



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 01:35 pm BST

PDB ID : 5MKC  
Title : Crystal structure of the RrgA Jo.In complex  
Authors : Bonnet, J.; Cartannaz, J.; Tourcier, G.; Contreras-Martel, C.; Kleman, J.P.;  
Fenel, D.; Schoehn, G.; Morlot, C.; Vernet, T.; Di Guilmi, A.M.  
Deposited on : 2016-12-03  
Resolution : 2.04 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

<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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

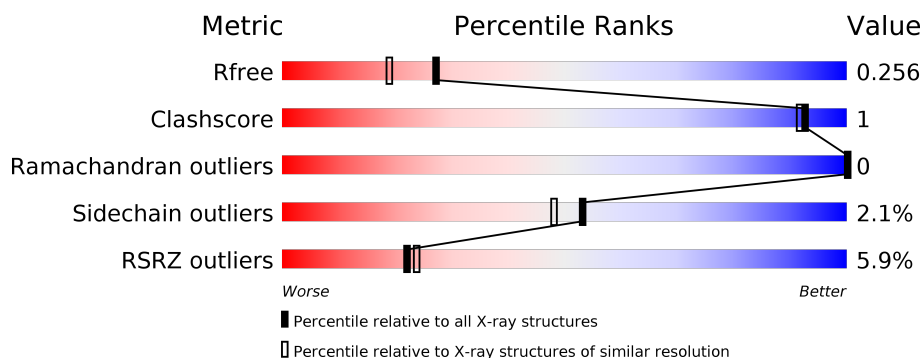
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.04 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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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 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  $\leq 5\%$ . 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	232	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>• • 11%</div> </div> </div>
1	B	232	<div> <div>7%</div> <div> <div></div> <div>85%</div> <div>• • 11%</div> </div> </div>
1	C	232	<div> <div>3%</div> <div> <div></div> <div>85%</div> <div>• • 11%</div> </div> </div>
1	D	232	<div> <div>3%</div> <div> <div></div> <div>85%</div> <div>• • 11%</div> </div> </div>
1	E	232	<div> <div>9%</div> <div> <div></div> <div>86%</div> <div>• 11%</div> </div> </div>
1	F	232	<div> <div>7%</div> <div> <div></div> <div>85%</div> <div>• • 11%</div> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10949 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 Cell wall surface anchor family protein (Jo), Cell wall surface anchor family protein (In).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	207	Total	C	N	O	S	0	2	0
			1644	1026	280	337	1			
1	B	206	Total	C	N	O	S	0	1	0
			1633	1017	279	336	1			
1	C	206	Total	C	N	O	S	0	2	0
			1638	1022	280	335	1			
1	D	207	Total	C	N	O	S	0	2	0
			1646	1026	281	338	1			
1	E	207	Total	C	N	O	S	0	1	0
			1642	1022	281	338	1			
1	F	207	Total	C	N	O	S	0	1	0
			1642	1022	281	338	1			

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	128	MET	-	initiating methionine	UNP A0A0Y1GQ57
A	129	GLY	-	expression tag	UNP A0A0Y1GQ57
A	130	SER	-	expression tag	UNP A0A0Y1GQ57
A	131	SER	-	expression tag	UNP A0A0Y1GQ57
A	132	HIS	-	expression tag	UNP A0A0Y1GQ57
A	133	HIS	-	expression tag	UNP A0A0Y1GQ57
A	134	HIS	-	expression tag	UNP A0A0Y1GQ57
A	135	HIS	-	expression tag	UNP A0A0Y1GQ57
A	136	HIS	-	expression tag	UNP A0A0Y1GQ57
A	137	HIS	-	expression tag	UNP A0A0Y1GQ57
A	138	SER	-	expression tag	UNP A0A0Y1GQ57
A	139	GLN	-	expression tag	UNP A0A0Y1GQ57
A	140	ASP	-	expression tag	UNP A0A0Y1GQ57
A	141	PRO	-	expression tag	UNP A0A0Y1GQ57
A	584	MET	-	linker	UNP A0A0Y1GQ57
A	585	ALA	-	linker	UNP A0A0Y1GQ57

