



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2017 – 07:49 AM EDT

PDB ID : 3B2U  
Title : Crystal structure of isolated domain III of the extracellular region of the epidermal growth factor receptor in complex with the Fab fragment of IMC-11F8  
Authors : Ferguson, K.M.; Li, S.; Kussie, P.  
Deposited on : unknown  
Resolution : 2.58 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

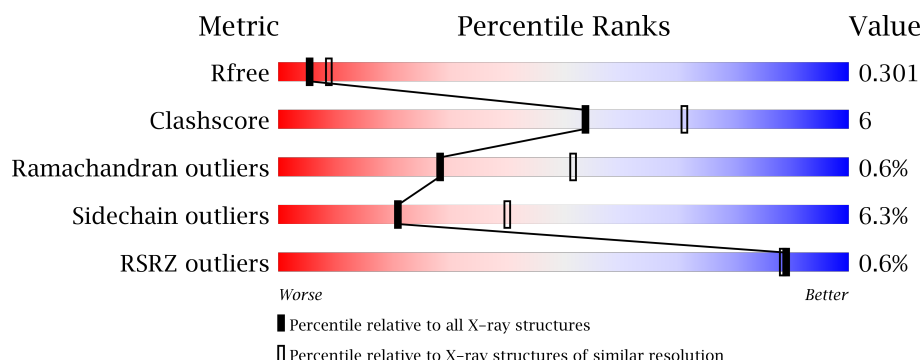
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.58 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.






















Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	2899 (2.60-2.56)
Clashscore	112137	3268 (2.60-2.56)
Ramachandran outliers	110173	3218 (2.60-2.56)
Sidechain outliers	110143	3218 (2.60-2.56)
RSRZ outliers	101464	2907 (2.60-2.56)

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	D	213	
1	G	213	
1	K	213	
1	L	213	
1	O	213	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	R	213	
1	U	213	
1	X	213	
2	C	223	
2	F	223	
2	H	223	
2	J	223	
2	N	223	
2	Q	223	
2	T	223	
2	W	223	
3	A	214	
3	B	214	
3	E	214	
3	I	214	
3	M	214	
3	P	214	
3	S	214	
3	V	214	

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
6	MAN	I	3284	X	-	-	-

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 39035 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 IMC-11F8 Fab Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	213	Total	C	N	O	S	0	0	0
			1614	1008	272	329	5			
1	D	213	Total	C	N	O	S	0	0	0
			1605	1003	270	327	5			
1	G	213	Total	C	N	O	S	0	0	0
			1601	1002	269	325	5			
1	K	213	Total	C	N	O	S	0	0	0
			1598	999	265	329	5			
1	O	213	Total	C	N	O	S	0	0	0
			1606	1003	271	327	5			
1	R	213	Total	C	N	O	S	0	0	0
			1610	1005	271	329	5			
1	U	213	Total	C	N	O	S	0	0	0
			1618	1012	274	327	5			
1	X	213	Total	C	N	O	S	0	0	0
			1612	1008	270	329	5			

- Molecule 2 is a protein called IMC-11F8 Fab Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	217	Total	C	N	O	S	0	0	0
			1603	1020	258	320	5			
2	C	213	Total	C	N	O	S	0	0	0
			1563	997	250	311	5			
2	F	215	Total	C	N	O	S	0	0	0
			1579	1004	254	316	5			
2	J	217	Total	C	N	O	S	0	0	0
			1591	1013	257	316	5			
2	N	218	Total	C	N	O	S	0	0	0
			1603	1018	258	322	5			
2	Q	217	Total	C	N	O	S	0	0	0
			1594	1015	256	318	5			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	216	Total	C	N	O	S	0	0	0
			1586	1011	253	317	5			
2	W	216	Total	C	N	O	S	0	0	0
			1595	1015	256	319	5			

