



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 07:26 pm GMT

PDB ID : 4BUJ  
Title : Crystal structure of the S. cerevisiae Ski2-3-8 complex  
Authors : Halbach, F.; Reichelt, P.; Rode, M.; Conti, E.  
Deposited on : 2013-06-20  
Resolution : 3.70 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)	:	trunk28620

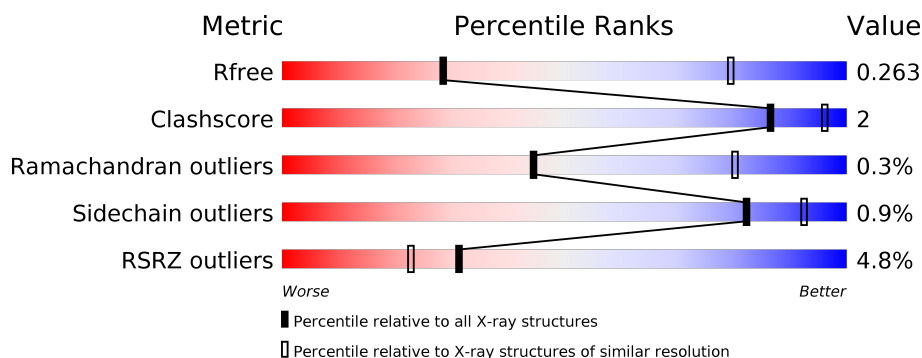
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.70 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	1234 (3.90-3.50)
Clashscore	112137	1377 (3.90-3.50)
Ramachandran outliers	110173	1323 (3.90-3.50)
Sidechain outliers	110143	1320 (3.90-3.50)
RSRZ outliers	101464	1262 (3.90-3.50)

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	1044	<div> <div></div> <div> <div></div> <div>80%</div> <div>6%</div> <div>14%</div> </div> </div>
1	E	1044	<div> <div>5%</div> <div> <div></div> <div>80%</div> <div>5%</div> <div>16%</div> </div> </div>
2	B	1436	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>8%</div> <div>5%</div> </div> </div>
2	F	1436	<div> <div>4%</div> <div> <div></div> <div>89%</div> <div>5%</div> <div>6%</div> </div> </div>
3	C	397	<div> <div></div> <div> <div></div> <div>90%</div> <div>8%</div> <div>..</div> </div> </div>
3	D	397	<div> <div>2%</div> <div> <div></div> <div>89%</div> <div>7%</div> <div>.</div> </div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	397	
3	H	397	

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	2004	-	-	-	X
4	SO4	C	1002	-	-	-	X
4	SO4	D	1001	-	-	-	X
4	SO4	D	1002	-	-	-	X
4	SO4	G	1001	-	-	-	X
4	SO4	G	1002	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 43019 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 ANTIVIRAL HELICASE SKI2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	903	Total	C	N	O	S	0	0	0
			6621	4236	1132	1223	30			
1	E	881	Total	C	N	O	S	0	0	0
			6016	3836	1046	1106	28			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	EXPRESSION TAG	UNP P35207
A	-2	PRO	-	EXPRESSION TAG	UNP P35207
A	-1	ASP	-	EXPRESSION TAG	UNP P35207
A	0	SER	-	EXPRESSION TAG	UNP P35207
A	835	GLY	-	LINKER	UNP P35207
A	836	SER	-	LINKER	UNP P35207
A	837	ARG	-	LINKER	UNP P35207
A	1085	GLY	-	LINKER	UNP P35207
E	-3	GLY	-	EXPRESSION TAG	UNP P35207
E	-2	PRO	-	EXPRESSION TAG	UNP P35207
E	-1	ASP	-	EXPRESSION TAG	UNP P35207
E	0	SER	-	EXPRESSION TAG	UNP P35207
E	835	GLY	-	LINKER	UNP P35207
E	836	SER	-	LINKER	UNP P35207
E	837	ARG	-	LINKER	UNP P35207
E	1085	GLY	-	LINKER	UNP P35207

- Molecule 2 is a protein called SUPERKILLER PROTEIN 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1367	Total	C	N	O	S	0	0	0
			9800	6281	1660	1822	37			
2	F	1351	Total	C	N	O	S	0	0	0
			9328	5966	1597	1733	32			