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	586	ASP	-	linker	UNP A0A0Y1GQ57
A	587	LEU	-	linker	UNP A0A0Y1GQ57
B	128	MET	-	initiating methionine	UNP A0A0Y1GQ57
B	129	GLY	-	expression tag	UNP A0A0Y1GQ57
B	130	SER	-	expression tag	UNP A0A0Y1GQ57
B	131	SER	-	expression tag	UNP A0A0Y1GQ57
B	132	HIS	-	expression tag	UNP A0A0Y1GQ57
B	133	HIS	-	expression tag	UNP A0A0Y1GQ57
B	134	HIS	-	expression tag	UNP A0A0Y1GQ57
B	135	HIS	-	expression tag	UNP A0A0Y1GQ57
B	136	HIS	-	expression tag	UNP A0A0Y1GQ57
B	137	HIS	-	expression tag	UNP A0A0Y1GQ57
B	138	SER	-	expression tag	UNP A0A0Y1GQ57
B	139	GLN	-	expression tag	UNP A0A0Y1GQ57
B	140	ASP	-	expression tag	UNP A0A0Y1GQ57
B	141	PRO	-	expression tag	UNP A0A0Y1GQ57
B	584	MET	-	linker	UNP A0A0Y1GQ57
B	585	ALA	-	linker	UNP A0A0Y1GQ57
B	586	ASP	-	linker	UNP A0A0Y1GQ57
B	587	LEU	-	linker	UNP A0A0Y1GQ57
C	128	MET	-	initiating methionine	UNP A0A0Y1GQ57
C	129	GLY	-	expression tag	UNP A0A0Y1GQ57
C	130	SER	-	expression tag	UNP A0A0Y1GQ57
C	131	SER	-	expression tag	UNP A0A0Y1GQ57
C	132	HIS	-	expression tag	UNP A0A0Y1GQ57
C	133	HIS	-	expression tag	UNP A0A0Y1GQ57
C	134	HIS	-	expression tag	UNP A0A0Y1GQ57
C	135	HIS	-	expression tag	UNP A0A0Y1GQ57
C	136	HIS	-	expression tag	UNP A0A0Y1GQ57
C	137	HIS	-	expression tag	UNP A0A0Y1GQ57
C	138	SER	-	expression tag	UNP A0A0Y1GQ57
C	139	GLN	-	expression tag	UNP A0A0Y1GQ57
C	140	ASP	-	expression tag	UNP A0A0Y1GQ57
C	141	PRO	-	expression tag	UNP A0A0Y1GQ57
C	584	MET	-	linker	UNP A0A0Y1GQ57
C	585	ALA	-	linker	UNP A0A0Y1GQ57
C	586	ASP	-	linker	UNP A0A0Y1GQ57
C	587	LEU	-	linker	UNP A0A0Y1GQ57
D	128	MET	-	initiating methionine	UNP A0A0Y1GQ57
D	129	GLY	-	expression tag	UNP A0A0Y1GQ57
D	130	SER	-	expression tag	UNP A0A0Y1GQ57
D	131	SER	-	expression tag	UNP A0A0Y1GQ57