- Molecule 3 is a protein called Epidermal growth factor receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	195	Total	C	N	O	S	0	0	0
			1452	909	255	279	9			
3	B	194	Total	C	N	O	S	0	0	0
			1424	893	249	273	9			
3	E	194	Total	C	N	O	S	0	0	0
			1441	903	252	277	9			
3	I	193	Total	C	N	O	S	0	0	0
			1454	912	254	279	9			
3	M	195	Total	C	N	O	S	0	0	0
			1444	905	253	277	9			
3	P	193	Total	C	N	O	S	0	0	0
			1441	903	253	276	9			
3	S	194	Total	C	N	O	S	0	0	0
			1427	889	254	275	9			
3	V	193	Total	C	N	O	S	0	0	0
			1422	894	243	276	9			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	307	LEU	-	EXPRESSION TAG	UNP P00533
A	308	GLU	-	EXPRESSION TAG	UNP P00533
A	309	GLU	-	EXPRESSION TAG	UNP P00533
A	310	LYS	-	EXPRESSION TAG	UNP P00533
A	515	HIS	-	EXPRESSION TAG	UNP P00533
A	516	HIS	-	EXPRESSION TAG	UNP P00533
A	517	HIS	-	EXPRESSION TAG	UNP P00533
A	518	HIS	-	EXPRESSION TAG	UNP P00533
A	519	HIS	-	EXPRESSION TAG	UNP P00533
A	520	HIS	-	EXPRESSION TAG	UNP P00533
B	307	LEU	-	EXPRESSION TAG	UNP P00533
B	308	GLU	-	EXPRESSION TAG	UNP P00533
B	309	GLU	-	EXPRESSION TAG	UNP P00533
B	310	LYS	-	EXPRESSION TAG	UNP P00533

*Continued on next page...*

*Continued from previous page...*

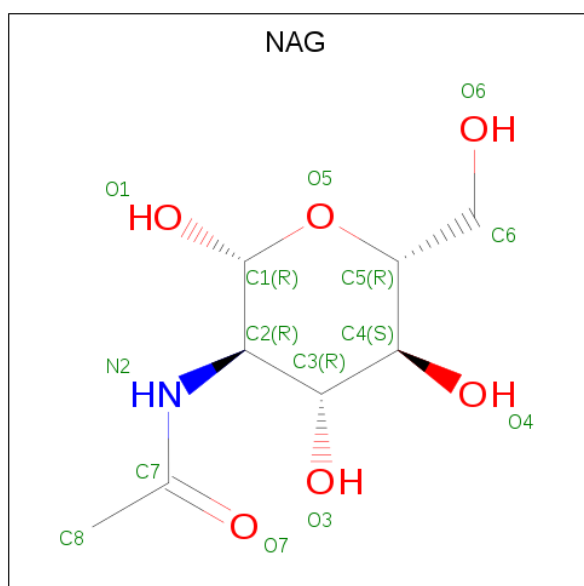
Chain	Residue	Modelled	Actual	Comment	Reference
B	515	HIS	-	EXPRESSION TAG	UNP P00533
B	516	HIS	-	EXPRESSION TAG	UNP P00533
B	517	HIS	-	EXPRESSION TAG	UNP P00533
B	518	HIS	-	EXPRESSION TAG	UNP P00533
B	519	HIS	-	EXPRESSION TAG	UNP P00533
B	520	HIS	-	EXPRESSION TAG	UNP P00533
E	307	LEU	-	EXPRESSION TAG	UNP P00533
E	308	GLU	-	EXPRESSION TAG	UNP P00533
E	309	GLU	-	EXPRESSION TAG	UNP P00533
E	310	LYS	-	EXPRESSION TAG	UNP P00533
E	515	HIS	-	EXPRESSION TAG	UNP P00533
E	516	HIS	-	EXPRESSION TAG	UNP P00533
E	517	HIS	-	EXPRESSION TAG	UNP P00533
E	518	HIS	-	EXPRESSION TAG	UNP P00533
E	519	HIS	-	EXPRESSION TAG	UNP P00533
E	520	HIS	-	EXPRESSION TAG	UNP P00533
I	307	LEU	-	EXPRESSION TAG	UNP P00533
I	308	GLU	-	EXPRESSION TAG	UNP P00533
I	309	GLU	-	EXPRESSION TAG	UNP P00533
I	310	LYS	-	EXPRESSION TAG	UNP P00533
I	515	HIS	-	EXPRESSION TAG	UNP P00533
I	516	HIS	-	EXPRESSION TAG	UNP P00533
I	517	HIS	-	EXPRESSION TAG	UNP P00533
I	518	HIS	-	EXPRESSION TAG	UNP P00533
I	519	HIS	-	EXPRESSION TAG	UNP P00533
I	520	HIS	-	EXPRESSION TAG	UNP P00533
M	307	LEU	-	EXPRESSION TAG	UNP P00533
M	308	GLU	-	EXPRESSION TAG	UNP P00533
M	309	GLU	-	EXPRESSION TAG	UNP P00533
M	310	LYS	-	EXPRESSION TAG	UNP P00533
M	515	HIS	-	EXPRESSION TAG	UNP P00533
M	516	HIS	-	EXPRESSION TAG	UNP P00533
M	517	HIS	-	EXPRESSION TAG	UNP P00533
M	518	HIS	-	EXPRESSION TAG	UNP P00533
M	519	HIS	-	EXPRESSION TAG	UNP P00533
M	520	HIS	-	EXPRESSION TAG	UNP P00533
P	307	LEU	-	EXPRESSION TAG	UNP P00533
P	308	GLU	-	EXPRESSION TAG	UNP P00533
P	309	GLU	-	EXPRESSION TAG	UNP P00533
P	310	LYS	-	EXPRESSION TAG	UNP P00533
P	515	HIS	-	EXPRESSION TAG	UNP P00533
P	516	HIS	-	EXPRESSION TAG	UNP P00533

