



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 12, 2026 – 03:02 PM EST

PDB ID : 4QFU / pdb\_00004qfu  
Title : Crystal structure of a glycoside hydrolase family 5 (BVU\_2644) from *Bacteroides vulgatus* ATCC 8482 at 1.90 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2014-05-21  
Resolution : 1.90 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.48

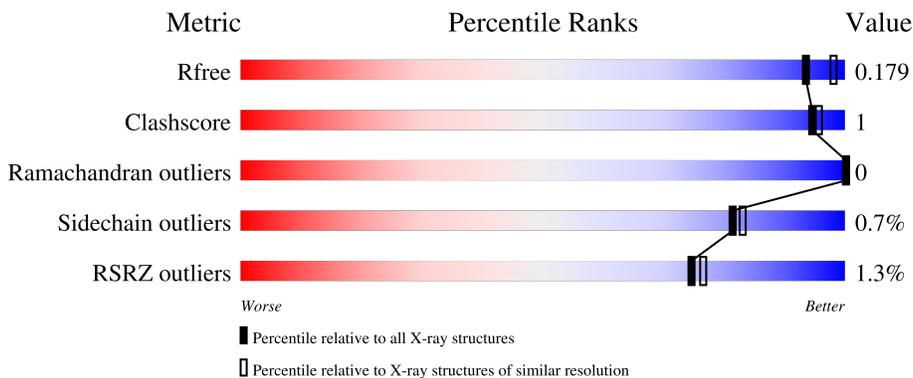
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

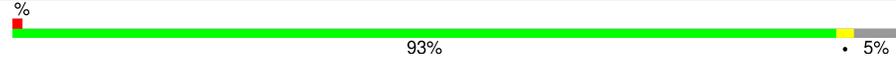
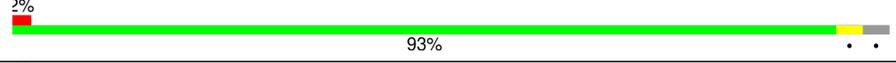
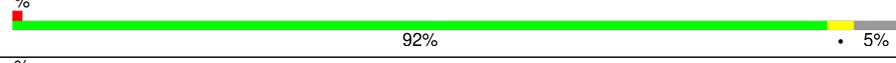
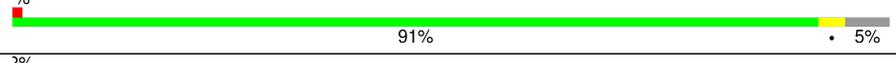
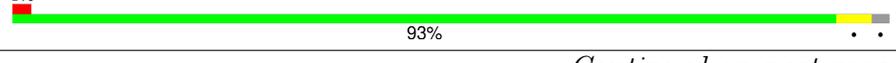
The reported resolution of this entry is 1.90 Å.

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))
$R_{free}$	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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 of 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	484	 93% • 5%
1	B	484	 93% • •
1	C	484	 92% • 5%
1	D	484	 91% • 5%
1	E	484	 93% • •

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	484	
1	G	484	
1	H	484	
1	I	484	
1	J	484	
1	K	484	
1	L	484	

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
5	SAR	C	502	-	-	X	-
5	SAR	G	503	-	-	X	-
5	SAR	K	503	-	-	X	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 51386 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 DUF4038 domain-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	461	3829	2454	652	706	2	15	0	12	0
1	B	468	3852	2460	661	715	2	14	0	7	0
1	C	459	3784	2420	644	704	2	14	0	7	0
1	D	458	3759	2403	644	696	2	14	0	3	0
1	E	472	3878	2479	663	719	2	15	0	9	0
1	F	465	3829	2447	655	710	2	15	0	6	0
1	G	465	3847	2457	666	708	2	14	0	9	0
1	H	461	3804	2434	651	703	2	14	0	8	0
1	I	465	3821	2442	651	712	2	14	0	7	0
1	J	463	3818	2439	655	708	2	14	0	7	0
1	K	463	3800	2431	652	701	2	14	0	5	0
1	L	463	3807	2432	652	707	2	14	0	7	0

