Q5U923

- Molecule 3 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).



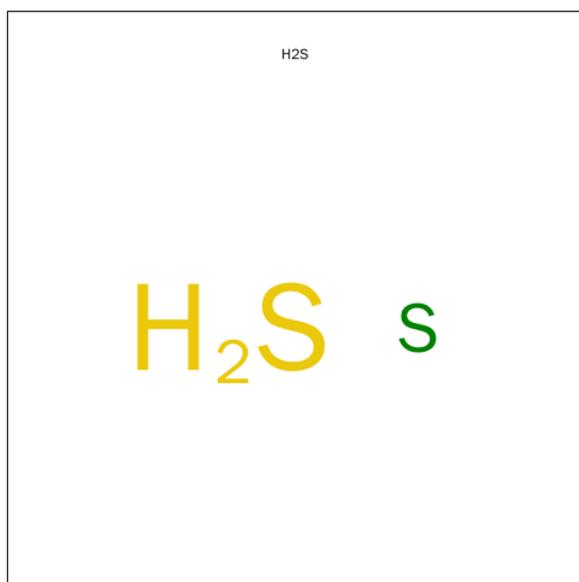
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	Fe	S	0	0
			8	4	4		
3	B	1	Total	Fe	S	0	0
			8	4	4		
3	C	1	Total	Fe	S	0	0
			8	4	4		
3	D	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 5 is HYDROSULFURIC ACID (three-letter code: H2S) (formula: H<sub>2</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total S 1 1	0	0
5	D	1	Total S 1 1	0	0

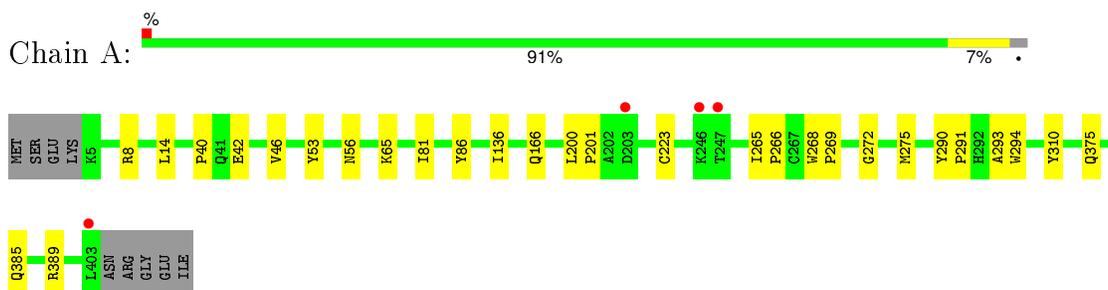
- Molecule 6 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	A	436	Total 442	O 442	0	6
6	B	349	Total 349	O 349	0	0
6	C	482	Total 484	O 484	0	2
6	D	416	Total 418	O 418	1	2

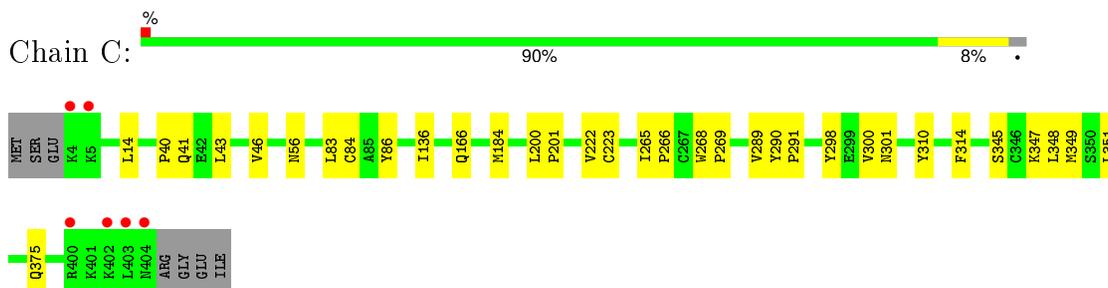
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