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	517	HIS	-	EXPRESSION TAG	UNP P00533
P	518	HIS	-	EXPRESSION TAG	UNP P00533
P	519	HIS	-	EXPRESSION TAG	UNP P00533
P	520	HIS	-	EXPRESSION TAG	UNP P00533
S	307	LEU	-	EXPRESSION TAG	UNP P00533
S	308	GLU	-	EXPRESSION TAG	UNP P00533
S	309	GLU	-	EXPRESSION TAG	UNP P00533
S	310	LYS	-	EXPRESSION TAG	UNP P00533
S	515	HIS	-	EXPRESSION TAG	UNP P00533
S	516	HIS	-	EXPRESSION TAG	UNP P00533
S	517	HIS	-	EXPRESSION TAG	UNP P00533
S	518	HIS	-	EXPRESSION TAG	UNP P00533
S	519	HIS	-	EXPRESSION TAG	UNP P00533
S	520	HIS	-	EXPRESSION TAG	UNP P00533
V	307	LEU	-	EXPRESSION TAG	UNP P00533
V	308	GLU	-	EXPRESSION TAG	UNP P00533
V	309	GLU	-	EXPRESSION TAG	UNP P00533
V	310	LYS	-	EXPRESSION TAG	UNP P00533
V	515	HIS	-	EXPRESSION TAG	UNP P00533
V	516	HIS	-	EXPRESSION TAG	UNP P00533
V	517	HIS	-	EXPRESSION TAG	UNP P00533
V	518	HIS	-	EXPRESSION TAG	UNP P00533
V	519	HIS	-	EXPRESSION TAG	UNP P00533
V	520	HIS	-	EXPRESSION TAG	UNP P00533

- Molecule 4 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	E	1	Total	C	N	O	0	0
			14	8	1	5		
4	I	1	Total	C	N	O	0	0
			14	8	1	5		
4	I	1	Total	C	N	O	0	0
			14	8	1	5		
4	I	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...