There are 228 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	6	MSE	-	expression tag	UNP A6L3N2
A	7	GLY	-	expression tag	UNP A6L3N2
A	8	SER	-	expression tag	UNP A6L3N2
A	9	ASP	-	expression tag	UNP A6L3N2
A	10	LYS	-	expression tag	UNP A6L3N2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	11	ILE	-	expression tag	UNP A6L3N2
A	12	HIS	-	expression tag	UNP A6L3N2
A	13	HIS	-	expression tag	UNP A6L3N2
A	14	HIS	-	expression tag	UNP A6L3N2
A	15	HIS	-	expression tag	UNP A6L3N2
A	16	HIS	-	expression tag	UNP A6L3N2
A	17	HIS	-	expression tag	UNP A6L3N2
A	18	GLU	-	expression tag	UNP A6L3N2
A	19	ASN	-	expression tag	UNP A6L3N2
A	20	LEU	-	expression tag	UNP A6L3N2
A	21	TYR	-	expression tag	UNP A6L3N2
A	22	PHE	-	expression tag	UNP A6L3N2
A	23	GLN	-	expression tag	UNP A6L3N2
A	24	GLY	-	expression tag	UNP A6L3N2
B	-18	MSE	-	expression tag	UNP A6L3N2
B	-17	GLY	-	expression tag	UNP A6L3N2
B	-16	SER	-	expression tag	UNP A6L3N2
B	-15	ASP	-	expression tag	UNP A6L3N2
B	-14	LYS	-	expression tag	UNP A6L3N2
B	-13	ILE	-	expression tag	UNP A6L3N2
B	-12	HIS	-	expression tag	UNP A6L3N2
B	-11	HIS	-	expression tag	UNP A6L3N2
B	-10	HIS	-	expression tag	UNP A6L3N2
B	-9	HIS	-	expression tag	UNP A6L3N2
B	-8	HIS	-	expression tag	UNP A6L3N2
B	-7	HIS	-	expression tag	UNP A6L3N2
B	-6	GLU	-	expression tag	UNP A6L3N2
B	-5	ASN	-	expression tag	UNP A6L3N2
B	-4	LEU	-	expression tag	UNP A6L3N2
B	21	TYR	-	expression tag	UNP A6L3N2
B	22	PHE	-	expression tag	UNP A6L3N2
B	23	GLN	-	expression tag	UNP A6L3N2
B	24	GLY	-	expression tag	UNP A6L3N2
C	6	MSE	-	expression tag	UNP A6L3N2
C	7	GLY	-	expression tag	UNP A6L3N2
C	8	SER	-	expression tag	UNP A6L3N2
C	9	ASP	-	expression tag	UNP A6L3N2
C	10	LYS	-	expression tag	UNP A6L3N2
C	11	ILE	-	expression tag	UNP A6L3N2
C	12	HIS	-	expression tag	UNP A6L3N2
C	13	HIS	-	expression tag	UNP A6L3N2
C	14	HIS	-	expression tag	UNP A6L3N2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	15	HIS	-	expression tag	UNP A6L3N2
C	16	HIS	-	expression tag	UNP A6L3N2
C	17	HIS	-	expression tag	UNP A6L3N2
C	18	GLU	-	expression tag	UNP A6L3N2
C	19	ASN	-	expression tag	UNP A6L3N2
C	20	LEU	-	expression tag	UNP A6L3N2
C	21	TYR	-	expression tag	UNP A6L3N2
C	22	PHE	-	expression tag	UNP A6L3N2
C	23	GLN	-	expression tag	UNP A6L3N2
C	24	GLY	-	expression tag	UNP A6L3N2
D	6	MSE	-	expression tag	UNP A6L3N2
D	7	GLY	-	expression tag	UNP A6L3N2
D	8	SER	-	expression tag	UNP A6L3N2
D	9	ASP	-	expression tag	UNP A6L3N2
D	10	LYS	-	expression tag	UNP A6L3N2
D	11	ILE	-	expression tag	UNP A6L3N2
D	12	HIS	-	expression tag	UNP A6L3N2
D	13	HIS	-	expression tag	UNP A6L3N2
D	14	HIS	-	expression tag	UNP A6L3N2
D	15	HIS	-	expression tag	UNP A6L3N2
D	16	HIS	-	expression tag	UNP A6L3N2
D	17	HIS	-	expression tag	UNP A6L3N2
D	18	GLU	-	expression tag	UNP A6L3N2
D	19	ASN	-	expression tag	UNP A6L3N2
D	20	LEU	-	expression tag	UNP A6L3N2
D	21	TYR	-	expression tag	UNP A6L3N2
D	22	PHE	-	expression tag	UNP A6L3N2
D	23	GLN	-	expression tag	UNP A6L3N2
D	24	GLY	-	expression tag	UNP A6L3N2
E	-18	MSE	-	expression tag	UNP A6L3N2
E	-17	GLY	-	expression tag	UNP A6L3N2
E	-16	SER	-	expression tag	UNP A6L3N2
E	-15	ASP	-	expression tag	UNP A6L3N2
E	-14	LYS	-	expression tag	UNP A6L3N2
E	-13	ILE	-	expression tag	UNP A6L3N2
E	-12	HIS	-	expression tag	UNP A6L3N2
E	-11	HIS	-	expression tag	UNP A6L3N2
E	-10	HIS	-	expression tag	UNP A6L3N2
E	-9	HIS	-	expression tag	UNP A6L3N2
E	-8	HIS	-	expression tag	UNP A6L3N2
E	-7	HIS	-	expression tag	UNP A6L3N2
E	-6	GLU	-	expression tag	UNP A6L3N2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-5	ASN	-	expression tag	UNP A6L3N2
E	-4	LEU	-	expression tag	UNP A6L3N2
E	-3	TYR	-	expression tag	UNP A6L3N2
E	-2	PHE	-	expression tag	UNP A6L3N2
E	23	GLN	-	expression tag	UNP A6L3N2
E	24	GLY	-	expression tag	UNP A6L3N2
F	-18	MSE	-	expression tag	UNP A6L3N2
F	-17	GLY	-	expression tag	UNP A6L3N2
F	-16	SER	-	expression tag	UNP A6L3N2
F	-15	ASP	-	expression tag	UNP A6L3N2
F	-14	LYS	-	expression tag	UNP A6L3N2
F	-13	ILE	-	expression tag	UNP A6L3N2
F	-12	HIS	-	expression tag	UNP A6L3N2
F	-11	HIS	-	expression tag	UNP A6L3N2
F	-10	HIS	-	expression tag	UNP A6L3N2
F	-9	HIS	-	expression tag	UNP A6L3N2
F	-8	HIS	-	expression tag	UNP A6L3N2
F	-7	HIS	-	expression tag	UNP A6L3N2
F	-6	GLU	-	expression tag	UNP A6L3N2
F	-5	ASN	-	expression tag	UNP A6L3N2
F	-4	LEU	-	expression tag	UNP A6L3N2
F	21	TYR	-	expression tag	UNP A6L3N2
F	22	PHE	-	expression tag	UNP A6L3N2
F	23	GLN	-	expression tag	UNP A6L3N2
F	24	GLY	-	expression tag	UNP A6L3N2
G	6	MSE	-	expression tag	UNP A6L3N2
G	7	GLY	-	expression tag	UNP A6L3N2
G	8	SER	-	expression tag	UNP A6L3N2
G	9	ASP	-	expression tag	UNP A6L3N2
G	10	LYS	-	expression tag	UNP A6L3N2
G	11	ILE	-	expression tag	UNP A6L3N2
G	12	HIS	-	expression tag	UNP A6L3N2
G	13	HIS	-	expression tag	UNP A6L3N2
G	14	HIS	-	expression tag	UNP A6L3N2
G	15	HIS	-	expression tag	UNP A6L3N2
G	16	HIS	-	expression tag	UNP A6L3N2
G	17	HIS	-	expression tag	UNP A6L3N2
G	18	GLU	-	expression tag	UNP A6L3N2
G	19	ASN	-	expression tag	UNP A6L3N2
G	20	LEU	-	expression tag	UNP A6L3N2
G	21	TYR	-	expression tag	UNP A6L3N2
G	22	PHE	-	expression tag	UNP A6L3N2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	23	GLN	-	expression tag	UNP A6L3N2
G	24	GLY	-	expression tag	UNP A6L3N2
H	6	MSE	-	expression tag	UNP A6L3N2
H	7	GLY	-	expression tag	UNP A6L3N2
H	8	SER	-	expression tag	UNP A6L3N2
H	9	ASP	-	expression tag	UNP A6L3N2
H	10	LYS	-	expression tag	UNP A6L3N2
H	11	ILE	-	expression tag	UNP A6L3N2
H	12	HIS	-	expression tag	UNP A6L3N2
H	13	HIS	-	expression tag	UNP A6L3N2
H	14	HIS	-	expression tag	UNP A6L3N2
H	15	HIS	-	expression tag	UNP A6L3N2
H	16	HIS	-	expression tag	UNP A6L3N2
H	17	HIS	-	expression tag	UNP A6L3N2
H	18	GLU	-	expression tag	UNP A6L3N2
H	19	ASN	-	expression tag	UNP A6L3N2
H	20	LEU	-	expression tag	UNP A6L3N2
H	21	TYR	-	expression tag	UNP A6L3N2
H	22	PHE	-	expression tag	UNP A6L3N2
H	23	GLN	-	expression tag	UNP A6L3N2
H	24	GLY	-	expression tag	UNP A6L3N2
I	6	MSE	-	expression tag	UNP A6L3N2
I	7	GLY	-	expression tag	UNP A6L3N2
I	8	SER	-	expression tag	UNP A6L3N2
I	9	ASP	-	expression tag	UNP A6L3N2
I	10	LYS	-	expression tag	UNP A6L3N2
I	11	ILE	-	expression tag	UNP A6L3N2
I	12	HIS	-	expression tag	UNP A6L3N2
I	13	HIS	-	expression tag	UNP A6L3N2
I	14	HIS	-	expression tag	UNP A6L3N2
I	15	HIS	-	expression tag	UNP A6L3N2
I	16	HIS	-	expression tag	UNP A6L3N2
I	17	HIS	-	expression tag	UNP A6L3N2
I	18	GLU	-	expression tag	UNP A6L3N2
I	19	ASN	-	expression tag	UNP A6L3N2
I	20	LEU	-	expression tag	UNP A6L3N2
I	21	TYR	-	expression tag	UNP A6L3N2
I	22	PHE	-	expression tag	UNP A6L3N2
I	23	GLN	-	expression tag	UNP A6L3N2
I	24	GLY	-	expression tag	UNP A6L3N2
J	6	MSE	-	expression tag	UNP A6L3N2
J	7	GLY	-	expression tag	UNP A6L3N2

*Continued on next page...*

*Continued from previous page...*

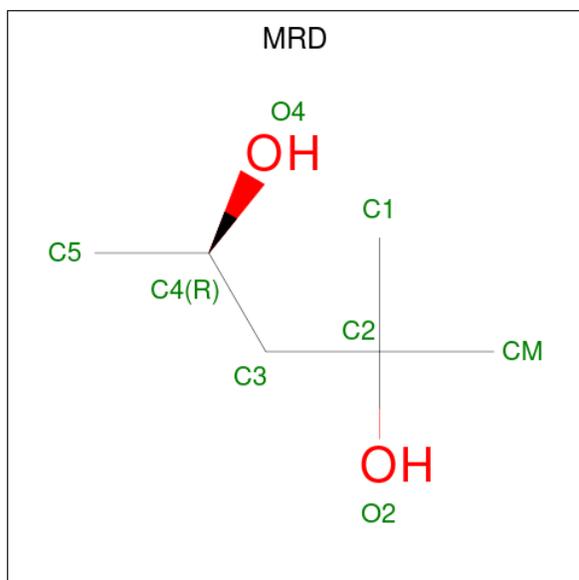
Chain	Residue	Modelled	Actual	Comment	Reference
J	8	SER	-	expression tag	UNP A6L3N2
J	9	ASP	-	expression tag	UNP A6L3N2
J	10	LYS	-	expression tag	UNP A6L3N2
J	11	ILE	-	expression tag	UNP A6L3N2
J	12	HIS	-	expression tag	UNP A6L3N2
J	13	HIS	-	expression tag	UNP A6L3N2
J	14	HIS	-	expression tag	UNP A6L3N2
J	15	HIS	-	expression tag	UNP A6L3N2
J	16	HIS	-	expression tag	UNP A6L3N2
J	17	HIS	-	expression tag	UNP A6L3N2
J	18	GLU	-	expression tag	UNP A6L3N2
J	19	ASN	-	expression tag	UNP A6L3N2
J	20	LEU	-	expression tag	UNP A6L3N2
J	21	TYR	-	expression tag	UNP A6L3N2
J	22	PHE	-	expression tag	UNP A6L3N2
J	23	GLN	-	expression tag	UNP A6L3N2
J	24	GLY	-	expression tag	UNP A6L3N2
K	6	MSE	-	expression tag	UNP A6L3N2
K	7	GLY	-	expression tag	UNP A6L3N2
K	8	SER	-	expression tag	UNP A6L3N2
K	9	ASP	-	expression tag	UNP A6L3N2
K	10	LYS	-	expression tag	UNP A6L3N2
K	11	ILE	-	expression tag	UNP A6L3N2
K	12	HIS	-	expression tag	UNP A6L3N2
K	13	HIS	-	expression tag	UNP A6L3N2
K	14	HIS	-	expression tag	UNP A6L3N2
K	15	HIS	-	expression tag	UNP A6L3N2
K	16	HIS	-	expression tag	UNP A6L3N2
K	17	HIS	-	expression tag	UNP A6L3N2
K	18	GLU	-	expression tag	UNP A6L3N2
K	19	ASN	-	expression tag	UNP A6L3N2
K	20	LEU	-	expression tag	UNP A6L3N2
K	21	TYR	-	expression tag	UNP A6L3N2
K	22	PHE	-	expression tag	UNP A6L3N2
K	23	GLN	-	expression tag	UNP A6L3N2
K	24	GLY	-	expression tag	UNP A6L3N2
L	6	MSE	-	expression tag	UNP A6L3N2
L	7	GLY	-	expression tag	UNP A6L3N2
L	8	SER	-	expression tag	UNP A6L3N2
L	9	ASP	-	expression tag	UNP A6L3N2
L	10	LYS	-	expression tag	UNP A6L3N2
L	11	ILE	-	expression tag	UNP A6L3N2

