



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:07 PM GMT

PDB ID : 4KMO  
Title : Crystal Structure of the Vps33-Vps16 HOPS subcomplex from Chaetomium thermophilum  
Authors : Baker, R.W.; Jeffrey, P.D.; Hughson, F.M.  
Deposited on : 2013-05-08  
Resolution : 2.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

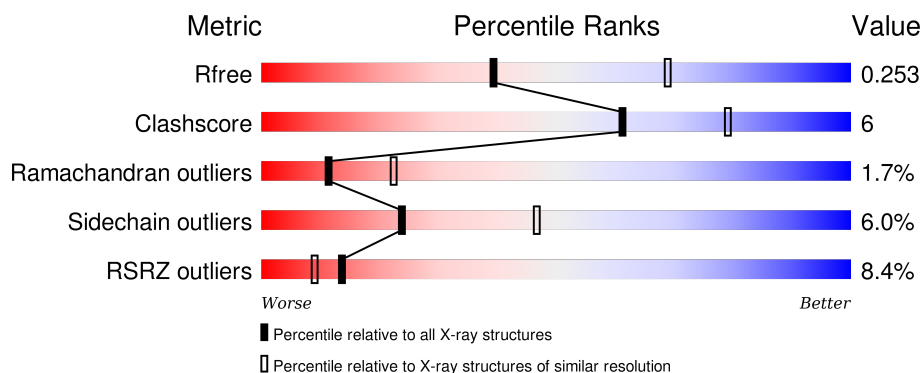
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

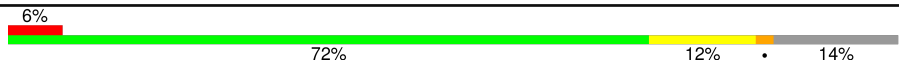
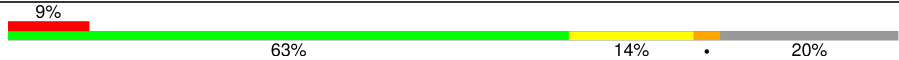
The reported resolution of this entry is 2.60 Å.

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	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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	669	
2	B	333	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6810 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 Small conjugating protein ligase-like protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	574	Total	C	N	O	S	Se	0	0	0
			4536	2870	803	852	4	7			

- Molecule 2 is a protein called Putative vacuolar protein sorting-associated protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	265	Total	C	N	O	S	Se	0	0	0
			2122	1347	374	392	2	7			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	O	S	0	0
			5	4	1		

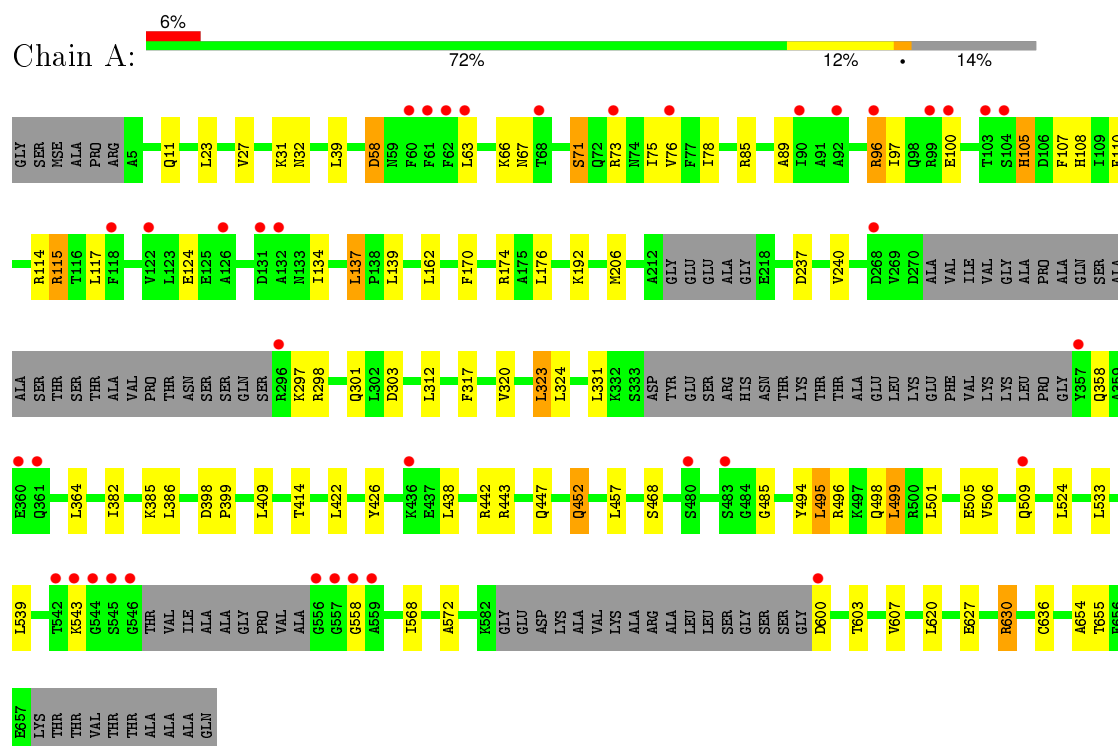
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	118	Total 118	O 118	0	0
4	B	29	Total 29	O 29	0	0