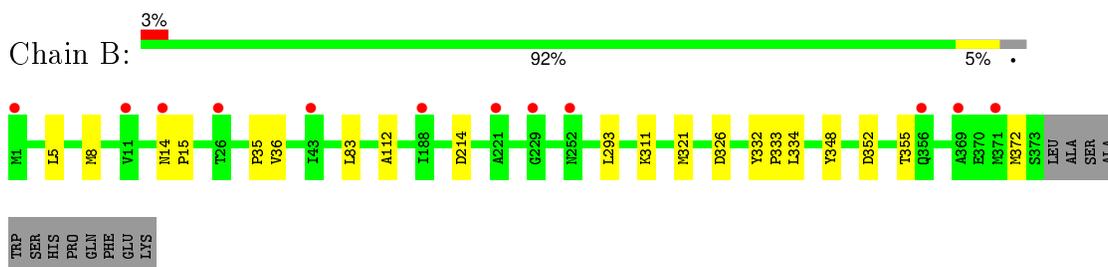
- Molecule 1: alpha subunit 2-hydroxyisocaproyl-CoA dehydratase



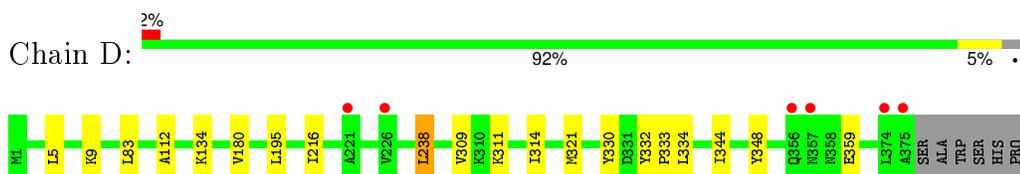
- Molecule 1: alpha subunit 2-hydroxyisocaproyl-CoA dehydratase



- Molecule 2: beta subunit 2-hydroxyacyl-CoA dehydratase



- Molecule 2: beta subunit 2-hydroxyacyl-CoA dehydratase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.76 Å 129.19 Å 179.17 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.70 – 1.82 19.70 – 1.82	Depositor EDS
% Data completeness (in resolution range)	99.2 (19.70-1.82) 99.2 (19.70-1.82)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.71 (at 1.82 Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.6.1_357)	Depositor
R, $R_{free}$	0.173 , 0.220 0.165 , 0.214	Depositor DCC
$R_{free}$ test set	7306 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	21.8	Xtrriage
Anisotropy	0.610	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 70.3	EDS
Estimated twinning fraction	No twinning to report.	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Outliers	1 of 146207 reflections (0.001%)	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	26633	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.64 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.5787e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: SF4, SO4, H2S

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	0.24	0/3319	0.39	0/4476
1	C	0.25	0/3335	0.40	0/4495
2	B	0.23	0/3038	0.39	0/4094
2	D	0.24	0/3080	0.39	0/4149
All	All	0.24	0/12772	0.39	0/17214

There are no bond length outliers.

There are no bond angle outliers.

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	3222	3188	3176	17	0
1	C	3239	3209	3207	23	0
2	B	2971	3004	2990	13	0
2	D	3004	3059	3052	15	0
3	A	8	0	0	0	0
3	B	8	0	0	0	0
3	C	8	0	0	0	0
3	D	8	0	0	0	0
4	A	5	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	5	0	0	0	0
5	B	1	0	0	1	0
5	D	1	0	0	0	0
6	A	442	0	0	2	0
6	B	349	0	0	1	0
6	C	484	0	0	0	0
6	D	418	0	0	2	0
All	All	14173	12460	12425	65	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 65 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:84:CYS:HA	1:C:345[B]:SER:HB2	1.83	0.61
2:D:83:LEU:HD21	2:D:112:ALA:HB1	1.83	0.59
2:D:134[B]:LYS:HG2	6:D:1034:HOH:O	2.02	0.59
2:D:321:MET:HG3	2:D:348:TYR:CE1	2.40	0.56
2:B:35:PRO:HB2	5:B:387:H2S:S	2.44	0.56