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	12	HIS	-	expression tag	UNP A6L3N2
L	13	HIS	-	expression tag	UNP A6L3N2
L	14	HIS	-	expression tag	UNP A6L3N2
L	15	HIS	-	expression tag	UNP A6L3N2
L	16	HIS	-	expression tag	UNP A6L3N2
L	17	HIS	-	expression tag	UNP A6L3N2
L	18	GLU	-	expression tag	UNP A6L3N2
L	19	ASN	-	expression tag	UNP A6L3N2
L	20	LEU	-	expression tag	UNP A6L3N2
L	21	TYR	-	expression tag	UNP A6L3N2
L	22	PHE	-	expression tag	UNP A6L3N2
L	23	GLN	-	expression tag	UNP A6L3N2
L	24	GLY	-	expression tag	UNP A6L3N2

- Molecule 2 is (4R)-2-METHYLPENTANE-2,4-DIOL (CCD ID: MRD) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			8	6	2		
2	A	1	Total	C	O	0	0
			8	6	2		
2	B	1	Total	C	O	0	0
			8	6	2		
2	B	1	Total	C	O	0	0
			8	6	2		
2	C	1	Total	C	O	0	0
			8	6	2		

Continued on next page...

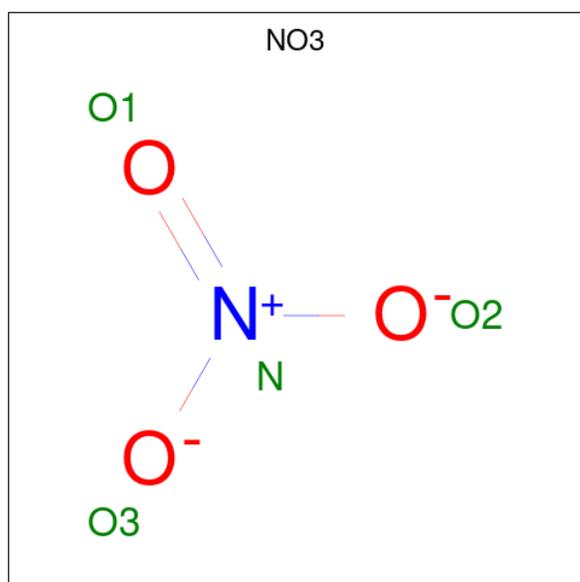
*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total	C	O	0	0
			8	6	2		
2	D	1	Total	C	O	0	0
			8	6	2		
2	E	1	Total	C	O	0	0
			8	6	2		
2	E	1	Total	C	O	0	0
			8	6	2		
2	E	1	Total	C	O	0	0
			8	6	2		
2	F	1	Total	C	O	0	0
			8	6	2		
2	F	1	Total	C	O	0	0
			8	6	2		
2	G	1	Total	C	O	0	0
			8	6	2		
2	G	1	Total	C	O	0	0
			8	6	2		
2	H	1	Total	C	O	0	0
			8	6	2		
2	H	1	Total	C	O	0	0
			8	6	2		
2	I	1	Total	C	O	0	0
			8	6	2		
2	I	1	Total	C	O	0	0
			8	6	2		
2	J	1	Total	C	O	0	0
			8	6	2		
2	J	1	Total	C	O	0	0
			8	6	2		
2	K	1	Total	C	O	0	0
			8	6	2		
2	K	1	Total	C	O	0	0
			8	6	2		
2	L	1	Total	C	O	0	0
			8	6	2		
2	L	1	Total	C	O	0	0
			8	6	2		

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

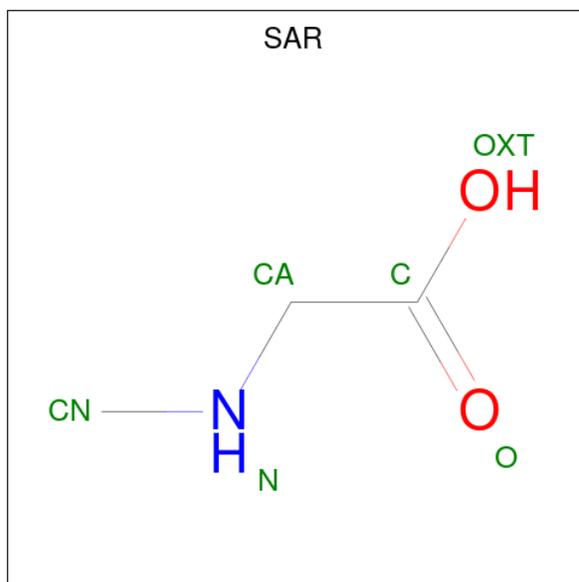
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Cl 1 1	0	0
3	B	1	Total Cl 1 1	0	0
3	C	1	Total Cl 1 1	0	0
3	D	1	Total Cl 1 1	0	0
3	E	1	Total Cl 1 1	0	0
3	F	1	Total Cl 1 1	0	0
3	G	1	Total Cl 1 1	0	0
3	H	1	Total Cl 1 1	0	0
3	I	1	Total Cl 1 1	0	0
3	J	1	Total Cl 1 1	0	0
3	K	1	Total Cl 1 1	0	0
3	L	1	Total Cl 1 1	0	0

- Molecule 4 is NITRATE ION (CCD ID: NO3) (formula: NO<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total N O 4 1 3	0	0
4	E	1	Total N O 4 1 3	0	0
4	F	1	Total N O 4 1 3	0	0

- Molecule 5 is SARCOSINE (CCD ID: SAR) (formula:  $C_3H_7NO_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total C N O 6 3 1 2	0	0
5	G	1	Total C N O 6 3 1 2	0	0
5	K	1	Total C N O 6 3 1 2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	538	Total O 539 539	0	1
6	B	540	Total O 540 540	0	0
6	C	554	Total O 554 554	0	0
6	D	461	Total O 461 461	0	0

Continued on next page...

