



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 13, 2018 – 06:25 pm GMT

PDB ID : 4RDO  
Title : Structure of YTH-YTHDF2 in the free state  
Authors : Li, F.D.; Zhao, D.B.; Wu, J.H.; Shi, Y.Y.  
Deposited on : 2014-09-19  
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](mailto: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 ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk31020  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk31020

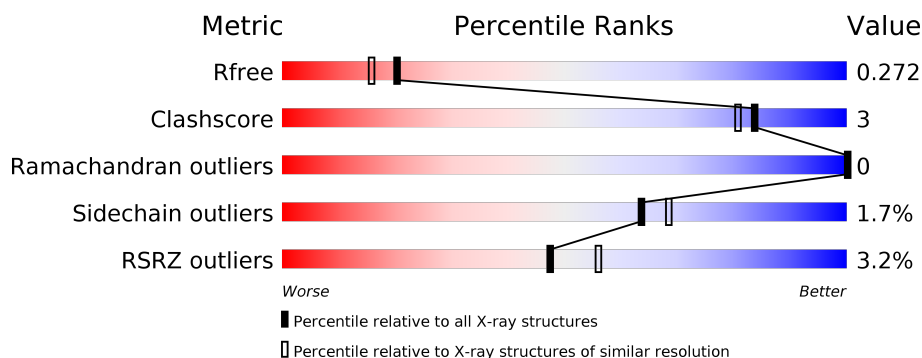
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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))
$R_{free}$	111664	1287 (2.16-2.16)
Clashscore	122126	1390 (2.16-2.16)
Ramachandran outliers	120053	1368 (2.16-2.16)
Sidechain outliers	120020	1367 (2.16-2.16)
RSRZ outliers	108989	1262 (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 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	167	<div> <div>78%</div> <div>10% • 11%</div> </div>
1	B	167	<div> <div>%</div> <div>84%</div> <div>• 13%</div> </div>
1	C	167	<div> <div>6%</div> <div>83%</div> <div>6% 11%</div> </div>
1	D	167	<div> <div>3%</div> <div>83%</div> <div>• 13%</div> </div>
1	E	167	<div> <div>4%</div> <div>83%</div> <div>• 16%</div> </div>
1	F	167	<div> <div>3%</div> <div>72%</div> <div>11% • 16%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7542 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 YTH domain-containing family protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	149	Total	C	N	O	S	0	0	0
			1217	770	221	220	6			
1	B	146	Total	C	N	O	S	0	0	0
			1192	755	217	214	6			
1	C	148	Total	C	N	O	S	0	0	0
			1212	767	220	219	6			
1	D	146	Total	C	N	O	S	0	0	0
			1195	758	217	214	6			
1	E	141	Total	C	N	O	S	0	0	0
			1141	720	208	207	6			
1	F	140	Total	C	N	O	S	0	0	0
			1133	715	207	206	5			