*Continued on next page...*

*Continued from previous page...*

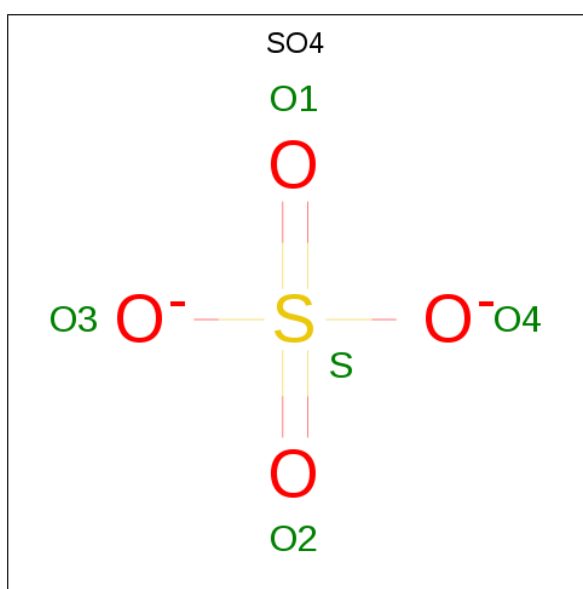
Chain	Residue	Modelled	Actual	Comment	Reference
D	132	HIS	-	expression tag	UNP A0A0Y1GQ57
D	133	HIS	-	expression tag	UNP A0A0Y1GQ57
D	134	HIS	-	expression tag	UNP A0A0Y1GQ57
D	135	HIS	-	expression tag	UNP A0A0Y1GQ57
D	136	HIS	-	expression tag	UNP A0A0Y1GQ57
D	137	HIS	-	expression tag	UNP A0A0Y1GQ57
D	138	SER	-	expression tag	UNP A0A0Y1GQ57
D	139	GLN	-	expression tag	UNP A0A0Y1GQ57
D	140	ASP	-	expression tag	UNP A0A0Y1GQ57
D	141	PRO	-	expression tag	UNP A0A0Y1GQ57
D	584	MET	-	linker	UNP A0A0Y1GQ57
D	585	ALA	-	linker	UNP A0A0Y1GQ57
D	586	ASP	-	linker	UNP A0A0Y1GQ57
D	587	LEU	-	linker	UNP A0A0Y1GQ57
E	128	MET	-	initiating methionine	UNP A0A0Y1GQ57
E	129	GLY	-	expression tag	UNP A0A0Y1GQ57
E	130	SER	-	expression tag	UNP A0A0Y1GQ57
E	131	SER	-	expression tag	UNP A0A0Y1GQ57
E	132	HIS	-	expression tag	UNP A0A0Y1GQ57
E	133	HIS	-	expression tag	UNP A0A0Y1GQ57
E	134	HIS	-	expression tag	UNP A0A0Y1GQ57
E	135	HIS	-	expression tag	UNP A0A0Y1GQ57
E	136	HIS	-	expression tag	UNP A0A0Y1GQ57
E	137	HIS	-	expression tag	UNP A0A0Y1GQ57
E	138	SER	-	expression tag	UNP A0A0Y1GQ57
E	139	GLN	-	expression tag	UNP A0A0Y1GQ57
E	140	ASP	-	expression tag	UNP A0A0Y1GQ57
E	141	PRO	-	expression tag	UNP A0A0Y1GQ57
E	584	MET	-	linker	UNP A0A0Y1GQ57
E	585	ALA	-	linker	UNP A0A0Y1GQ57
E	586	ASP	-	linker	UNP A0A0Y1GQ57
E	587	LEU	-	linker	UNP A0A0Y1GQ57
F	128	MET	-	initiating methionine	UNP A0A0Y1GQ57
F	129	GLY	-	expression tag	UNP A0A0Y1GQ57
F	130	SER	-	expression tag	UNP A0A0Y1GQ57
F	131	SER	-	expression tag	UNP A0A0Y1GQ57
F	132	HIS	-	expression tag	UNP A0A0Y1GQ57
F	133	HIS	-	expression tag	UNP A0A0Y1GQ57
F	134	HIS	-	expression tag	UNP A0A0Y1GQ57
F	135	HIS	-	expression tag	UNP A0A0Y1GQ57
F	136	HIS	-	expression tag	UNP A0A0Y1GQ57
F	137	HIS	-	expression tag	UNP A0A0Y1GQ57

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	138	SER	-	expression tag	UNP A0A0Y1GQ57
F	139	GLN	-	expression tag	UNP A0A0Y1GQ57
F	140	ASP	-	expression tag	UNP A0A0Y1GQ57
F	141	PRO	-	expression tag	UNP A0A0Y1GQ57
F	584	MET	-	linker	UNP A0A0Y1GQ57
F	585	ALA	-	linker	UNP A0A0Y1GQ57
F	586	ASP	-	linker	UNP A0A0Y1GQ57
F	587	LEU	-	linker	UNP A0A0Y1GQ57

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



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

Continued on next page...