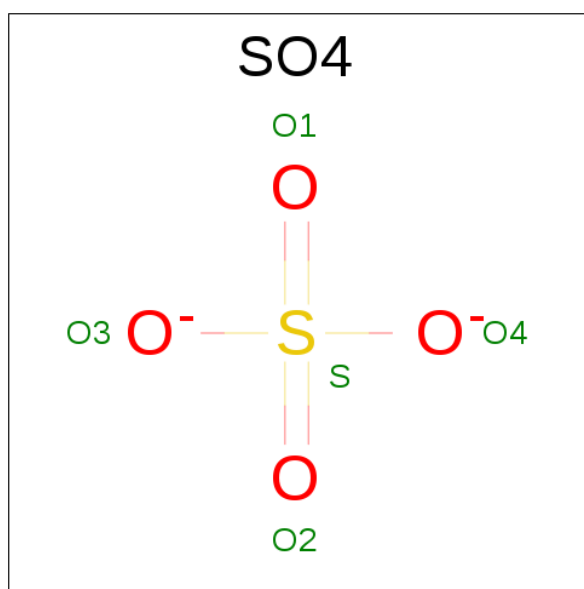
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	EXPRESSION TAG	UNP P17883
B	-2	PRO	-	EXPRESSION TAG	UNP P17883
B	-1	ASP	-	EXPRESSION TAG	UNP P17883
B	0	SER	-	EXPRESSION TAG	UNP P17883
F	-3	GLY	-	EXPRESSION TAG	UNP P17883
F	-2	PRO	-	EXPRESSION TAG	UNP P17883
F	-1	ASP	-	EXPRESSION TAG	UNP P17883
F	0	SER	-	EXPRESSION TAG	UNP P17883

- Molecule 3 is a protein called ANTIVIRAL PROTEIN SKI8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	392	Total	C	N	O	S	0	0	0
			2931	1860	499	558	14			
3	D	381	Total	C	N	O	S	0	0	0
			2790	1778	478	521	13			
3	G	387	Total	C	N	O	S	0	0	0
			2904	1846	497	547	14			
3	H	351	Total	C	N	O	S	0	0	0
			2529	1615	427	474	13			

- 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	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

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

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 ( $\text{RSRZ} > 2$ ). 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.

Chain A:

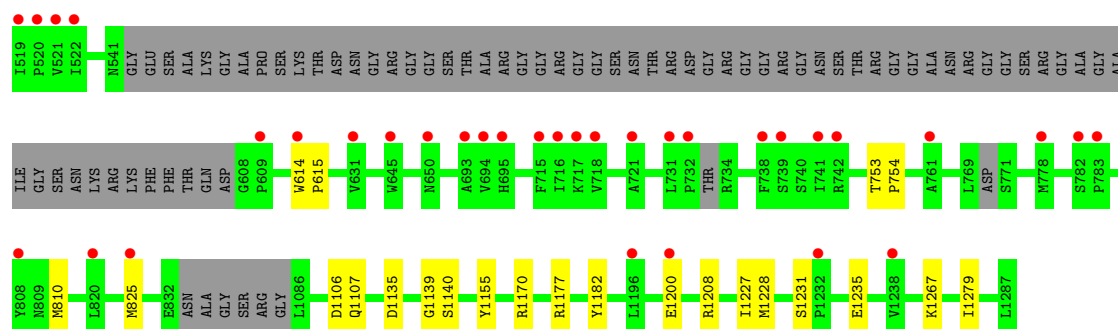
Sequence logo for Chain A showing amino acid conservation across 1000 positions. The y-axis represents information content in bits (0.00 to 1.73). The x-axis shows positions 1 to 1000. Amino acids are color-coded: yellow for common, green for conserved, and red for rare. The sequence is: GLY, PRO, ASP, SER, MET, SER, GLU, GLY, PHE, SER, SER, SER, S8, Y14, K18, A24, ASP, V26, E39, SER, THR, ASP, GLU, PRO, K45, W64, V70, P74, S78, PRO, GLU, ASP, CYS, SER, G84, E111, Y117, I162, V167, THR, ASP, SER, GLN, ASN, GLY, SER, D397, L417, L434, E445, V446, H447, Y448, V449, R454, F471, L472, L473, L474, S475, A476, T477, L488, P502, L512, W513, A514, K535, H536, L539, L540, H541, GLY, GLU, SER, ALA, LYS, GLY, ALA, PRO, LYS, THR, ASP, ASN, ARG, GLY, GLY, H638, E639, I702, F715, F720, L737, F738, D745, G746, G765, S782, P783, Y808, L820, H833, A834, GLY, SER, ARG, G1085, L1089, P1090, R1095, V1098, F1104, L1112, L1113, K1114, G1124, F1125, E1126, L1127, S1151, V1152, L1171, M1185, V1188, R1208, M1212, H1213, V1214, W1218, F1224, L1237.