There are 132 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	386	MET	-	EXPRESSION TAG	UNP Q9Y5A9
A	387	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
A	388	SER	-	EXPRESSION TAG	UNP Q9Y5A9
A	389	SER	-	EXPRESSION TAG	UNP Q9Y5A9
A	390	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
A	391	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
A	392	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
A	393	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
A	394	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
A	395	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
A	396	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
A	397	SER	-	EXPRESSION TAG	UNP Q9Y5A9
A	398	SER	-	EXPRESSION TAG	UNP Q9Y5A9
A	399	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
A	400	GLU	-	EXPRESSION TAG	UNP Q9Y5A9
A	401	ASN	-	EXPRESSION TAG	UNP Q9Y5A9
A	402	LEU	-	EXPRESSION TAG	UNP Q9Y5A9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	403	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
A	404	PHE	-	EXPRESSION TAG	UNP Q9Y5A9
A	405	GLN	-	EXPRESSION TAG	UNP Q9Y5A9
A	406	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
A	407	MET	-	EXPRESSION TAG	UNP Q9Y5A9
B	386	MET	-	EXPRESSION TAG	UNP Q9Y5A9
B	387	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
B	388	SER	-	EXPRESSION TAG	UNP Q9Y5A9
B	389	SER	-	EXPRESSION TAG	UNP Q9Y5A9
B	390	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
B	391	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
B	392	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
B	393	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
B	394	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
B	395	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
B	396	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
B	397	SER	-	EXPRESSION TAG	UNP Q9Y5A9
B	398	SER	-	EXPRESSION TAG	UNP Q9Y5A9
B	399	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
B	400	GLU	-	EXPRESSION TAG	UNP Q9Y5A9
B	401	ASN	-	EXPRESSION TAG	UNP Q9Y5A9
B	402	LEU	-	EXPRESSION TAG	UNP Q9Y5A9
B	403	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
B	404	PHE	-	EXPRESSION TAG	UNP Q9Y5A9
B	405	GLN	-	EXPRESSION TAG	UNP Q9Y5A9
B	406	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
B	407	MET	-	EXPRESSION TAG	UNP Q9Y5A9
C	386	MET	-	EXPRESSION TAG	UNP Q9Y5A9
C	387	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
C	388	SER	-	EXPRESSION TAG	UNP Q9Y5A9
C	389	SER	-	EXPRESSION TAG	UNP Q9Y5A9
C	390	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
C	391	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
C	392	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
C	393	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
C	394	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
C	395	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
C	396	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
C	397	SER	-	EXPRESSION TAG	UNP Q9Y5A9
C	398	SER	-	EXPRESSION TAG	UNP Q9Y5A9
C	399	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
C	400	GLU	-	EXPRESSION TAG	UNP Q9Y5A9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	401	ASN	-	EXPRESSION TAG	UNP Q9Y5A9
C	402	LEU	-	EXPRESSION TAG	UNP Q9Y5A9
C	403	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
C	404	PHE	-	EXPRESSION TAG	UNP Q9Y5A9
C	405	GLN	-	EXPRESSION TAG	UNP Q9Y5A9
C	406	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
C	407	MET	-	EXPRESSION TAG	UNP Q9Y5A9
D	386	MET	-	EXPRESSION TAG	UNP Q9Y5A9
D	387	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
D	388	SER	-	EXPRESSION TAG	UNP Q9Y5A9
D	389	SER	-	EXPRESSION TAG	UNP Q9Y5A9
D	390	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
D	391	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
D	392	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
D	393	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
D	394	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
D	395	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
D	396	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
D	397	SER	-	EXPRESSION TAG	UNP Q9Y5A9
D	398	SER	-	EXPRESSION TAG	UNP Q9Y5A9
D	399	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
D	400	GLU	-	EXPRESSION TAG	UNP Q9Y5A9
D	401	ASN	-	EXPRESSION TAG	UNP Q9Y5A9
D	402	LEU	-	EXPRESSION TAG	UNP Q9Y5A9
D	403	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
D	404	PHE	-	EXPRESSION TAG	UNP Q9Y5A9
D	405	GLN	-	EXPRESSION TAG	UNP Q9Y5A9
D	406	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
D	407	MET	-	EXPRESSION TAG	UNP Q9Y5A9
E	386	MET	-	EXPRESSION TAG	UNP Q9Y5A9
E	387	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
E	388	SER	-	EXPRESSION TAG	UNP Q9Y5A9
E	389	SER	-	EXPRESSION TAG	UNP Q9Y5A9
E	390	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
E	391	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
E	392	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
E	393	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
E	394	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
E	395	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
E	396	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
E	397	SER	-	EXPRESSION TAG	UNP Q9Y5A9
E	398	SER	-	EXPRESSION TAG	UNP Q9Y5A9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	399	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
E	400	GLU	-	EXPRESSION TAG	UNP Q9Y5A9
E	401	ASN	-	EXPRESSION TAG	UNP Q9Y5A9
E	402	LEU	-	EXPRESSION TAG	UNP Q9Y5A9
E	403	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
E	404	PHE	-	EXPRESSION TAG	UNP Q9Y5A9
E	405	GLN	-	EXPRESSION TAG	UNP Q9Y5A9
E	406	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
E	407	MET	-	EXPRESSION TAG	UNP Q9Y5A9
F	386	MET	-	EXPRESSION TAG	UNP Q9Y5A9
F	387	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
F	388	SER	-	EXPRESSION TAG	UNP Q9Y5A9
F	389	SER	-	EXPRESSION TAG	UNP Q9Y5A9
F	390	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
F	391	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
F	392	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
F	393	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
F	394	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
F	395	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
F	396	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
F	397	SER	-	EXPRESSION TAG	UNP Q9Y5A9
F	398	SER	-	EXPRESSION TAG	UNP Q9Y5A9
F	399	GLY	-	EXPRESSION TAG	UNP Q9Y5A9
F	400	GLU	-	EXPRESSION TAG	UNP Q9Y5A9
F	401	ASN	-	EXPRESSION TAG	UNP Q9Y5A9
F	402	LEU	-	EXPRESSION TAG	UNP Q9Y5A9
F	403	TYR	-	EXPRESSION TAG	UNP Q9Y5A9
F	404	PHE	-	EXPRESSION TAG	UNP Q9Y5A9
F	405	GLN	-	EXPRESSION TAG	UNP Q9Y5A9
F	406	HIS	-	EXPRESSION TAG	UNP Q9Y5A9
F	407	MET	-	EXPRESSION TAG	UNP Q9Y5A9