*Continued from previous page...*

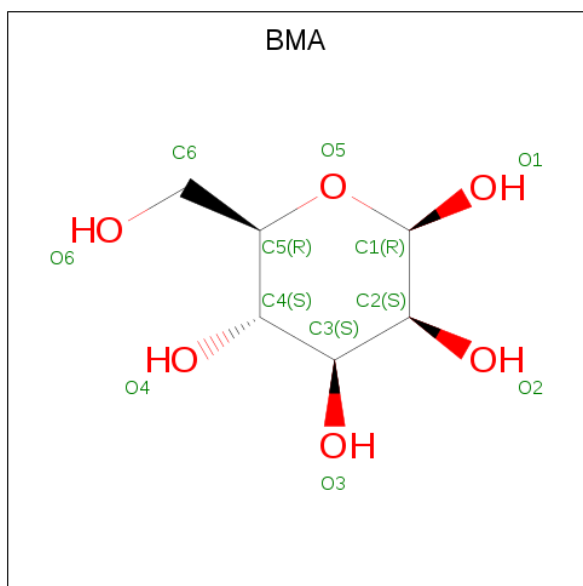
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	I	1	Total	C	N	O	0	0
			14	8	1	5		
4	I	1	Total	C	N	O	0	0
			14	8	1	5		
4	I	1	Total	C	N	O	0	0
			14	8	1	5		
4	M	1	Total	C	N	O	0	0
			14	8	1	5		
4	M	1	Total	C	N	O	0	0
			14	8	1	5		
4	M	1	Total	C	N	O	0	0
			14	8	1	5		
4	M	1	Total	C	N	O	0	0
			14	8	1	5		
4	M	1	Total	C	N	O	0	0
			14	8	1	5		
4	M	1	Total	C	N	O	0	0
			14	8	1	5		
4	P	1	Total	C	N	O	0	0
			14	8	1	5		
4	P	1	Total	C	N	O	0	0
			14	8	1	5		
4	P	1	Total	C	N	O	0	0
			14	8	1	5		
4	P	1	Total	C	N	O	0	0
			14	8	1	5		
4	P	1	Total	C	N	O	0	0
			14	8	1	5		
4	S	1	Total	C	N	O	0	0
			14	8	1	5		
4	S	1	Total	C	N	O	0	0
			14	8	1	5		
4	S	1	Total	C	N	O	0	0
			14	8	1	5		
4	S	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	S	1	Total	C	N	O	0	0
			14	8	1	5		
4	V	1	Total	C	N	O	0	0
			14	8	1	5		
4	V	1	Total	C	N	O	0	0
			14	8	1	5		
4	V	1	Total	C	N	O	0	0
			14	8	1	5		
4	V	1	Total	C	N	O	0	0
			14	8	1	5		
4	V	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is BETA-D-MANNOSE (three-letter code: BMA) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



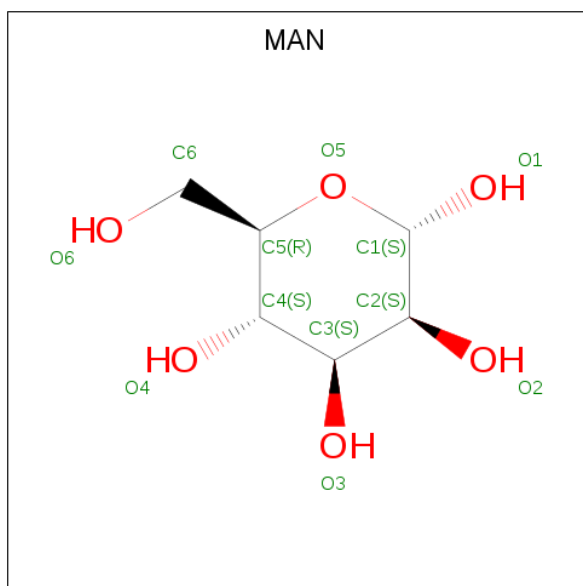
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			11	6	5		
5	B	1	Total	C	O	0	0
			11	6	5		
5	E	1	Total	C	O	0	0
			11	6	5		
5	I	1	Total	C	O	0	0
			11	6	5		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	I	1	Total	C	O	0	0
			11	6	5		
5	M	1	Total	C	O	0	0
			11	6	5		
5	M	1	Total	C	O	0	0
			11	6	5		
5	M	1	Total	C	O	0	0
			11	6	5		
5	P	1	Total	C	O	0	0
			11	6	5		
5	S	1	Total	C	O	0	0
			11	6	5		
5	V	1	Total	C	O	0	0
			11	6	5		