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	408/408 (100%)	400 (98%)	8 (2%)	0	100	100
1	C	410/408 (100%)	401 (98%)	9 (2%)	0	100	100
2	B	375/385 (97%)	365 (97%)	10 (3%)	0	100	100
2	D	381/385 (99%)	372 (98%)	9 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1574/1586 (99%)	1538 (98%)	36 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	351/348 (101%)	348 (99%)	3 (1%)	84	79
1	C	353/348 (101%)	352 (100%)	1 (0%)	94	93
2	B	330/336 (98%)	327 (99%)	3 (1%)	84	79
2	D	335/336 (100%)	334 (100%)	1 (0%)	94	93
All	All	1369/1368 (100%)	1361 (99%)	8 (1%)	92	88

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	214	ASP
2	D	238	LEU
2	B	372	MET
1	A	65	LYS
2	B	326	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	41	GLN
1	C	357	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](#)

Of 8 ligands modelled in this entry, 2 are modelled with single atom - leaving 6 for Mogul analysis.

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
3	SF4	A	409	1,6	0,12,12	0.00	-	0,24,24	0.00	-
4	SO4	A	410	-	4,4,4	0.23	0	6,6,6	0.08	0
3	SF4	B	386	2,5	0,12,12	0.00	-	0,24,24	0.00	-
3	SF4	C	409	1,6	0,12,12	0.00	-	0,24,24	0.00	-
4	SO4	C	410	-	4,4,4	0.22	0	6,6,6	0.07	0
3	SF4	D	386	2,5	0,12,12	0.00	-	0,24,24	0.00	-

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
3	SF4	A	409	1,6	-	0/0/48/48	0/6/5/5
4	SO4	A	410	-	-	0/0/0/0	0/0/0/0
3	SF4	B	386	2,5	-	0/0/48/48	0/6/5/5
3	SF4	C	409	1,6	-	0/0/48/48	0/6/5/5
4	SO4	C	410	-	-	0/0/0/0	0/0/0/0
3	SF4	D	386	2,5	-	0/0/48/48	0/6/5/5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	399/408 (97%)	-0.27	4 (1%) 84 82	14, 28, 51, 92	0
1	C	401/408 (98%)	-0.34	6 (1%) 76 72	13, 24, 42, 125	0
2	B	373/385 (96%)	-0.00	12 (3%) 51 45	15, 34, 71, 175	0
2	D	375/385 (97%)	-0.27	6 (1%) 74 71	14, 26, 52, 156	0
All	All	1548/1586 (97%)	-0.22	28 (1%) 71 67	13, 27, 58, 175	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	356	GLN	8.3
1	C	403	LEU	5.0
2	B	356	GLN	4.6
2	D	357	ASN	4.6
1	C	4	LYS	4.5

### 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(Å <sup>2</sup> )	Q<0.9
4	SO4	C	410	5/5	0.85	0.20	4.30	53,54,55,58	5
4	SO4	A	410	5/5	0.89	0.17	3.39	56,56,61,61	5
5	H2S	D	387	1/1	0.94	0.12	2.99	22,22,22,22	0
5	H2S	B	387	1/1	0.81	0.13	2.04	26,26,26,26	0
3	SF4	A	409	8/8	0.97	0.06	-0.82	19,23,31,32	1
3	SF4	D	386	8/8	0.99	0.06	-1.07	14,15,16,18	0
3	SF4	B	386	8/8	0.99	0.05	-1.54	16,17,19,21	0
3	SF4	C	409	8/8	0.98	0.05	-1.83	18,25,30,30	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.