Chain E:

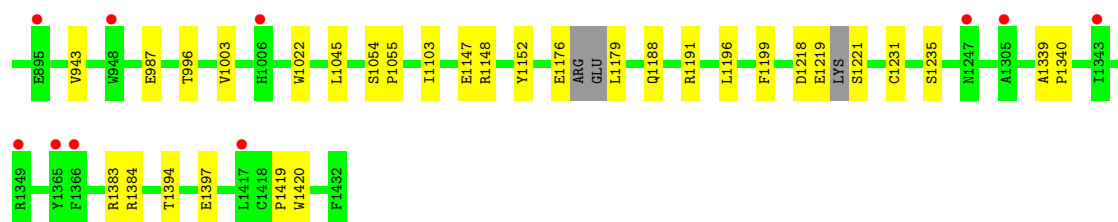
5% 80% 5% 16%

GLY PRO ASP MET SER SER GLU GLY PHE SER SER SER S9 L13 A24 ASP V26 R32 F39 SER THR ASP GLU PRO LYS H46 L66 M69 P74 HIS THR SER SER PRO GLU ASP CYS SER GLY LYS LEU ASP TYR R89 I98 F105 K106 R107 E111

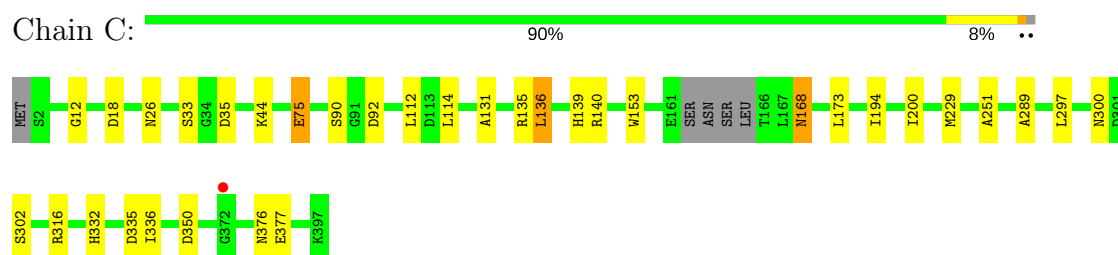
UNK X219 W302 F348 Y360 Y377 F399 D397 Q409 I410 M411 P412 L417 I418 R425 S426 M427 G431 F440 D444 W458 E459 E460 K470 T477 I499 S500 P505 L508 E509 I510 M511 I512 W513 A514 K515