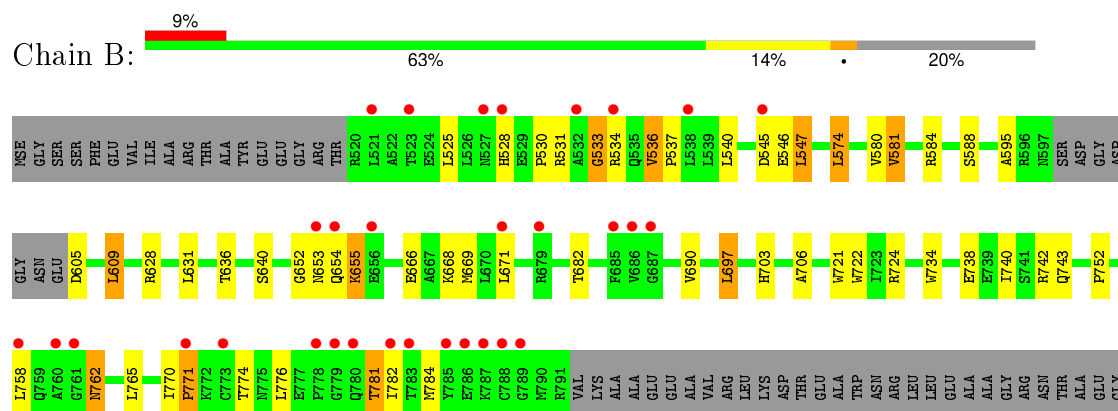
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Small conjugating protein ligase-like protein



- Molecule 2: Putative vacuolar protein sorting-associated protein



ARG  
GLU  
ILE  
GLU  
ARG  
LEU  
GLY  
ALA  
THR  
VAL  
PHE  
LYS  
LYS

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.27Å 100.27Å 176.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.66 – 2.60 48.66 – 2.60	Depositor EDS
% Data completeness (in resolution range)	98.1 (48.66-2.60) 98.1 (48.66-2.60)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.65 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.225 , 0.254 0.222 , 0.253	Depositor DCC
$R_{free}$ test set	1599 reflections (5.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	56.8	Xtriage
Anisotropy	0.197	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 45.3	EDS
Estimated twinning fraction	0.027 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 31609 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6810	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.21	0/4597	0.38	0/6185
2	B	0.21	0/2153	0.39	0/2894
All	All	0.21	0/6750	0.38	0/9079

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

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	4536	0	4597	44	0
2	B	2122	0	2138	33	0
3	B	5	0	0	0	0
4	A	118	0	0	1	0
4	B	29	0	0	1	0
All	All	6810	0	6735	74	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.