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	E	1	Total	O	S	0	0
			5	4	1		
2	F	1	Total	O	S	0	0
			5	4	1		


- Molecule 3 is water.

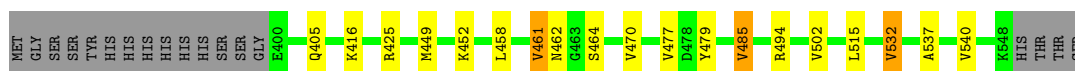
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	90	Total	O	0	0
			90	90		
3	B	85	Total	O	0	0
			85	85		
3	C	65	Total	O	0	0
			65	65		
3	D	69	Total	O	0	0
			69	69		
3	E	57	Total	O	0	0
			57	57		
3	F	56	Total	O	0	0
			56	56		

### 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 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.

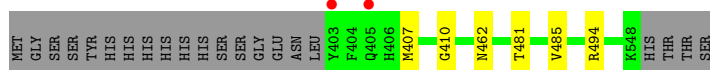
- Molecule 1: YTH domain-containing family protein 2

Chain A: 




- Molecule 1: YTH domain-containing family protein 2

Chain B: 




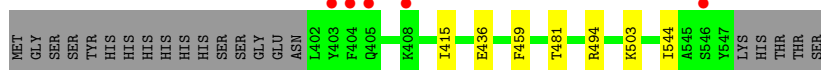
- Molecule 1: YTH domain-containing family protein 2

Chain C: 




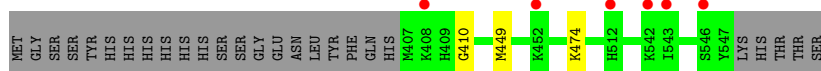
- Molecule 1: YTH domain-containing family protein 2

Chain D: 



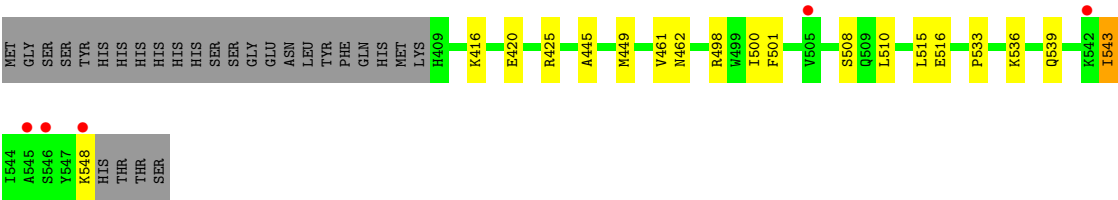
- Molecule 1: YTH domain-containing family protein 2

Chain E: 



- Molecule 1: YTH domain-containing family protein 2

Chain F: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	139.28Å 139.28Å 112.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.28 – 2.15 42.28 – 2.15	Depositor EDS
% Data completeness (in resolution range)	97.0 (42.28-2.15) 97.0 (42.28-2.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.87 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.213 , 0.270 0.217 , 0.272	Depositor DCC
$R_{free}$ test set	3329 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.7	Xtriage
Anisotropy	0.030	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 48.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.045 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7542	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.28 % 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 5.9815e-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, 0.333 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	0.47	0/1247	0.67	0/1681
1	B	0.47	0/1222	0.64	0/1647
1	C	0.47	0/1242	0.62	0/1674
1	D	0.46	0/1225	0.67	0/1651
1	E	0.46	0/1168	0.62	0/1575
1	F	0.48	0/1160	0.64	0/1565
All	All	0.47	0/7264	0.64	0/9793

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	1217	0	1187	11	0
1	B	1192	0	1164	3	0
1	C	1212	0	1185	6	0
1	D	1195	0	1173	3	0
1	E	1141	0	1118	2	0
1	F	1133	0	1109	11	0
2	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
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	5	0	0	0	0
3	A	90	0	0	1	0
3	B	85	0	0	1	0
3	C	65	0	0	0	0
3	D	69	0	0	0	1
3	E	57	0	0	1	0
3	F	56	0	0	2	0
All	All	7542	0	6936	36	1