*Continued from previous page...*

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

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ni	0	0
			1	1		
3	D	1	Total	Ni	0	0
			1	1		
3	F	1	Total	Ni	0	0
			1	1		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	1	Total	Ca	0	0
			1	1		
4	F	1	Total	Ca	0	0
			1	1		
4	E	1	Total	Ca	0	0
			1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	197	Total	O	0	0
			197	197		

*Continued on next page...*

*Continued from previous page...*

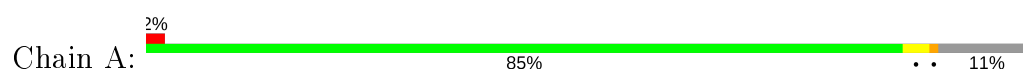
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	120	Total 120	O 120	0	0
5	C	191	Total 191	O 191	0	0
5	D	198	Total 198	O 198	0	0
5	E	146	Total 146	O 146	0	0
5	F	171	Total 171	O 171	0	0



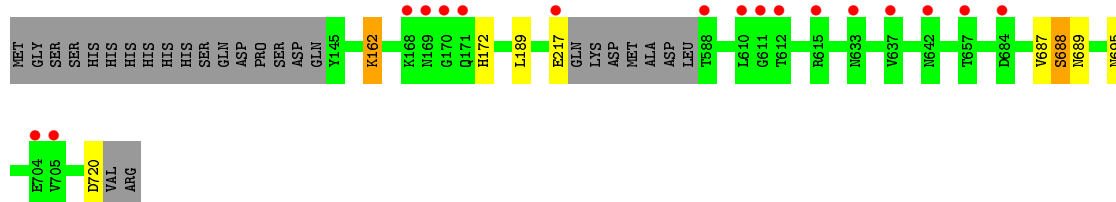
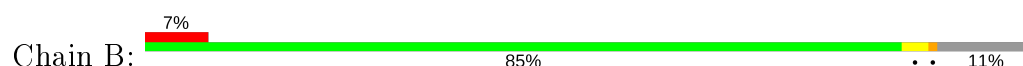
### 3 Residue-property plots [i](#)

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.

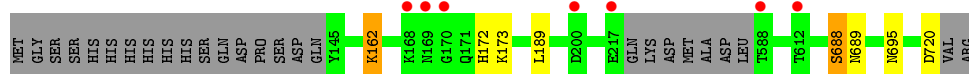
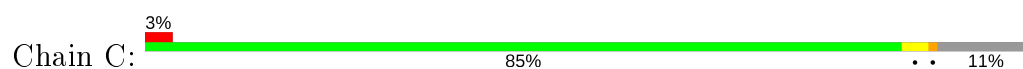
- Molecule 1: Cell wall surface anchor family protein (Jo),Cell wall surface anchor family protein (In)



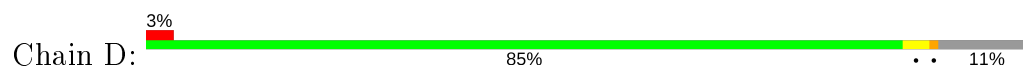
- Molecule 1: Cell wall surface anchor family protein (Jo),Cell wall surface anchor family protein (In)




- Molecule 1: Cell wall surface anchor family protein (Jo),Cell wall surface anchor family protein (In)

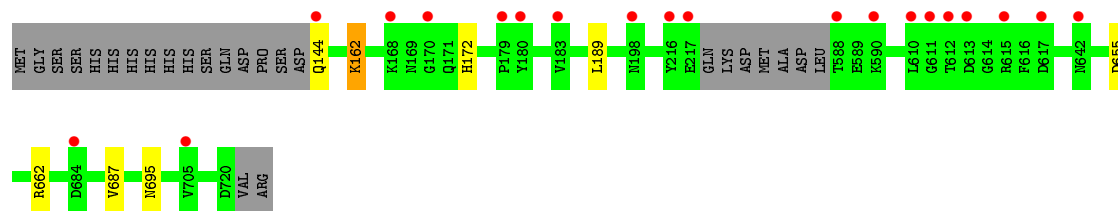


- Molecule 1: Cell wall surface anchor family protein (Jo),Cell wall surface anchor family protein (In)




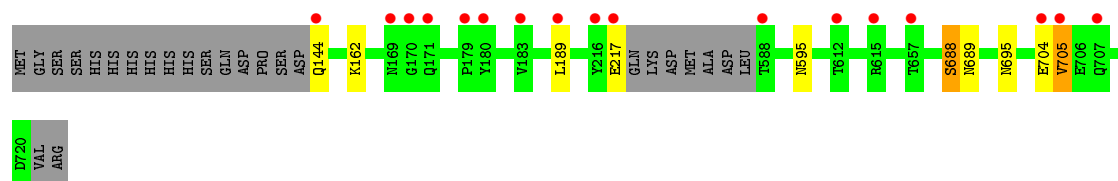
- Molecule 1: Cell wall surface anchor family protein (Jo),Cell wall surface anchor family protein (In)