- Molecule 6 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			11	6	5		
6	E	1	Total	C	O	0	0
			11	6	5		
6	I	1	Total	C	O	0	0
			11	6	5		
6	M	1	Total	C	O	0	0
			11	6	5		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	M	1	Total	C	O	0	0
			11	6	5		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	O	S	0	0
			5	4	1		
7	G	1	Total	O	S	0	0
			5	4	1		
7	E	1	Total	O	S	0	0
			5	4	1		
7	K	1	Total	O	S	0	0
			5	4	1		
7	O	1	Total	O	S	0	0
			5	4	1		
7	U	1	Total	O	S	0	0
			5	4	1		
7	X	1	Total	O	S	0	0
			5	4	1		
7	V	1	Total	O	S	0	0
			5	4	1		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	L	56	Total O 56 56	0	0
8	H	47	Total O 47 47	0	0
8	A	30	Total O 30 30	0	0
8	D	55	Total O 55 55	0	0
8	C	45	Total O 45 45	0	0
8	B	22	Total O 22 22	0	0
8	G	50	Total O 50 50	0	0
8	F	40	Total O 40 40	0	0
8	E	30	Total O 30 30	0	0
8	K	63	Total O 63 63	0	0
8	J	44	Total O 44 44	0	0
8	I	41	Total O 41 41	0	0
8	O	57	Total O 57 57	0	0
8	N	40	Total O 40 40	0	0
8	M	35	Total O 35 35	0	0
8	R	62	Total O 62 62	0	0
8	Q	51	Total O 51 51	0	0
8	P	41	Total O 41 41	0	0
8	U	42	Total O 42 42	0	0
8	T	41	Total O 41 41	0	0
8	S	23	Total O 23 23	0	0
8	X	47	Total O 47 47	0	0

*Continued on next page...*

*Continued from previous page...*

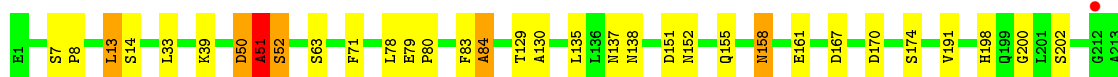
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	W	50	Total	O	0	0
			50	50		
8	V	24	Total	O	0	0
			24	24		