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 36 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:528:ASP:O	1:C:529:THR:HB	1.84	0.78
1:B:481:THR:HG21	1:B:494:ARG:HG2	1.71	0.71
1:A:532:VAL:CG2	1:A:537:ALA:HB2	2.21	0.70
1:C:481:THR:OG1	1:C:492:LYS:O	2.10	0.70
1:F:516:GLU:OE1	1:F:536:LYS:NZ	2.22	0.69

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:733:HOH:O	3:D:767:HOH:O[6_554]	2.18	0.02

## 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	147/167 (88%)	144 (98%)	3 (2%)	0	100	100
1	B	144/167 (86%)	142 (99%)	2 (1%)	0	100	100
1	C	146/167 (87%)	139 (95%)	7 (5%)	0	100	100
1	D	144/167 (86%)	139 (96%)	5 (4%)	0	100	100
1	E	139/167 (83%)	134 (96%)	5 (4%)	0	100	100
1	F	138/167 (83%)	134 (97%)	4 (3%)	0	100	100
All	All	858/1002 (86%)	832 (97%)	26 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	132/149 (89%)	125 (95%)	7 (5%)	25	20
1	B	129/149 (87%)	127 (98%)	2 (2%)	65	71
1	C	132/149 (89%)	131 (99%)	1 (1%)	83	87
1	D	130/149 (87%)	129 (99%)	1 (1%)	83	87
1	E	124/149 (83%)	124 (100%)	0	100	100
1	F	123/149 (83%)	121 (98%)	2 (2%)	65	71
All	All	770/894 (86%)	757 (98%)	13 (2%)	63	68

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	515	LEU
1	A	532	VAL
1	D	436	GLU
1	A	502	VAL
1	C	475	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	405	GLN
1	F	539	GLN
1	F	462	ASN
1	A	480	ASN
1	E	509	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](#)

6 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	$\# Z  > 2$
2	SO4	A	601	-	4,4,4	0.43	0	6,6,6	0.43	0
2	SO4	B	601	-	4,4,4	0.38	0	6,6,6	0.33	0
2	SO4	C	601	-	4,4,4	0.47	0	6,6,6	0.46	0
2	SO4	D	601	-	4,4,4	0.48	0	6,6,6	0.50	0
2	SO4	E	601	-	4,4,4	0.41	0	6,6,6	0.29	0
2	SO4	F	601	-	4,4,4	0.47	0	6,6,6	0.23	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
2	SO4	A	601	-	-	0/0/0/0	0/0/0/0
2	SO4	B	601	-	-	0/0/0/0	0/0/0/0
2	SO4	C	601	-	-	0/0/0/0	0/0/0/0
2	SO4	D	601	-	-	0/0/0/0	0/0/0/0
2	SO4	E	601	-	-	0/0/0/0	0/0/0/0
2	SO4	F	601	-	-	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 [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	149/167 (89%)	-0.18	0 <b>100</b> <b>100</b>	17, 28, 43, 52	0
1	B	146/167 (87%)	-0.02	2 (1%) 75 80	16, 27, 48, 69	0
1	C	148/167 (88%)	0.25	10 (6%) 17 23	18, 37, 57, 86	0
1	D	146/167 (87%)	-0.01	5 (3%) 45 53	17, 32, 58, 69	0
1	E	141/167 (84%)	0.09	6 (4%) 35 44	20, 32, 54, 70	0
1	F	140/167 (83%)	0.23	5 (3%) 42 50	17, 33, 56, 66	0
All	All	870/1002 (86%)	0.06	28 (3%) 47 56	16, 31, 54, 86	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	404	PHE	6.6
1	D	403	TYR	3.7
1	C	405	GLN	3.5
1	C	501	PHE	3.4
1	C	547	TYR	3.4

### 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	SO4	F	601	5/5	0.91	0.10	54,58,60,62	0
2	SO4	E	601	5/5	0.93	0.11	56,57,58,62	0
2	SO4	A	601	5/5	0.96	0.07	40,41,44,44	0
2	SO4	C	601	5/5	0.96	0.11	41,42,44,46	0
2	SO4	D	601	5/5	0.98	0.09	38,41,42,42	0
2	SO4	B	601	5/5	0.98	0.10	33,34,34,35	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.