Chain E: 



- Molecule 1: Cell wall surface anchor family protein (Jo),Cell wall surface anchor family protein (In)

Chain F: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.54Å 134.57Å 144.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.90 – 2.04 45.90 – 2.04	Depositor EDS
% Data completeness (in resolution range)	94.0 (45.90-2.04) 94.0 (45.90-2.04)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.63 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, $R_{free}$	0.213 , 0.243 0.225 , 0.256	Depositor DCC
$R_{free}$ test set	3898 reflections (2.54%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.6	Xtriage
Anisotropy	0.208	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 45.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.012 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10949	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.66% 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: NI, CA, 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	0.52	0/1678	0.69	0/2278
1	B	0.44	0/1664	0.65	0/2258
1	C	0.49	0/1672	0.68	0/2269
1	D	0.55	0/1680	0.71	0/2280
1	E	0.48	0/1673	0.66	0/2270
1	F	0.49	0/1673	0.68	0/2270
All	All	0.50	0/10040	0.68	0/13625

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	1644	0	1607	6	0
1	B	1633	0	1589	6	0
1	C	1638	0	1602	5	0
1	D	1646	0	1606	5	0
1	E	1642	0	1597	4	0
1	F	1642	0	1597	4	0
2	A	15	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	10	0	0	1	0
2	C	10	0	0	1	0
2	D	15	0	0	1	0
2	E	10	0	0	0	0
2	F	15	0	0	0	0
3	A	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
4	C	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
5	A	197	0	0	0	0
5	B	120	0	0	1	0
5	C	191	0	0	0	1
5	D	198	0	0	0	0
5	E	146	0	0	0	0
5	F	171	0	0	0	1
All	All	10949	0	9598	29	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 1.

All (29) 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:162:LYS:NZ	2:D:803:SO4:O4	2.06	0.88
1:C:173:LYS:NZ	1:C:720:ASP:OD1	2.39	0.56
1:B:720:ASP:OD1	1:B:720:ASP:C	2.47	0.53
1:F:595:ASN:HB3	1:F:704:GLU:HG2	1.90	0.52
1:F:704:GLU:HG3	1:F:705:VAL:N	2.24	0.51
1:B:162:LYS:NZ	2:B:802:SO4:O1	2.44	0.50
1:B:162:LYS:HD3	1:B:172:HIS:CG	2.48	0.49
1:A:162:LYS:NZ	2:A:803:SO4:O1	2.46	0.48
1:C:162:LYS:NZ	2:C:802:SO4:O1	2.49	0.45
1:D:688:SER:O	1:D:689:ASN:HB2	2.17	0.45
1:A:630:ARG:HA	1:A:647:LEU:HD12	1.99	0.44
1:D:189:LEU:C	1:D:189:LEU:HD12	2.38	0.44
1:C:688:SER:O	1:C:689:ASN:HB2	2.18	0.44
1:F:688:SER:O	1:F:689:ASN:HB2	2.16	0.43
1:A:189:LEU:HD12	1:A:189:LEU:C	2.38	0.43
1:C:162:LYS:HD3	1:C:172:HIS:CG	2.54	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:189:LEU:HD12	1:B:189:LEU:C	2.39	0.43
1:B:687:VAL:HG23	5:B:906:HOH:O	2.19	0.43
1:E:189:LEU:C	1:E:189:LEU:HD12	2.39	0.43
1:F:189:LEU:HD12	1:F:189:LEU:C	2.39	0.43
1:E:162:LYS:HD3	1:E:172:HIS:CG	2.54	0.42
1:A:705:VAL:HA	1:E:687:VAL:HG11	2.01	0.42
1:E:655:ASP:OD2	1:E:662:ARG:HD3	2.19	0.42
1:C:189:LEU:C	1:C:189:LEU:HD12	2.39	0.42
1:D:162:LYS:HD3	1:D:172:HIS:CG	2.55	0.41
1:B:688:SER:O	1:B:689:ASN:HB2	2.20	0.41
1:A:688:SER:O	1:A:689:ASN:HB2	2.21	0.41
1:A:162:LYS:HD3	1:A:172:HIS:CG	2.56	0.40
1:D:216:TYR:CE1	1:D:589:GLU:HG2	2.57	0.40