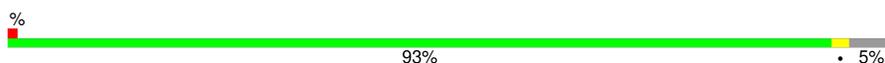
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	E	469	Total 469	O 469	0	0
6	F	410	Total 410	O 410	0	0
6	G	386	Total 386	O 386	0	1
6	H	409	Total 409	O 409	0	0
6	I	344	Total 344	O 344	0	0
6	J	434	Total 434	O 434	0	0
6	K	428	Total 428	O 428	0	0
6	L	350	Total 350	O 350	0	0

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

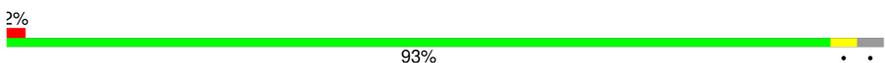
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: DUF4038 domain-containing protein

Chain A: 



- Molecule 1: DUF4038 domain-containing protein

Chain B: 



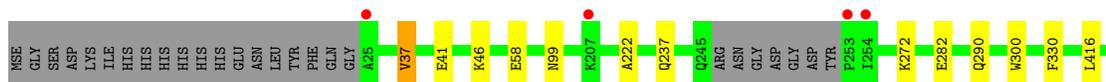
- Molecule 1: DUF4038 domain-containing protein

Chain C: 

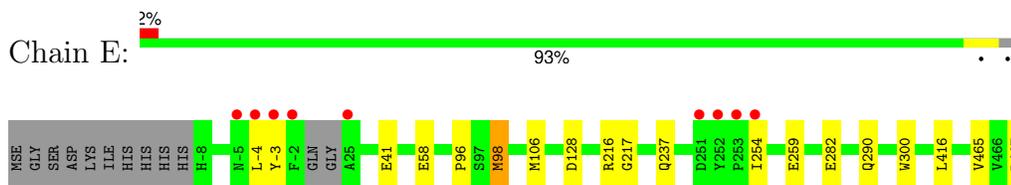


- Molecule 1: DUF4038 domain-containing protein

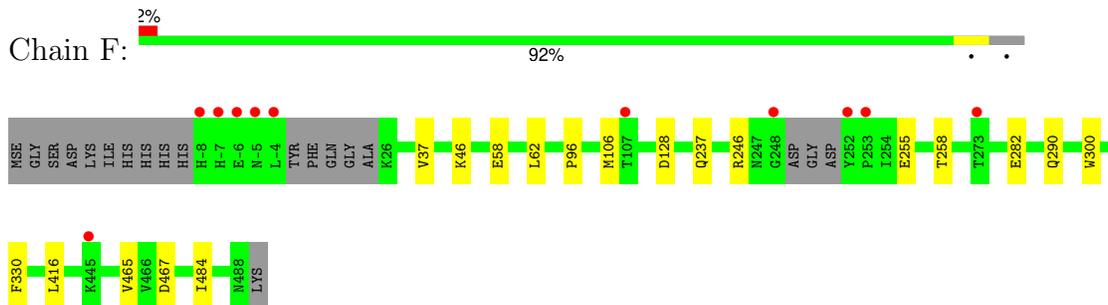
Chain D: 



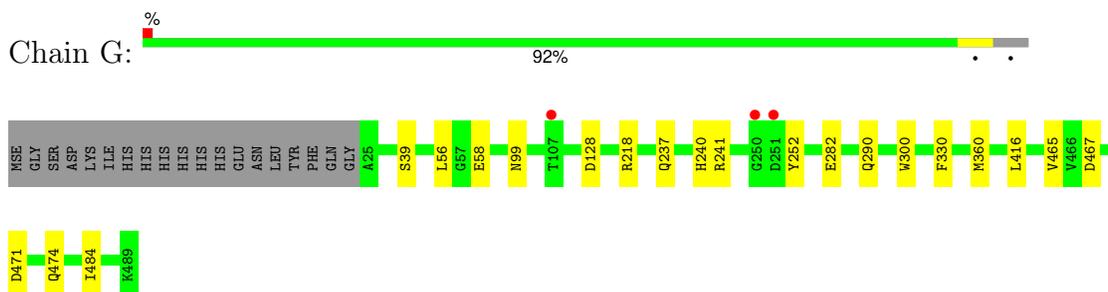
- Molecule 1: DUF4038 domain-containing protein



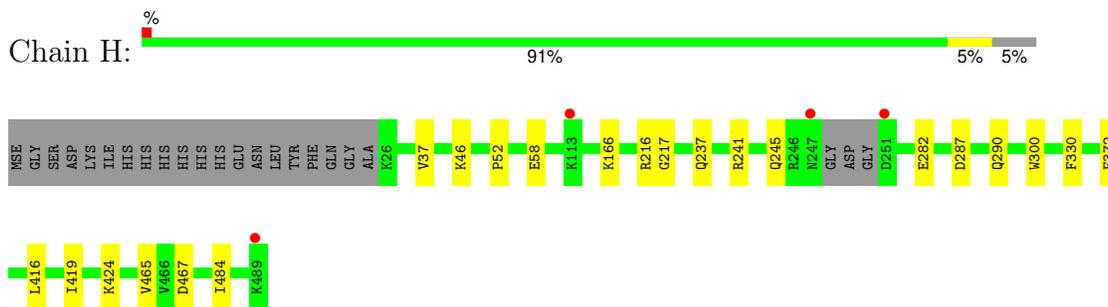
- Molecule 1: DUF4038 domain-containing protein



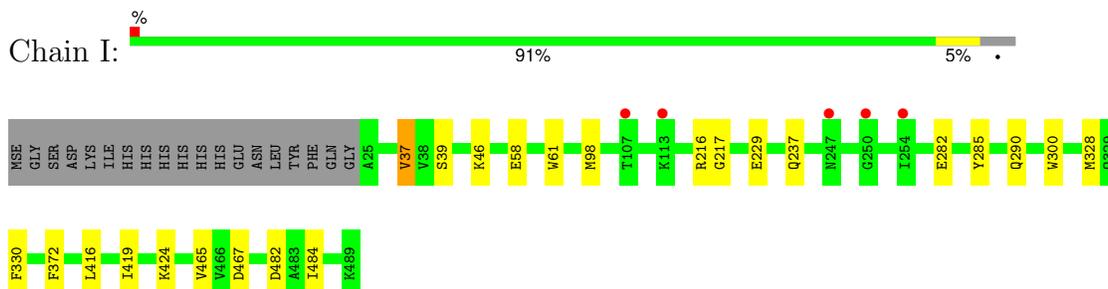
- Molecule 1: DUF4038 domain-containing protein



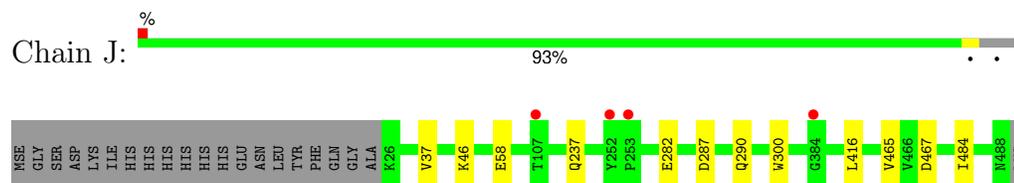
- Molecule 1: DUF4038 domain-containing protein



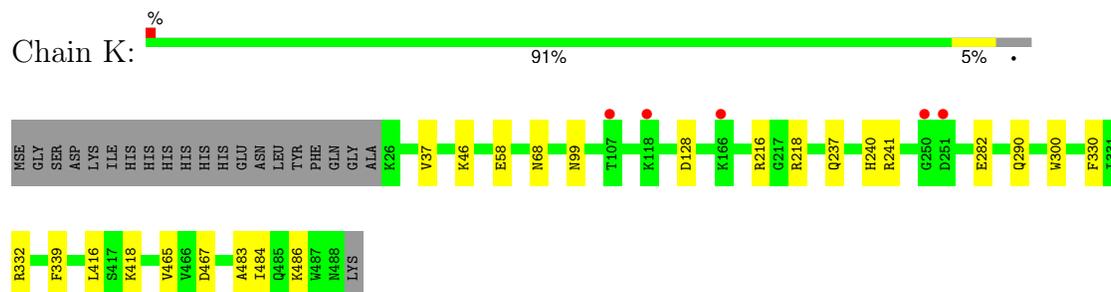
- Molecule 1: DUF4038 domain-containing protein



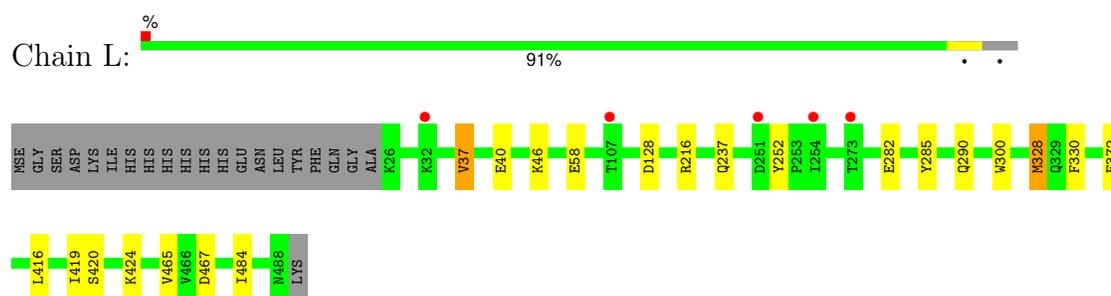
- Molecule 1: DUF4038 domain-containing protein