All (74) 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:452:GLN:NE2	1:A:654:ALA:O	2.28	0.67
1:A:312:LEU:HD23	1:A:323:LEU:HD23	1.81	0.62
1:A:539:LEU:O	1:A:543:LYS:N	2.30	0.59
2:B:762:ASN:N	2:B:762:ASN:OD1	2.35	0.58
1:A:600:ASP:OD1	1:A:600:ASP:N	2.36	0.58
1:A:358:GLN:NE2	4:A:814:HOH:O	2.37	0.57
2:B:636:THR:HB	2:B:671:LEU:HD13	1.86	0.57
2:B:697:LEU:HB3	2:B:706:ALA:HB2	1.87	0.56
1:A:607:VAL:HG22	1:A:636:CYS:HB2	1.87	0.56
2:B:776:LEU:HD11	2:B:784:MSE:HE1	1.87	0.55
1:A:447:GLN:HB2	2:B:690:VAL:HB	1.89	0.55
2:B:536:VAL:HG13	2:B:537:PRO:HD3	1.90	0.54
2:B:581:VAL:HG13	2:B:588:SER:HA	1.89	0.52
1:A:73:ARG:HG3	1:A:105:HIS:HA	1.92	0.51
1:A:499:LEU:HD13	1:A:568:ILE:HG21	1.91	0.51
2:B:721:TRP:HZ3	2:B:740:ILE:HD13	1.76	0.50
1:A:627:GLU:OE2	1:A:630:ARG:NH1	2.44	0.50
1:A:174:ARG:HA	1:A:206:MSE:HE1	1.93	0.50
1:A:301:GLN:NE2	1:A:303:ASP:OD2	2.34	0.50
1:A:78:ILE:HG12	1:A:110:PHE:HB2	1.94	0.49
1:A:496:ARG:HA	1:A:501:LEU:HB2	1.94	0.49
2:B:546:GLU:HG2	2:B:580:VAL:HG13	1.95	0.48
1:A:137:LEU:HD22	1:A:139:LEU:HG	1.94	0.48
1:A:23:LEU:HD12	1:A:137:LEU:HG	1.95	0.48
2:B:652:GLY:O	2:B:654:GLN:N	2.46	0.47
1:A:443:ARG:O	1:A:447:GLN:HG2	2.15	0.46
2:B:722:TRP:CH2	2:B:752:PRO:HB3	2.50	0.46
2:B:740:ILE:O	2:B:743:GLN:HG2	2.15	0.46
1:A:66:LYS:NZ	1:A:89:ALA:O	2.48	0.46
1:A:494:TYR:HD2	1:A:495:LEU:HD13	1.80	0.46
1:A:115:ARG:HG3	1:A:134:ILE:HG21	1.96	0.46
1:A:442:ARG:HG2	1:A:457:LEU:HD13	1.99	0.45
1:A:96:ARG:HG3	1:A:100:GLU:OE1	2.16	0.45
2:B:574:LEU:HD13	2:B:574:LEU:HA	1.81	0.45
1:A:170:PHE:CE2	1:A:174:ARG:HD2	2.52	0.45
2:B:738:GLU:HG2	2:B:742:ARG:HH21	1.81	0.45
2:B:640:SER:HB2	2:B:668:LYS:HD2	1.97	0.45
2:B:628:ARG:NH1	4:B:1018:HOH:O	2.50	0.45
2:B:545:ASP:O	2:B:547:LEU:N	2.48	0.45
2:B:724:ARG:HE	2:B:740:ILE:HD11	1.82	0.45
2:B:770:ILE:HD11	2:B:781:THR:HG23	1.98	0.45
1:A:237:ASP:HB3	1:A:240:VAL:HG23	1.99	0.45