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)
5:C:1050:HOH:O	5:F:986:HOH:O[2_555]	1.94	0.26

## 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	205/232 (88%)	203 (99%)	2 (1%)	0	100	100
1	B	203/232 (88%)	200 (98%)	3 (2%)	0	100	100
1	C	204/232 (88%)	202 (99%)	2 (1%)	0	100	100
1	D	205/232 (88%)	204 (100%)	1 (0%)	0	100	100
1	E	204/232 (88%)	201 (98%)	3 (2%)	0	100	100
1	F	204/232 (88%)	201 (98%)	3 (2%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1225/1392 (88%)	1211 (99%)	14 (1%)	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	185/206 (90%)	182 (98%)	3 (2%)	62	59
1	B	183/206 (89%)	179 (98%)	4 (2%)	52	46
1	C	184/206 (89%)	181 (98%)	3 (2%)	62	59
1	D	185/206 (90%)	181 (98%)	4 (2%)	52	46
1	E	184/206 (89%)	181 (98%)	3 (2%)	62	59
1	F	184/206 (89%)	178 (97%)	6 (3%)	38	31
All	All	1105/1236 (89%)	1082 (98%)	23 (2%)	53	48

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	162	LYS
1	A	688	SER
1	A	695	ASN
1	B	162	LYS
1	B	217	GLU
1	B	688	SER
1	B	695	ASN
1	C	162	LYS
1	C	688	SER
1	C	695	ASN
1	D	144	GLN
1	D	162	LYS
1	D	688	SER
1	D	695	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	144	GLN
1	E	162	LYS
1	E	695	ASN
1	F	144	GLN
1	F	162	LYS
1	F	217	GLU
1	F	688	SER
1	F	695	ASN
1	F	705	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	707	GLN
1	B	203	GLN
1	C	707	GLN
1	D	144	GLN
1	E	144	GLN
1	F	144	GLN
1	F	689	ASN

### 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 21 ligands modelled in this entry, 6 are monoatomic - leaving 15 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	$\# Z  > 2$
2	SO4	D	802	-	4,4,4	0.39	0	6,6,6	0.12	0
2	SO4	A	802	-	4,4,4	0.37	0	6,6,6	0.14	0
2	SO4	E	801	-	4,4,4	0.41	0	6,6,6	0.27	0
2	SO4	C	802	-	4,4,4	0.36	0	6,6,6	0.09	0
2	SO4	F	801	3	4,4,4	0.37	0	6,6,6	0.21	0
2	SO4	D	801	3	4,4,4	0.43	0	6,6,6	0.20	0
2	SO4	B	801	-	4,4,4	0.41	0	6,6,6	0.13	0
2	SO4	F	802	-	4,4,4	0.39	0	6,6,6	0.14	0
2	SO4	A	803	-	4,4,4	0.36	0	6,6,6	0.20	0
2	SO4	D	803	-	4,4,4	0.36	0	6,6,6	0.26	0
2	SO4	A	801	3	4,4,4	0.43	0	6,6,6	0.41	0
2	SO4	F	803	-	4,4,4	0.34	0	6,6,6	0.19	0
2	SO4	C	801	-	4,4,4	0.35	0	6,6,6	0.09	0
2	SO4	E	802	-	4,4,4	0.32	0	6,6,6	0.25	0
2	SO4	B	802	-	4,4,4	0.39	0	6,6,6	0.17	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.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	802	SO4	1	0
2	A	803	SO4	1	0
2	D	803	SO4	1	0
2	B	802	SO4	1	0

## 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	207/232 (89%)	0.11	5 (2%) 59 63	31, 41, 74, 132	0
1	B	206/232 (88%)	0.49	17 (8%) 11 11	35, 59, 94, 123	0
1	C	206/232 (88%)	0.02	7 (3%) 45 49	33, 47, 78, 113	0
1	D	207/232 (89%)	0.01	7 (3%) 45 49	32, 42, 74, 119	0
1	E	207/232 (89%)	0.64	20 (9%) 7 8	37, 55, 95, 126	0
1	F	207/232 (89%)	0.39	17 (8%) 11 12	32, 47, 83, 128	0
All	All	1240/1392 (89%)	0.28	73 (5%) 22 24	31, 48, 89, 132	0