- Molecule 1: DUF4038 domain-containing protein



- Molecule 1: DUF4038 domain-containing protein



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.95Å 117.59Å 125.25Å 91.60° 92.82° 98.92°	Depositor
Resolution (Å)	46.87 – 1.90 46.87 – 1.90	Depositor EDS
% Data completeness (in resolution range)	89.5 (46.87-1.90) 89.7 (46.87-1.90)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 1.91Å)	Xtrriage
Refinement program	BUSTER-TNT 2.10.0, BUSTER 2.10.0	Depositor
R, $R_{free}$	0.142 , 0.172 0.148 , 0.179	Depositor DCC
$R_{free}$ test set	22984 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.6	Xtrriage
Anisotropy	0.211	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 52.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.035 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	51386	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.89% 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: NO3, MRD, CL, SAR

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	0.86	0/3963	1.02	3/5338 (0.1%)
1	B	0.88	0/3973	1.03	3/5359 (0.1%)
1	C	0.86	0/3903	1.01	2/5266 (0.0%)
1	D	0.82	0/3865	1.04	3/5208 (0.1%)
1	E	0.83	0/4005	1.04	2/5403 (0.0%)
1	F	0.84	0/3946	1.04	4/5321 (0.1%)
1	G	0.82	0/3973	1.04	6/5356 (0.1%)
1	H	0.82	0/3926	1.04	3/5293 (0.1%)
1	I	0.82	0/3941	1.05	5/5315 (0.1%)
1	J	0.80	0/3938	1.05	2/5312 (0.0%)
1	K	0.79	0/3914	1.06	5/5279 (0.1%)
1	L	0.78	0/3927	1.07	5/5299 (0.1%)
All	All	0.83	0/47274	1.04	43/63749 (0.1%)

There are no bond length outliers.

All (43) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	330	PHE	CA-CB-CG	-6.48	107.32	113.80
1	J	467	ASP	CA-CB-CG	6.38	118.98	112.60
1	G	467	ASP	CA-CB-CG	6.14	118.74	112.60
1	K	467	ASP	CA-CB-CG	6.05	118.65	112.60
1	H	467	ASP	CA-CB-CG	6.03	118.63	112.60
1	A	467	ASP	CA-CB-CG	5.95	118.55	112.60
1	L	467	ASP	CA-CB-CG	5.89	118.49	112.60
1	E	467	ASP	CA-CB-CG	5.82	118.42	112.60
1	C	467	ASP	CA-CB-CG	5.80	118.40	112.60
1	K	128	ASP	CA-CB-CG	5.79	118.39	112.60
1	I	467	ASP	CA-CB-CG	5.78	118.38	112.60
1	K	68	ASN	CA-CB-CG	5.73	118.33	112.60

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	330	PHE	CA-CB-CG	-5.66	108.14	113.80
1	H	330	PHE	CA-CB-CG	-5.50	108.30	113.80
1	B	330	PHE	CA-CB-CG	-5.46	108.33	113.80
1	A	330	PHE	CA-CB-CG	-5.45	108.35	113.80
1	D	99	ASN	CA-CB-CG	5.45	118.05	112.60
1	L	128	ASP	CA-CB-CG	5.44	118.04	112.60
1	I	330	PHE	CA-CB-CG	-5.43	108.37	113.80
1	F	467	ASP	CA-CB-CG	5.42	118.02	112.60
1	D	467	ASP	CA-CB-CG	5.41	118.01	112.60
1	C	330	PHE	CA-CB-CG	-5.38	108.42	113.80
1	J	287	ASP	CA-CB-CG	5.37	117.97	112.60
1	B	467	ASP	CA-CB-CG	5.37	117.97	112.60
1	G	39	SER	CA-C-N	5.29	127.63	120.38
1	G	39	SER	C-N-CA	5.29	127.63	120.38
1	B	99	ASN	CA-CB-CG	5.29	117.89	112.60
1	G	99	ASN	CA-CB-CG	5.28	117.88	112.60
1	D	330	PHE	CA-CB-CG	-5.25	108.55	113.80
1	K	99	ASN	CA-CB-CG	5.21	117.81	112.60
1	E	128	ASP	CA-CB-CG	5.21	117.81	112.60
1	A	128	ASP	CA-CB-CG	5.18	117.78	112.60
1	L	372	PHE	N-CA-C	5.16	117.69	111.40
1	G	128	ASP	CA-CB-CG	5.11	117.71	112.60
1	L	330	PHE	CA-CB-CG	-5.11	108.69	113.80
1	I	372	PHE	N-CA-C	5.07	117.59	111.40
1	H	287	ASP	CA-CB-CG	5.05	117.65	112.60
1	L	328	MSE	N-CA-CB	-5.02	102.01	110.49
1	F	128	ASP	CA-CB-CG	5.01	117.61	112.60
1	F	246	ARG	CA-C-N	5.00	129.47	122.36
1	F	246	ARG	C-N-CA	5.00	129.47	122.36
1	I	39	SER	CA-C-N	5.00	128.32	120.82
1	I	39	SER	C-N-CA	5.00	128.32	120.82

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	3829	0	3672	4	0
1	B	3852	0	3644	5	0
1	C	3784	0	3583	13	0
1	D	3759	0	3569	5	0
1	E	3878	0	3674	12	0
1	F	3829	0	3623	9	0
1	G	3847	0	3666	11	0
1	H	3804	0	3616	9	0
1	I	3821	0	3615	10	0
1	J	3818	0	3621	4	0
1	K	3800	0	3607	14	0
1	L	3807	0	3602	9	0
2	A	16	0	28	0	0
2	B	16	0	28	0	0
2	C	8	0	14	0	0
2	D	16	0	28	0	0
2	E	24	0	42	5	0
2	F	16	0	28	1	0
2	G	16	0	28	0	0
2	H	16	0	28	0	0
2	I	16	0	28	0	0
2	J	16	0	28	0	0
2	K	16	0	28	1	0
2	L	16	0	28	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	0	0
3	I	1	0	0	0	0
3	J	1	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
4	B	4	0	0	0	0
4	E	4	0	0	0	0
4	F	4	0	0	0	0
5	C	6	0	6	9	0
5	G	6	0	6	6	0
5	K	6	0	6	7	0
6	A	539	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	540	0	0	1	0
6	C	554	0	0	1	0
6	D	461	0	0	0	0
6	E	469	0	0	0	0
6	F	410	0	0	4	0
6	G	386	0	0	0	0
6	H	409	0	0	1	0
6	I	344	0	0	0	0
6	J	434	0	0	0	0
6	K	428	0	0	0	0
6	L	350	0	0	1	0
All	All	51386	0	43846	106	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 1.

All (106) 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:G:218:ARG:HH12	5:G:503:SAR:HN3	1.24	1.03
1:G:241:ARG:HE	5:G:503:SAR:HA3	1.26	0.98
1:C:241:ARG:HE	5:C:502:SAR:HA3	1.26	0.98
1:E:-3:TYR:HB2	2:E:501:MRD:HMC3	1.49	0.91
1:F:62:LEU:HB2	6:F:989:HOH:O	1.70	0.89
1:K:241:ARG:HE	5:K:503:SAR:HA3	1.37	0.88
1:C:218:ARG:HH12	5:C:502:SAR:HN3	1.37	0.88
1:G:240:HIS:H	5:G:503:SAR:HN1	1.42	0.85
1:K:218:ARG:HH12	5:K:503:SAR:HN3	1.44	0.82
1:F:330:PHE:HD2	6:F:989:HOH:O	1.62	0.81
1:F:330:PHE:CD2	6:F:989:HOH:O	2.38	0.73
1:C:240:HIS:H	5:C:502:SAR:HN1	1.56	0.71
1:E:-3:TYR:HB2	2:E:501:MRD:CM	2.19	0.70
1:E:-3:TYR:CB	2:E:501:MRD:HMC3	2.20	0.69
1:G:218:ARG:NH1	5:G:503:SAR:HN3	2.03	0.67
1:G:240:HIS:N	5:G:503:SAR:HN1	2.14	0.62
1:C:336:GLY:HA3	1:E:-4:LEU:HD11	1.81	0.62
1:E:416:LEU:HD22	1:E:465:VAL:HG11	1.81	0.62
1:C:218:ARG:NH1	5:C:502:SAR:HN3	2.11	0.60
1:K:218:ARG:NH1	5:K:503:SAR:HN3	2.14	0.59
6:B:1111:HOH:O	1:K:418:LYS:HE2	2.03	0.59
1:H:416:LEU:HD22	1:H:465:VAL:HG11	1.85	0.58