### 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.

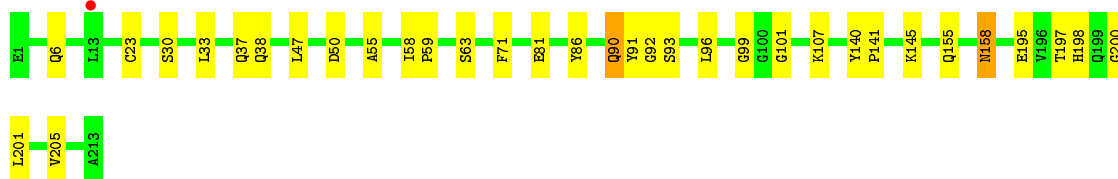
- Molecule 1: IMC-11F8 Fab Light chain

Chain L:  85% 13%




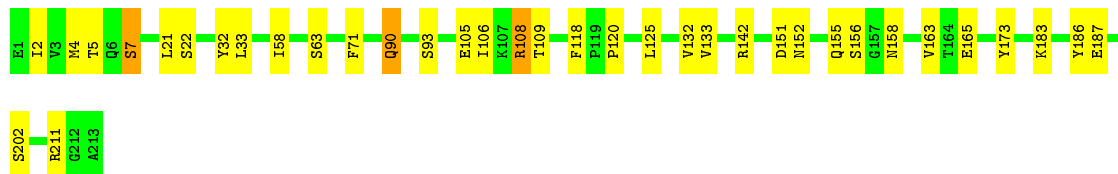
- Molecule 1: IMC-11F8 Fab Light chain

Chain D:  84% 15%




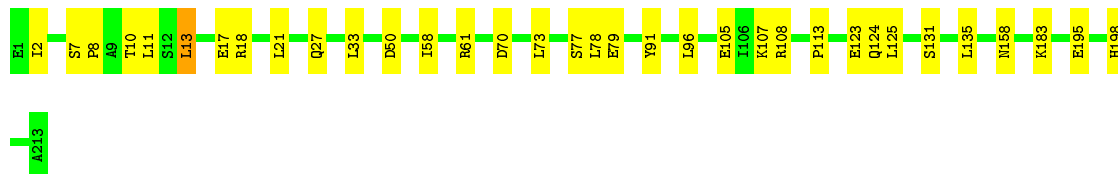
- Molecule 1: IMC-11F8 Fab Light chain

Chain G:  83% 15%




- Molecule 1: IMC-11F8 Fab Light chain

Chain K:  84% 15%




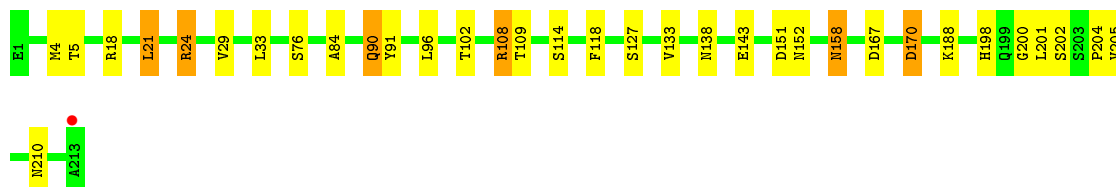
- Molecule 1: IMC-11F8 Fab Light chain

Chain O:  91% 8% .




• Molecule 1: IMC-11F8 Fab Light chain

Chain R:  84% 13% .




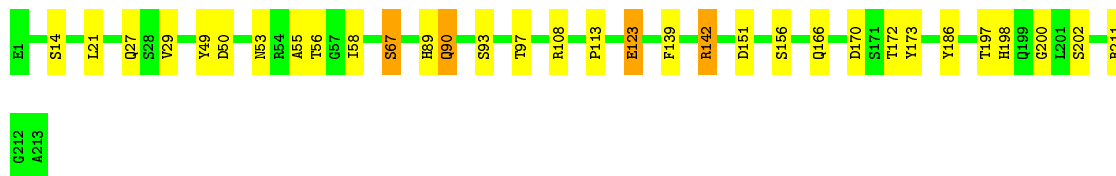
• Molecule 1: IMC-11F8 Fab Light chain

Chain U:  88% 10% .




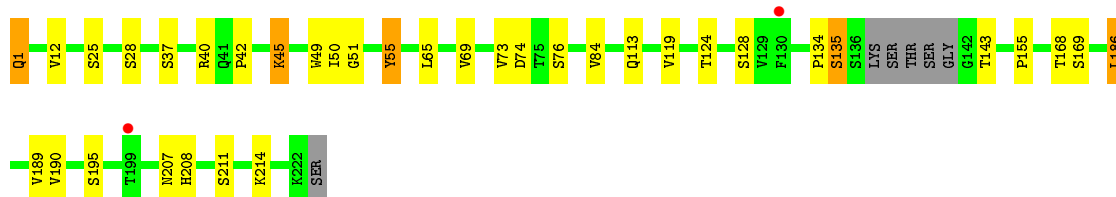
• Molecule 1: IMC-11F8 Fab Light chain

Chain X:  85% 13% .



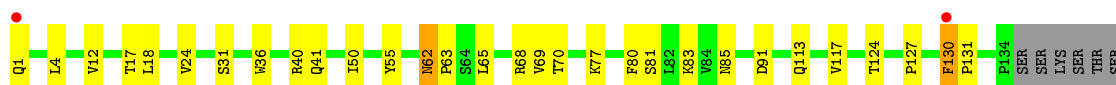
• Molecule 2: IMC-11F8 Fab Heavy chain

Chain H:  81% 14% . .