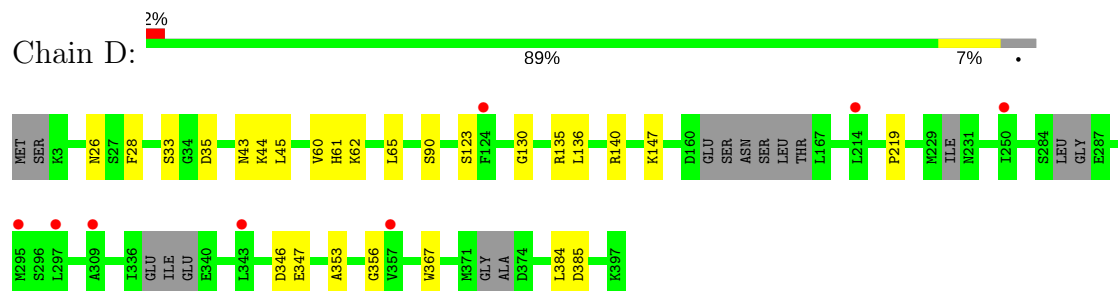




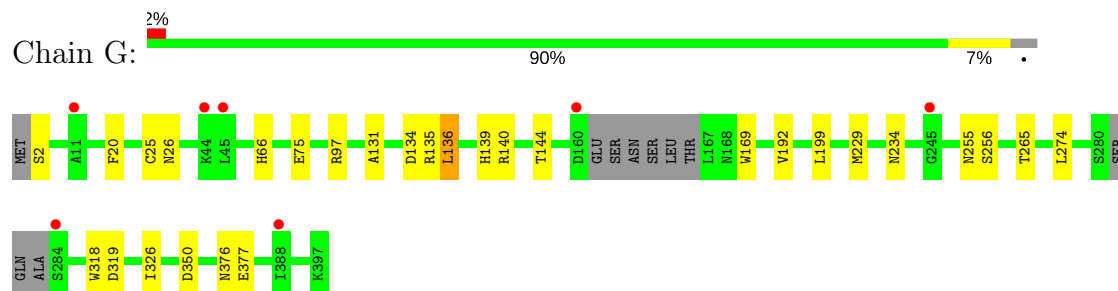
● Molecule 3: ANTIVIRAL PROTEIN SKI8



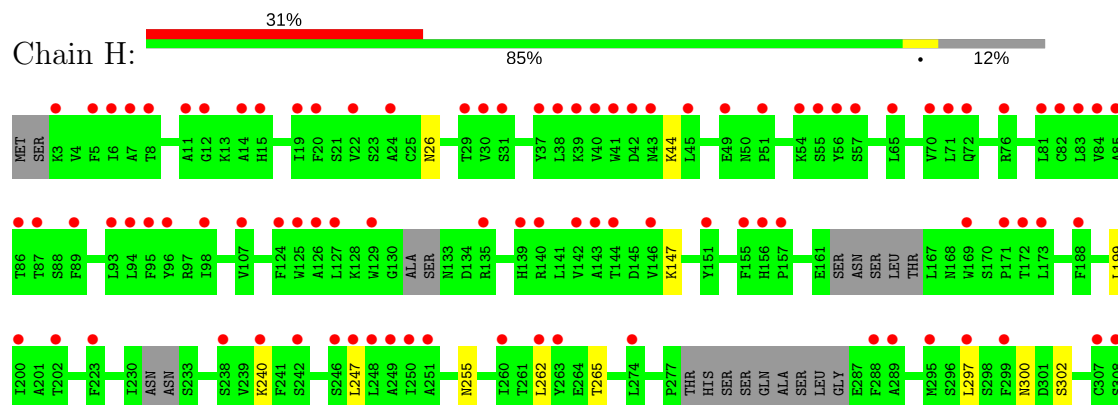
● Molecule 3: ANTIVIRAL PROTEIN SKI8

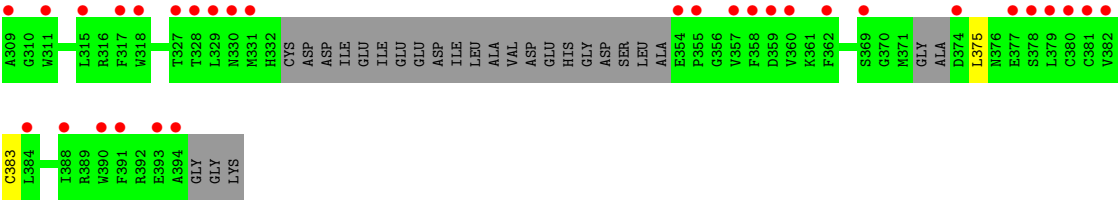


● Molecule 3: ANTIVIRAL PROTEIN SKI8



● Molecule 3: ANTIVIRAL PROTEIN SKI8





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	183.38Å 200.41Å 340.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	96.12 – 3.70 96.12 – 3.70	Depositor EDS
% Data completeness (in resolution range)	98.8 (96.12-3.70) 98.6 (96.12-3.70)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 3.67Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.231 , 0.265 0.228 , 0.263	Depositor DCC
$R_{free}$ test set	6658 reflections (5.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	139.9	Xtriage
Anisotropy	0.156	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 137.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	43019	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.78% 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: 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	0.28	0/6488	0.47	0/8829
1	E	0.25	0/5886	0.43	0/8055
2	B	0.27	0/10000	0.44	0/13675
2	F	0.26	0/9511	0.41	0/13043
3	C	0.29	0/2999	0.49	0/4079
3	D	0.27	0/2855	0.47	0/3885
3	G	0.28	0/2972	0.50	0/4039
3	H	0.24	0/2588	0.43	0/3536
All	All	0.27	0/43299	0.45	0/59141

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	6621	0	5996	41	0
1	E	6016	0	4947	25	0
2	B	9800	0	8525	61	0
2	F	9328	0	7703	38	0
3	C	2931	0	2691	22	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	2790	0	2487	11	0
3	G	2904	0	2658	14	0
3	H	2529	0	2218	6	0
4	A	20	0	0	2	0
4	B	5	0	0	0	0
4	C	20	0	0	0	0
4	D	10	0	0	0	0
4	E	15	0	0	0	0
4	G	20	0	0	0	0
4	H	10	0	0	0	0
All	All	43019	0	37225	199	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 2.