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:58:ASP:N	1:A:58:ASP:OD1	2.50	0.45
1:A:385:LYS:HB3	1:A:409:LEU:HD21	2.00	0.44
1:A:443:ARG:NH2	2:B:666:GLU:OE1	2.51	0.44
1:A:438:LEU:HD22	1:A:468:SER:HB2	2.00	0.44
1:A:67:ASN:ND2	1:A:67:ASN:O	2.51	0.44
2:B:776:LEU:HA	2:B:776:LEU:HD12	1.71	0.44
1:A:494:TYR:CE2	1:A:498:GLN:HG3	2.53	0.43
1:A:452:GLN:HB2	1:A:452:GLN:HE21	1.62	0.43
2:B:697:LEU:HD12	2:B:697:LEU:HA	1.86	0.43
1:A:382:ILE:HD13	1:A:414:THR:HG22	2.00	0.43
1:A:75:ILE:HB	1:A:107:PHE:HD1	1.83	0.43
1:A:654:ALA:HB3	2:B:631:LEU:HD13	1.99	0.42
2:B:595:ALA:O	2:B:605:ASP:HB2	2.19	0.42
1:A:317:PHE:HA	1:A:320:VAL:HG23	2.01	0.42
2:B:609:LEU:HA	2:B:609:LEU:HD12	1.84	0.42
2:B:770:ILE:HD13	2:B:784:MSE:SE	2.70	0.42
2:B:540:LEU:HA	2:B:540:LEU:HD12	1.90	0.42
1:A:655:THR:O	1:A:655:THR:OG1	2.33	0.42
1:A:409:LEU:O	1:A:414:THR:HG23	2.19	0.42
2:B:655:LYS:HD3	2:B:655:LYS:H	1.85	0.41
1:A:32:ASN:HD21	1:A:71:SER:HB3	1.85	0.41
1:A:398:ASP:HA	1:A:399:PRO:HD3	1.91	0.41
1:A:76:VAL:HG22	1:A:108:HIS:HB2	2.03	0.41
1:A:192:LYS:HE2	1:A:572:ALA:HB3	2.02	0.41
2:B:734:TRP:HB3	2:B:765:LEU:HD22	2.03	0.41
1:A:386:LEU:HD23	1:A:422:LEU:HD11	2.03	0.41
1:A:331:LEU:HD12	1:A:364:LEU:HD22	2.03	0.41
2:B:770:ILE:HD12	2:B:771:PRO:HA	2.03	0.40
2:B:531:ARG:O	2:B:533:GLY:N	2.49	0.40
2:B:782:ILE:HG13	2:B:782:ILE:H	1.56	0.40
2:B:534:ARG:HA	2:B:534:ARG:HD3	1.67	0.40
1:A:499:LEU:HA	1:A:499:LEU:HD12	1.84	0.40

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	562/669 (84%)	528 (94%)	27 (5%)	7 (1%)	16	33
2	B	261/333 (78%)	239 (92%)	15 (6%)	7 (3%)	6	10
All	All	823/1002 (82%)	767 (93%)	42 (5%)	14 (2%)	11	22

All (14) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	506	VAL
2	B	530	PRO
2	B	774	THR
1	A	71	SER
1	A	97	ILE
1	A	505	GLU
2	B	781	THR
1	A	105	HIS
1	A	558	GLY
2	B	533	GLY
2	B	653	ASN
2	B	703	HIS
1	A	485	GLY
2	B	771	PRO

### 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	494/556 (89%)	465 (94%)	29 (6%)	24	47
2	B	226/272 (83%)	212 (94%)	14 (6%)	23	45
All	All	720/828 (87%)	677 (94%)	43 (6%)	24	47