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:416:LEU:HD22	1:C:465:VAL:HG11	1.85	0.58
1:J:416:LEU:HD22	1:J:465:VAL:HG11	1.84	0.58
1:D:416:LEU:HD22	1:D:465:VAL:HG11	1.86	0.58
1:L:416:LEU:HD22	1:L:465:VAL:HG11	1.85	0.58
1:F:416:LEU:HD22	1:F:465:VAL:HG11	1.86	0.57
1:I:416:LEU:HD22	1:I:465:VAL:HG11	1.87	0.57
1:A:416:LEU:HD22	1:A:465:VAL:HG11	1.86	0.56
1:G:416:LEU:HD22	1:G:465:VAL:HG11	1.89	0.55
1:C:216:ARG:NH2	5:C:502:SAR:OXT	2.40	0.55
1:B:237:GLN:HG3	1:B:282:GLU:HB3	1.88	0.54
1:C:240:HIS:H	5:C:502:SAR:CN	2.21	0.54
1:C:240:HIS:N	5:C:502:SAR:HN1	2.23	0.53
1:I:482:ASP:OD1	1:I:484:ILE:HG22	2.07	0.53
1:G:240:HIS:H	5:G:503:SAR:CN	2.18	0.52
5:C:502:SAR:HN2	6:C:715:HOH:O	2.10	0.52
1:K:416:LEU:HD22	1:K:465:VAL:HG11	1.91	0.52
1:L:285:TYR:CE2	1:L:328:MSE:HG3	2.45	0.52
1:C:237:GLN:HG3	1:C:282:GLU:HB3	1.92	0.51
1:E:237:GLN:HG3	1:E:282:GLU:HB3	1.92	0.51
1:I:61:TRP:CD1	1:I:328:MSE:HG2	2.45	0.51
1:D:222:ALA:HB1	1:D:272:LYS:HD2	1.94	0.50
1:H:237:GLN:HG3	1:H:282:GLU:HB3	1.92	0.50
2:F:502:MRD:H1C2	6:F:990:HOH:O	2.12	0.49
1:B:416:LEU:HD22	1:B:465:VAL:HG11	1.95	0.48
1:I:419:ILE:HG13	1:I:424:LYS:HE2	1.96	0.48
1:J:237:GLN:HG3	1:J:282:GLU:HB3	1.96	0.48
1:K:240:HIS:H	5:K:503:SAR:HN1	1.78	0.48
1:K:237:GLN:HG3	1:K:282:GLU:HB3	1.96	0.48
1:H:419:ILE:HG13	1:H:424:LYS:HE2	1.95	0.47
1:D:290:GLN:HG2	1:D:300:TRP:CE2	2.50	0.47
1:F:96:PRO:HA	1:F:106[B]:MSE:HE3	1.97	0.47
1:G:237:GLN:HG3	1:G:282:GLU:HB3	1.95	0.47
1:B:290:GLN:HG2	1:B:300:TRP:CE2	2.50	0.46
1:L:420:SER:O	1:L:424:LYS:HE3	2.15	0.46
1:L:237:GLN:HG3	1:L:282:GLU:HB3	1.97	0.46
1:L:37:VAL:HG13	1:L:46:LYS:HG3	1.98	0.46
1:F:290:GLN:HG2	1:F:300:TRP:CE2	2.50	0.46
1:I:285:TYR:CE2	1:I:328:MSE:HG3	2.51	0.46
1:L:290:GLN:HG2	1:L:300:TRP:CE2	2.51	0.46
1:L:419:ILE:HG13	1:L:424:LYS:HE2	1.98	0.46
1:E:96:PRO:HA	1:E:106[B]:MSE:HE3	1.98	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:98:MSE:HE2	1:E:98:MSE:HB2	1.89	0.45
1:G:290:GLN:HG2	1:G:300:TRP:CE2	2.51	0.45
1:L:40[A]:GLU:H	1:L:40[A]:GLU:CD	2.24	0.45
1:B:37:VAL:HG13	1:B:46:LYS:HG3	1.98	0.45
1:K:290:GLN:HG2	1:K:300:TRP:CE2	2.52	0.45
1:I:237:GLN:HG3	1:I:282:GLU:HB3	1.99	0.44
1:K:241:ARG:HE	5:K:503:SAR:CA	2.20	0.44
1:K:241:ARG:NE	5:K:503:SAR:HA3	2.19	0.44
1:I:290:GLN:HG2	1:I:300:TRP:CE2	2.52	0.44
1:E:290:GLN:HG2	1:E:300:TRP:CE2	2.52	0.44
1:F:37[A]:VAL:HG13	1:F:46:LYS:HG3	1.99	0.44
1:G:471:ASP:HB3	1:G:474:GLN:NE2	2.33	0.44
1:I:37:VAL:HG13	1:I:46:LYS:HG3	2.00	0.44
1:L:216:ARG:HD2	6:L:747:HOH:O	2.17	0.44
1:K:218:ARG:HH12	5:K:503:SAR:CN	2.24	0.43
1:D:237:GLN:HG3	1:D:282:GLU:HB3	2.01	0.43
1:A:290:GLN:HG2	1:A:300:TRP:CE2	2.53	0.43
1:C:241:ARG:NE	5:C:502:SAR:HA3	2.11	0.43
1:F:237:GLN:HG3	1:F:282:GLU:HB3	2.00	0.43
1:J:290:GLN:HG2	1:J:300:TRP:CE2	2.53	0.43
1:C:290:GLN:HG2	1:C:300:TRP:CE2	2.54	0.43
1:H:37[A]:VAL:HG13	1:H:46:LYS:HG3	1.99	0.43
1:B:443:ASP:OD1	1:B:445:LYS:HG2	2.18	0.43
1:F:255[A]:GLU:HG3	1:F:258:THR:HG21	2.01	0.43
1:K:339:PHE:CZ	2:K:502:MRD:HMC2	2.54	0.42
1:A:270:GLN:NE2	6:A:835:HOH:O	2.32	0.42
1:H:166[A]:LYS:HE3	6:H:995:HOH:O	2.20	0.42
1:E:-3:TYR:HD1	2:E:501:MRD:H1C3	1.85	0.41
1:I:98:MSE:HE2	1:I:98:MSE:HB2	2.02	0.41
1:K:483:ALA:O	1:K:486:LYS:HG2	2.20	0.41
1:A:140:TYR:HA	1:A:177:ASN:HB2	2.02	0.41
1:H:216:ARG:NH1	1:H:217:GLY:H	2.17	0.41
1:H:290:GLN:HG2	1:H:300:TRP:CE2	2.56	0.41
1:H:52:PRO:HB3	1:H:372:PHE:CD2	2.55	0.41
1:I:216:ARG:NH1	1:I:217:GLY:H	2.19	0.41
1:C:339:PHE:CZ	2:E:501:MRD:HMC2	2.56	0.41
1:K:37:VAL:HG23	1:K:46:LYS:HG3	2.03	0.40
1:H:241:ARG:HD3	1:H:245:GLN:O	2.22	0.40
1:E:216:ARG:HH11	1:E:217:GLY:H	1.70	0.40
1:J:37[A]:VAL:HG13	1:J:46:LYS:HG3	2.03	0.40
1:D:37:VAL:HG13	1:D:46:LYS:HG3	2.04	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:254:ILE:HD11	1:E:259:GLU:HG3	2.04	0.40
1:G:56:LEU:HD11	1:G:360:MSE:HB3	2.04	0.40