The worst 5 of 199 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:680:GLN:NE2	1:A:1235:GLU:OE1	2.17	0.77
3:D:356:GLY:O	3:D:384:LEU:N	2.22	0.72
2:B:1224:THR:O	2:B:1228:ASN:ND2	2.25	0.70
2:B:1383:ARG:NH1	3:C:350:ASP:OD2	2.30	0.64
1:A:380:PRO:O	1:A:454:ARG:NH2	2.31	0.63

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	835/1044 (80%)	797 (95%)	36 (4%)	2 (0%)	51	85
1	E	810/1044 (78%)	785 (97%)	24 (3%)	1 (0%)	55	88

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	1357/1436 (94%)	1298 (96%)	52 (4%)	7 (0%)	32	73
2	F	1333/1436 (93%)	1284 (96%)	48 (4%)	1 (0%)	55	88
3	C	388/397 (98%)	364 (94%)	23 (6%)	1 (0%)	44	80
3	D	369/397 (93%)	347 (94%)	17 (5%)	5 (1%)	13	55
3	G	381/397 (96%)	358 (94%)	22 (6%)	1 (0%)	44	80
3	H	337/397 (85%)	313 (93%)	23 (7%)	1 (0%)	44	80
All	All	5810/6548 (89%)	5546 (96%)	245 (4%)	19 (0%)	44	80

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	484	VAL
2	B	1239	LEU
3	C	75	GLU
3	D	44	LYS
3	H	255	ASN

### 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	614/832 (74%)	609 (99%)	5 (1%)	85	93
1	E	466/832 (56%)	461 (99%)	5 (1%)	78	90
2	B	836/1282 (65%)	826 (99%)	10 (1%)	75	90
2	F	715/1282 (56%)	712 (100%)	3 (0%)	93	97
3	C	295/347 (85%)	293 (99%)	2 (1%)	87	94
3	D	266/347 (77%)	263 (99%)	3 (1%)	78	90
3	G	289/347 (83%)	285 (99%)	4 (1%)	71	89
3	H	240/347 (69%)	238 (99%)	2 (1%)	85	93
All	All	3721/5616 (66%)	3687 (99%)	34 (1%)	82	92

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

Mol	Chain	Res	Type
3	C	136	LEU
3	D	136	LEU
3	G	234	ASN
3	D	28	PHE
2	B	458	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

20 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	2001	-	4,4,4	0.13	0	6,6,6	0.06	0
4	SO4	A	2002	-	4,4,4	0.17	0	6,6,6	0.07	0
4	SO4	A	2003	-	4,4,4	0.13	0	6,6,6	0.14	0
4	SO4	A	2004	-	4,4,4	0.15	0	6,6,6	0.06	0
4	SO4	B	2001	-	4,4,4	0.14	0	6,6,6	0.06	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SO4	C	1001	-	4,4,4	0.13	0	6,6,6	0.07	0
4	SO4	C	1002	-	4,4,4	0.15	0	6,6,6	0.07	0
4	SO4	C	1003	-	4,4,4	0.12	0	6,6,6	0.07	0
4	SO4	C	1004	-	4,4,4	0.15	0	6,6,6	0.11	0
4	SO4	D	1001	-	4,4,4	0.13	0	6,6,6	0.05	0
4	SO4	D	1002	-	4,4,4	0.14	0	6,6,6	0.10	0
4	SO4	E	2001	-	4,4,4	0.14	0	6,6,6	0.06	0
4	SO4	E	2002	-	4,4,4	0.16	0	6,6,6	0.06	0
4	SO4	E	2003	-	4,4,4	0.14	0	6,6,6	0.06	0
4	SO4	G	1001	-	4,4,4	0.13	0	6,6,6	0.07	0
4	SO4	G	1002	-	4,4,4	0.15	0	6,6,6	0.10	0
4	SO4	G	1003	-	4,4,4	0.15	0	6,6,6	0.05	0
4	SO4	G	1004	-	4,4,4	0.16	0	6,6,6	0.08	0
4	SO4	H	1001	-	4,4,4	0.14	0	6,6,6	0.06	0
4	SO4	H	1002	-	4,4,4	0.14	0	6,6,6	0.06	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
4	SO4	A	2001	-	-	0/0/0/0	0/0/0/0
4	SO4	A	2002	-	-	0/0/0/0	0/0/0/0
4	SO4	A	2003	-	-	0/0/0/0	0/0/0/0
4	SO4	A	2004	-	-	0/0/0/0	0/0/0/0
4	SO4	B	2001	-	-	0/0/0/0	0/0/0/0
4	SO4	C	1001	-	-	0/0/0/0	0/0/0/0
4	SO4	C	1002	-	-	0/0/0/0	0/0/0/0
4	SO4	C	1003	-	-	0/0/0/0	0/0/0/0
4	SO4	C	1004	-	-	0/0/0/0	0/0/0/0
4	SO4	D	1001	-	-	0/0/0/0	0/0/0/0
4	SO4	D	1002	-	-	0/0/0/0	0/0/0/0
4	SO4	E	2001	-	-	0/0/0/0	0/0/0/0
4	SO4	E	2002	-	-	0/0/0/0	0/0/0/0
4	SO4	E	2003	-	-	0/0/0/0	0/0/0/0
4	SO4	G	1001	-	-	0/0/0/0	0/0/0/0
4	SO4	G	1002	-	-	0/0/0/0	0/0/0/0
4	SO4	G	1003	-	-	0/0/0/0	0/0/0/0
4	SO4	G	1004	-	-	0/0/0/0	0/0/0/0
4	SO4	H	1001	-	-	0/0/0/0	0/0/0/0
4	SO4	H	1002	-	-	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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	2001	SO4	1	0
4	A	2002	SO4	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 ⓘ