All (73) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	721	VAL	11.4
1	E	705	VAL	9.6
1	B	705	VAL	6.9
1	F	144	GLN	5.8
1	F	705	VAL	5.5
1	B	612	THR	5.3
1	B	217	GLU	5.1
1	F	217	GLU	4.6
1	E	612	THR	4.5
1	B	588	THR	4.5
1	F	588	THR	4.3
1	D	217	GLU	4.3
1	B	615	ARG	4.1
1	E	217	GLU	4.0
1	C	588	THR	4.0
1	E	610	LEU	4.0
1	D	144	GLN	3.9
1	A	588	THR	3.7
1	B	169	ASN	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	144	GLN	3.6
1	B	642	ASN	3.4
1	E	216	TYR	3.4
1	A	170	GLY	3.3
1	E	613	ASP	3.3
1	F	171	GLN	3.3
1	B	170	GLY	3.2
1	E	588	THR	3.2
1	B	610	LEU	3.1
1	B	168	LYS	3.1
1	F	615	ARG	3.0
1	E	168	LYS	3.0
1	C	217	GLU	2.9
1	F	180	TYR	2.9
1	D	170	GLY	2.8
1	F	704	GLU	2.8
1	F	216	TYR	2.8
1	E	179	PRO	2.8
1	E	180	TYR	2.7
1	F	183	VAL	2.7
1	F	170	GLY	2.7
1	F	612	THR	2.7
1	F	169	ASN	2.7
1	B	611	GLY	2.7
1	B	704	GLU	2.6
1	E	615	ARG	2.6
1	C	169	ASN	2.6
1	C	612	THR	2.6
1	C	168	LYS	2.6
1	D	169	ASN	2.6
1	A	217	GLU	2.6
1	F	707	GLN	2.6
1	E	590	LYS	2.5
1	E	611	GLY	2.5
1	C	170	GLY	2.5
1	B	633	ASN	2.5
1	C	200	ASP	2.5
1	E	617	ASP	2.4
1	F	189	LEU	2.4
1	B	171	GLN	2.4
1	E	642	ASN	2.3
1	B	684	ASP	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	612	THR	2.3
1	D	183	VAL	2.3
1	F	657	THR	2.2
1	E	183	VAL	2.2
1	F	179	PRO	2.2
1	E	170	GLY	2.2
1	E	198	ASN	2.1
1	E	684	ASP	2.1
1	D	171	GLN	2.1
1	B	657	THR	2.1
1	B	637	VAL	2.0
1	D	180	TYR	2.0

## 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. 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	B	801	5/5	0.82	0.22	88,100,108,119	0
4	CA	E	803	1/1	0.87	0.25	101,101,101,101	0
2	SO4	D	802	5/5	0.90	0.26	69,91,116,132	0
2	SO4	A	802	5/5	0.90	0.29	68,82,127,138	0
2	SO4	E	801	5/5	0.91	0.19	67,82,102,104	0
2	SO4	F	801	5/5	0.91	0.24	76,77,130,139	0
2	SO4	C	801	5/5	0.91	0.37	89,94,124,155	0
2	SO4	B	802	5/5	0.92	0.13	70,95,97,104	0
4	CA	C	803	1/1	0.93	0.20	80,80,80,80	0
2	SO4	F	802	5/5	0.93	0.18	70,84,96,122	0
2	SO4	D	803	5/5	0.94	0.12	79,81,87,88	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	F	803	5/5	0.95	0.10	84,88,93,109	0
3	NI	A	804	1/1	0.95	0.16	113,113,113,113	0
4	CA	F	805	1/1	0.96	0.16	86,86,86,86	0
2	SO4	A	803	5/5	0.96	0.14	72,75,85,103	0
2	SO4	C	802	5/5	0.96	0.13	77,85,91,110	0
2	SO4	E	802	5/5	0.96	0.17	71,90,117,128	0
2	SO4	D	801	5/5	0.96	0.16	44,64,84,106	0
3	NI	D	804	1/1	0.97	0.10	97,97,97,97	0
3	NI	F	804	1/1	0.97	0.10	107,107,107,107	0
2	SO4	A	801	5/5	0.98	0.12	50,53,63,68	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.