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	469/484 (97%)	454 (97%)	15 (3%)	0	100	100
1	B	471/484 (97%)	455 (97%)	16 (3%)	0	100	100
1	C	462/484 (96%)	449 (97%)	13 (3%)	0	100	100
1	D	457/484 (94%)	444 (97%)	13 (3%)	0	100	100
1	E	477/484 (99%)	462 (97%)	15 (3%)	0	100	100
1	F	465/484 (96%)	449 (97%)	16 (3%)	0	100	100
1	G	472/484 (98%)	458 (97%)	14 (3%)	0	100	100
1	H	465/484 (96%)	452 (97%)	13 (3%)	0	100	100
1	I	470/484 (97%)	457 (97%)	13 (3%)	0	100	100
1	J	468/484 (97%)	453 (97%)	15 (3%)	0	100	100
1	K	466/484 (96%)	450 (97%)	16 (3%)	0	100	100
1	L	468/484 (97%)	452 (97%)	16 (3%)	0	100	100
All	All	5610/5808 (97%)	5435 (97%)	175 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	403/396 (102%)	402 (100%)	1 (0%)	92	93
1	B	401/396 (101%)	397 (99%)	4 (1%)	73	74
1	C	396/396 (100%)	394 (100%)	2 (0%)	86	88
1	D	391/396 (99%)	387 (99%)	4 (1%)	73	74
1	E	404/396 (102%)	400 (99%)	4 (1%)	73	74
1	F	400/396 (101%)	398 (100%)	2 (0%)	86	88
1	G	400/396 (101%)	397 (99%)	3 (1%)	79	80
1	H	397/396 (100%)	395 (100%)	2 (0%)	86	88
1	I	397/396 (100%)	393 (99%)	4 (1%)	73	74
1	J	398/396 (100%)	396 (100%)	2 (0%)	86	88
1	K	394/396 (100%)	389 (99%)	5 (1%)	65	65
1	L	396/396 (100%)	392 (99%)	4 (1%)	73	74
All	All	4777/4752 (100%)	4740 (99%)	37 (1%)	81	80

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	58	GLU
1	B	37	VAL
1	B	58	GLU
1	B	252	TYR
1	B	484	ILE
1	C	58	GLU
1	C	484	ILE
1	D	37	VAL
1	D	41	GLU
1	D	58	GLU
1	D	484	ILE
1	E	41	GLU
1	E	58	GLU
1	E	98	MSE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	484	ILE
1	F	58	GLU
1	F	484	ILE
1	G	58	GLU
1	G	252	TYR
1	G	484	ILE
1	H	58	GLU
1	H	484	ILE
1	I	37	VAL
1	I	58	GLU
1	I	229[A]	GLU
1	I	229[B]	GLU
1	J	58	GLU
1	J	484	ILE
1	K	58	GLU
1	K	216	ARG
1	K	332[A]	ARG
1	K	332[B]	ARG
1	K	484	ILE
1	L	37	VAL
1	L	58	GLU
1	L	252	TYR
1	L	484	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (58) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	190	ASN
1	A	234	ASN
1	A	325	ASN
1	B	-8	HIS
1	B	88	GLN
1	B	226	ASN
1	B	234	ASN
1	B	270	GLN
1	B	325	ASN
1	C	88	GLN
1	C	111	ASN
1	C	270	GLN
1	C	325	ASN
1	C	450	GLN
1	D	111	ASN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	D	200	ASN
1	D	234	ASN
1	D	270	GLN
1	D	325	ASN
1	E	-8	HIS
1	E	270	GLN
1	E	325	ASN
1	F	-8	HIS
1	F	88	GLN
1	F	111	ASN
1	F	200	ASN
1	F	270	GLN
1	G	200	ASN
1	G	226	ASN
1	G	234	ASN
1	G	270	GLN
1	G	325	ASN
1	G	474	GLN
1	H	111	ASN
1	H	200	ASN
1	H	226	ASN
1	H	234	ASN
1	H	325	ASN
1	H	399	ASN
1	I	200	ASN
1	I	226	ASN
1	I	234	ASN
1	I	325	ASN
1	I	461	GLN
1	J	111	ASN
1	J	200	ASN
1	J	234	ASN
1	J	270	GLN
1	J	325	ASN
1	K	200	ASN
1	K	226	ASN
1	K	234	ASN
1	K	270	GLN
1	K	325	ASN
1	L	88	GLN
1	L	200	ASN
1	L	234	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	L	325	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 oligosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 42 ligands modelled in this entry, 12 are monoatomic - leaving 30 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	MRD	H	501	-	7,7,7	0.59	0	9,10,10	0.49	0
2	MRD	I	501	-	7,7,7	0.65	0	9,10,10	0.35	0
5	SAR	G	503	-	5,5,5	1.52	1 (20%)	3,5,5	0.95	0
2	MRD	G	501	-	7,7,7	0.77	0	9,10,10	0.32	0
2	MRD	J	501	-	7,7,7	0.62	0	9,10,10	0.30	0
2	MRD	L	501	-	7,7,7	0.69	0	9,10,10	0.38	0
2	MRD	B	502	-	7,7,7	0.51	0	9,10,10	0.40	0
2	MRD	E	502	-	7,7,7	0.63	0	9,10,10	0.43	0
2	MRD	D	502	-	7,7,7	0.69	0	9,10,10	0.35	0
2	MRD	F	502	-	7,7,7	0.59	0	9,10,10	0.70	0
2	MRD	K	502	-	7,7,7	0.62	0	9,10,10	0.39	0
2	MRD	G	502	-	7,7,7	0.67	0	9,10,10	0.32	0
2	MRD	J	502	-	7,7,7	0.58	0	9,10,10	0.31	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	MRD	A	501	-	7,7,7	0.44	0	9,10,10	0.34	0
2	MRD	L	502	-	7,7,7	0.56	0	9,10,10	0.39	0
5	SAR	K	503	-	5,5,5	1.16	0	3,5,5	1.02	0
2	MRD	E	501	-	7,7,7	0.52	0	9,10,10	1.08	1 (11%)
5	SAR	C	502	-	5,5,5	1.14	0	3,5,5	1.10	0
2	MRD	K	501	-	7,7,7	0.52	0	9,10,10	0.45	0
2	MRD	B	501	-	7,7,7	0.62	0	9,10,10	0.42	0
2	MRD	H	502	-	7,7,7	0.60	0	9,10,10	0.39	0
2	MRD	I	502	-	7,7,7	0.54	0	9,10,10	0.43	0
2	MRD	C	501	-	7,7,7	0.71	0	9,10,10	0.41	0
2	MRD	D	501	-	7,7,7	0.45	0	9,10,10	0.43	0
4	NO3	F	503	-	1,3,3	0.30	0	0,3,3	-	-
4	NO3	E	504	-	1,3,3	0.86	0	0,3,3	-	-
4	NO3	B	503	-	1,3,3	0.65	0	0,3,3	-	-
2	MRD	F	501	-	7,7,7	0.49	0	9,10,10	0.49	0
2	MRD	A	502	-	7,7,7	0.53	0	9,10,10	0.37	0
2	MRD	E	503	-	7,7,7	0.42	0	9,10,10	0.35	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	MRD	H	501	-	-	0/5/5/5	-
2	MRD	I	501	-	-	0/5/5/5	-
5	SAR	G	503	-	-	2/2/3/3	-
2	MRD	G	501	-	-	0/5/5/5	-
2	MRD	J	501	-	-	0/5/5/5	-
2	MRD	L	501	-	-	0/5/5/5	-
2	MRD	B	502	-	-	0/5/5/5	-
2	MRD	E	502	-	-	0/5/5/5	-
2	MRD	D	502	-	-	0/5/5/5	-
2	MRD	F	502	-	-	3/5/5/5	-
2	MRD	K	502	-	-	0/5/5/5	-
2	MRD	G	502	-	-	0/5/5/5	-
2	MRD	J	502	-	-	0/5/5/5	-
2	MRD	A	501	-	-	0/5/5/5	-
2	MRD	L	502	-	-	0/5/5/5	-
5	SAR	K	503	-	-	2/2/3/3	-
2	MRD	E	501	-	-	1/5/5/5	-
5	SAR	C	502	-	-	2/2/3/3	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	MRD	K	501	-	-	0/5/5/5	-
2	MRD	B	501	-	-	0/5/5/5	-
2	MRD	H	502	-	-	0/5/5/5	-
2	MRD	I	502	-	-	1/5/5/5	-
2	MRD	C	501	-	-	0/5/5/5	-
2	MRD	D	501	-	-	0/5/5/5	-
2	MRD	F	501	-	-	0/5/5/5	-
2	MRD	A	502	-	-	0/5/5/5	-
2	MRD	E	503	-	-	1/5/5/5	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	503	SAR	CA-C	2.28	1.56	1.50

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	501	MRD	CM-C2-C3	-2.12	101.04	110.20

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	K	503	SAR	OXT-C-CA-N
5	K	503	SAR	O-C-CA-N
5	C	502	SAR	OXT-C-CA-N
5	C	502	SAR	O-C-CA-N
2	F	502	MRD	C1-C2-C3-C4
5	G	503	SAR	OXT-C-CA-N
5	G	503	SAR	O-C-CA-N
2	F	502	MRD	O2-C2-C3-C4
2	E	503	MRD	C2-C3-C4-O4
2	E	501	MRD	C1-C2-C3-C4
2	F	502	MRD	CM-C2-C3-C4
2	I	502	MRD	C1-C2-C3-C4

There are no ring outliers.