All (43) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	GLN
1	A	27	VAL
1	A	31	LYS
1	A	39	LEU
1	A	58	ASP
1	A	63	LEU
1	A	85	ARG
1	A	96	ARG
1	A	114	ARG
1	A	115	ARG
1	A	117	LEU
1	A	124	GLU
1	A	137	LEU
1	A	162	LEU
1	A	176	LEU
1	A	297	LYS
1	A	298	ARG
1	A	323	LEU
1	A	324	LEU
1	A	426	TYR
1	A	452	GLN
1	A	495	LEU
1	A	499	LEU
1	A	509	GLN
1	A	524	LEU
1	A	533	LEU
1	A	603	THR
1	A	620	LEU
1	A	630	ARG
2	B	525	LEU
2	B	528	HIS
2	B	536	VAL
2	B	547	LEU
2	B	574	LEU
2	B	581	VAL
2	B	584	ARG
2	B	609	LEU
2	B	655	LYS
2	B	669	MSE
2	B	682	THR
2	B	697	LEU
2	B	758	LEU
2	B	762	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	452	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](#)

1 ligand is 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	$\# Z  > 2$
3	SO4	B	901	-	4,4,4	0.24	0	6,6,6	0.08	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
3	SO4	B	901	-	-	0/0/0/0	0/0/0/0

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

### 6.1 Protein, DNA and RNA chains

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	567/669 (84%)	0.27	38 (6%)	21 15	26, 48, 115, 157	0
2	B	258/333 (77%)	0.41	31 (12%)	6 4	39, 62, 115, 134	0
All	All	825/1002 (82%)	0.32	69 (8%)	14 9	26, 54, 115, 157	0

All (69) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	558	GLY	15.6
1	A	357	TYR	8.5
1	A	546	GLY	7.7
1	A	544	GLY	5.4
1	A	545	SER	5.3
1	A	104	SER	5.2
2	B	780	GLN	5.0
1	A	73	ARG	4.8
1	A	100	GLU	4.5
1	A	296	ARG	4.4
1	A	543	LYS	4.3
1	A	559	ALA	4.3
2	B	528	HIS	4.3
1	A	99	ARG	4.1
2	B	679	ARG	3.8
2	B	778	PRO	3.6
1	A	96	ARG	3.6
2	B	656	GLU	3.5
1	A	557	GLY	3.5
1	A	600	ASP	3.3
2	B	789	GLY	3.3
2	B	538	LEU	3.2
1	A	268	ASP	3.2
2	B	779	GLY	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	90	ILE	3.1
2	B	760	ALA	3.1
2	B	671	LEU	3.0
1	A	483	SER	3.0
1	A	131	ASP	2.9
1	A	480	SER	2.8
2	B	545	ASP	2.8
1	A	60	PHE	2.8
2	B	686	VAL	2.8
2	B	783	THR	2.8
2	B	788	CYS	2.7
1	A	556	GLY	2.7
2	B	653	ASN	2.7
1	A	361	GLN	2.7
2	B	532	ALA	2.7
2	B	773	CYS	2.6
1	A	76	VAL	2.6
2	B	534	ARG	2.6
2	B	521	LEU	2.5
2	B	523	THR	2.5
2	B	761	GLY	2.5
2	B	758	LEU	2.5
1	A	436	LYS	2.4
1	A	103	THR	2.4
2	B	687	GLY	2.4
1	A	61	PHE	2.4
2	B	787	LYS	2.4
1	A	122	VAL	2.4
1	A	132	ALA	2.4
1	A	360	GLU	2.3
2	B	786	GLU	2.3
1	A	126	ALA	2.3
2	B	782	ILE	2.3
2	B	685	PHE	2.2
2	B	654	GLN	2.2
1	A	118	PHE	2.2
2	B	785	TYR	2.1
1	A	509	GLN	2.1
1	A	68	THR	2.1
1	A	542	THR	2.1
2	B	771	PRO	2.1
1	A	62	PHE	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	527	ASN	2.0
1	A	63	LEU	2.0
1	A	92	ALA	2.0

## 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	B	901	5/5	0.96	0.13	-0.76	79,80,81,81	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.