### 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	851/1044 (81%)	0.19	15 (1%) 69 56	49, 86, 153, 361	0
1	E	832/1044 (79%)	0.26	53 (6%) 20 14	79, 150, 224, 363	0
2	B	1367/1436 (95%)	0.11	25 (1%) 69 56	51, 101, 164, 277	0
2	F	1351/1436 (94%)	0.18	52 (3%) 41 30	61, 128, 199, 310	0
3	C	392/397 (98%)	0.35	1 (0%) 93 90	40, 67, 140, 322	0
3	D	381/397 (95%)	0.32	8 (2%) 64 51	56, 92, 169, 265	0
3	G	387/397 (97%)	0.37	7 (1%) 69 56	45, 87, 151, 235	0
3	H	351/397 (88%)	1.50	124 (35%) 0 1	115, 180, 260, 351	0
All	All	5912/6548 (90%)	0.29	285 (4%) 31 23	40, 112, 203, 363	0

The worst 5 of 285 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	H	11	ALA	8.6
3	H	12	GLY	6.8
3	H	355	PRO	6.7
3	H	24	ALA	6.6
3	H	394	ALA	6.0

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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	1002	5/5	0.83	0.79	6.52	162,162,162,162	0
4	SO4	A	2004	5/5	0.44	0.60	3.50	205,205,205,205	0
4	SO4	G	1002	5/5	0.77	0.69	3.08	151,151,151,151	0
4	SO4	G	1001	5/5	0.86	0.44	1.53	178,178,178,178	0
4	SO4	G	1004	5/5	0.68	0.37	1.46	191,191,191,191	0
4	SO4	C	1001	5/5	0.84	0.39	0.75	173,173,173,173	0
4	SO4	D	1001	5/5	0.77	0.46	0.52	174,174,174,174	0
4	SO4	D	1002	5/5	0.87	0.48	0.30	142,142,142,142	0
4	SO4	E	2003	5/5	0.80	0.28	0.06	252,252,252,252	0
4	SO4	G	1003	5/5	0.85	0.30	-0.24	142,142,142,142	0
4	SO4	A	2003	5/5	0.94	0.21	-0.67	97,97,97,97	0
4	SO4	A	2001	5/5	0.84	0.22	-0.77	175,175,175,175	0
4	SO4	C	1004	5/5	0.89	0.26	-0.92	182,182,182,182	0
4	SO4	E	2001	5/5	0.81	0.16	-1.50	238,238,238,238	0
4	SO4	H	1001	5/5	0.37	0.22	-1.50	256,256,256,256	0
4	SO4	A	2002	5/5	0.91	0.16	-	133,133,133,133	0
4	SO4	B	2001	5/5	0.87	0.40	-	212,212,212,212	0
4	SO4	E	2002	5/5	0.92	0.13	-	208,208,208,208	0
4	SO4	H	1002	5/5	0.77	0.18	-	267,267,267,267	0
4	SO4	C	1003	5/5	0.89	0.24	-	122,122,122,122	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.