6 monomers are involved in 29 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	503	SAR	6	0
2	F	502	MRD	1	0
2	K	502	MRD	1	0
5	K	503	SAR	7	0
2	E	501	MRD	5	0
5	C	502	SAR	9	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	447/484 (92%)	-0.68	7 (1%) 70 72	13, 21, 39, 88	11 (2%)
1	B	454/484 (93%)	-0.66	11 (2%) 59 61	12, 20, 44, 82	7 (1%)
1	C	445/484 (91%)	-0.69	3 (0%) 84 86	11, 21, 42, 77	7 (1%)
1	D	444/484 (91%)	-0.56	4 (0%) 81 82	16, 25, 47, 71	3 (0%)
1	E	458/484 (94%)	-0.47	10 (2%) 62 64	14, 26, 49, 96	8 (1%)
1	F	451/484 (93%)	-0.44	11 (2%) 59 61	15, 25, 50, 84	5 (1%)
1	G	451/484 (93%)	-0.48	3 (0%) 84 86	13, 26, 50, 74	9 (1%)
1	H	447/484 (92%)	-0.51	4 (0%) 81 82	15, 26, 50, 81	8 (1%)
1	I	451/484 (93%)	-0.41	5 (1%) 77 79	14, 30, 53, 92	7 (1%)
1	J	449/484 (92%)	-0.43	4 (0%) 81 82	14, 28, 53, 78	7 (1%)
1	K	449/484 (92%)	-0.41	5 (1%) 77 79	15, 29, 52, 79	5 (1%)
1	L	449/484 (92%)	-0.20	5 (1%) 77 79	18, 35, 60, 82	7 (1%)
All	All	5395/5808 (92%)	-0.50	72 (1%) 74 76	11, 26, 51, 96	84 (1%)

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	-4	LEU	6.1
1	E	-2	PHE	5.9
1	F	-4	LEU	5.9
1	F	248	GLY	4.2
1	F	-5	ASN	4.2
1	K	251	ASP	3.8
1	L	273[A]	THR	3.5
1	D	254	ILE	3.4
1	B	107[A]	THR	3.4
1	C	26	LYS	3.3
1	F	-6	GLU	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	J	384	GLY	3.2
1	B	-6	GLU	3.2
1	A	254	ILE	3.2
1	K	107	THR	3.1
1	A	107[A]	THR	3.1
1	A	253	PRO	3.0
1	B	250	GLY	3.0
1	E	489	LYS	3.0
1	A	252	TYR	2.9
1	J	107	THR	2.9
1	D	25	ALA	2.9
1	E	251	ASP	2.9
1	B	-7	HIS	2.9
1	F	273	THR	2.9
1	F	-7	HIS	2.8
1	D	253	PRO	2.8
1	K	118[A]	LYS	2.8
1	I	247	ASN	2.8
1	K	250	GLY	2.8
1	E	-3	TYR	2.8
1	B	247	ASN	2.8
1	H	247	ASN	2.8
1	B	251	ASP	2.8
1	B	-5	ASN	2.7
1	E	253	PRO	2.7
1	E	25	ALA	2.7
1	G	251	ASP	2.6
1	J	252	TYR	2.6
1	C	253	PRO	2.6
1	E	-4	LEU	2.6
1	F	-8	HIS	2.6
1	F	107	THR	2.5
1	B	-8	HIS	2.5
1	E	254	ILE	2.5
1	L	251	ASP	2.5
1	G	250	GLY	2.4
1	G	107[A]	THR	2.4
1	A	247	ASN	2.4
1	L	107	THR	2.4
1	K	166[A]	LYS	2.4
1	H	251	ASP	2.4
1	F	252	TYR	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	L	32[A]	LYS	2.3
1	H	113	LYS	2.3
1	H	489	LYS	2.3
1	I	113	LYS	2.3
1	F	253	PRO	2.3
1	I	250	GLY	2.3
1	I	107	THR	2.3
1	A	25	ALA	2.3
1	B	246	ARG	2.3
1	L	254	ILE	2.2
1	B	248	GLY	2.2
1	A	207[A]	LYS	2.2
1	J	253	PRO	2.1
1	E	-5	ASN	2.1
1	F	445	LYS	2.1
1	E	252	TYR	2.1
1	D	207	LYS	2.0
1	I	254	ILE	2.0
1	C	247	ASN	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 oligosaccharides 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
5	SAR	K	503	6/6	0.70	0.19	32,43,63,78	0
2	MRD	E	501	8/8	0.77	0.19	35,46,52,52	0
5	SAR	G	503	6/6	0.78	0.17	23,33,61,62	0
2	MRD	H	502	8/8	0.81	0.16	31,44,52,55	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	MRD	J	502	8/8	0.82	0.24	49,57,63,64	0
2	MRD	E	503	8/8	0.84	0.18	46,56,63,64	0
2	MRD	G	502	8/8	0.84	0.17	37,50,58,58	0
2	MRD	L	502	8/8	0.85	0.18	53,58,61,62	0
2	MRD	I	502	8/8	0.85	0.17	42,55,61,62	0
2	MRD	B	502	8/8	0.85	0.17	33,48,56,57	0
2	MRD	K	502	8/8	0.86	0.18	36,51,59,59	0
2	MRD	D	502	8/8	0.88	0.14	36,49,58,58	0
2	MRD	E	502	8/8	0.91	0.11	28,33,34,35	0
5	SAR	C	502	6/6	0.92	0.12	18,31,54,73	0
2	MRD	A	502	8/8	0.92	0.13	34,44,51,53	0
2	MRD	K	501	8/8	0.92	0.11	38,40,42,43	0
2	MRD	F	501	8/8	0.93	0.09	24,30,32,33	0
2	MRD	L	501	8/8	0.94	0.09	35,37,38,38	0
2	MRD	J	501	8/8	0.94	0.09	35,37,38,39	0
2	MRD	F	502	8/8	0.94	0.12	36,47,54,56	0
2	MRD	I	501	8/8	0.94	0.08	29,33,35,35	0
2	MRD	A	501	8/8	0.94	0.07	26,29,31,34	0
2	MRD	G	501	8/8	0.95	0.08	25,29,30,32	0
2	MRD	B	501	8/8	0.95	0.07	23,26,26,26	0
4	NO3	B	503	4/4	0.95	0.11	43,45,46,49	0
4	NO3	E	504	4/4	0.95	0.11	42,45,45,49	0
4	NO3	F	503	4/4	0.95	0.12	38,38,39,40	0
2	MRD	H	501	8/8	0.95	0.09	23,30,32,33	0
2	MRD	C	501	8/8	0.95	0.09	20,25,27,28	0
2	MRD	D	501	8/8	0.95	0.07	29,33,34,36	0
3	CL	E	505	1/1	0.96	0.15	41,41,41,41	0
3	CL	K	504	1/1	0.96	0.08	54,54,54,54	0
3	CL	I	503	1/1	0.97	0.06	48,48,48,48	0
3	CL	J	503	1/1	0.97	0.11	52,52,52,52	0
3	CL	H	503	1/1	0.97	0.07	52,52,52,52	0
3	CL	D	503	1/1	0.98	0.06	38,38,38,38	0
3	CL	A	503	1/1	0.98	0.06	37,37,37,37	0
3	CL	G	504	1/1	0.98	0.05	41,41,41,41	0
3	CL	L	503	1/1	0.98	0.05	48,48,48,48	0
3	CL	C	503	1/1	0.98	0.06	42,42,42,42	0
3	CL	B	504	1/1	0.99	0.06	37,37,37,37	0
3	CL	F	504	1/1	0.99	0.07	39,39,39,39	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.