• Molecule 2: IMC-11F8 Fab Heavy chain

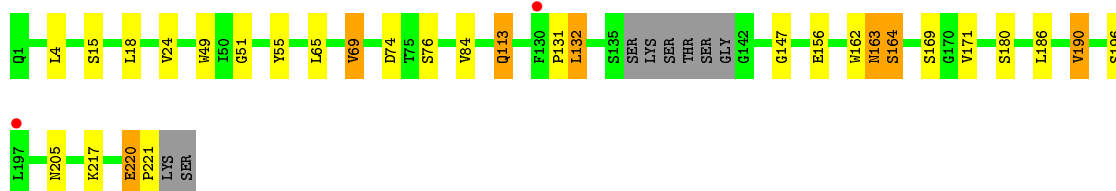
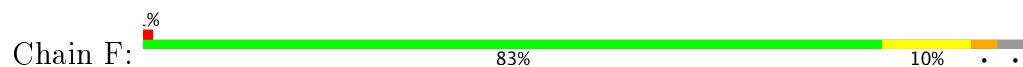
Chain C:  72% 21% . .



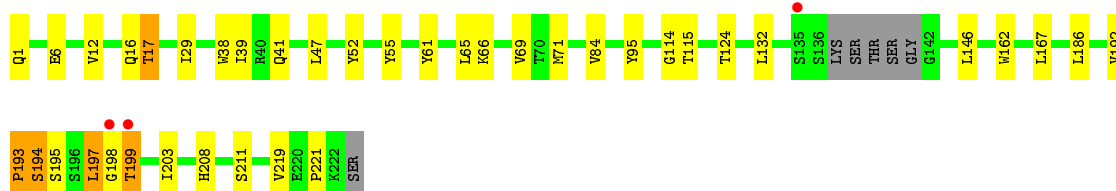
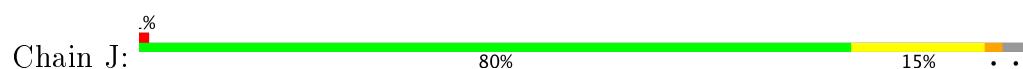




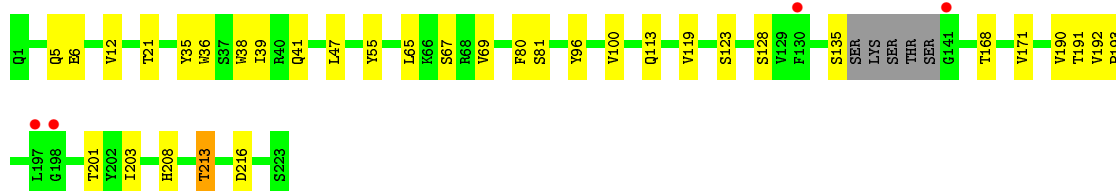
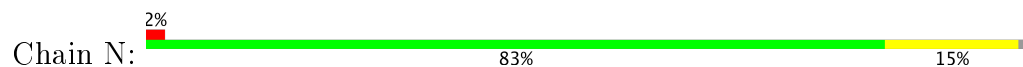
- Molecule 2: IMC-11F8 Fab Heavy chain



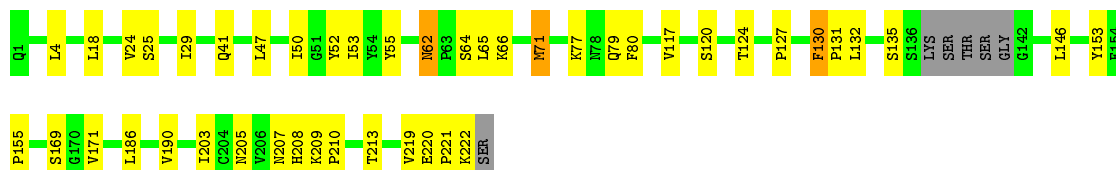
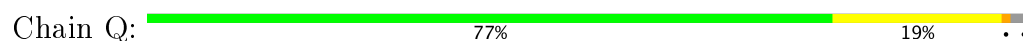
- Molecule 2: IMC-11F8 Fab Heavy chain



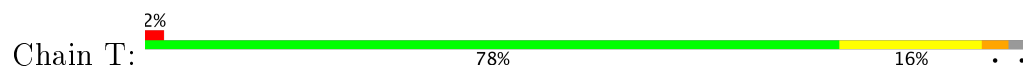
- Molecule 2: IMC-11F8 Fab Heavy chain



- Molecule 2: IMC-11F8 Fab Heavy chain



- Molecule 2: IMC-11F8 Fab Heavy chain

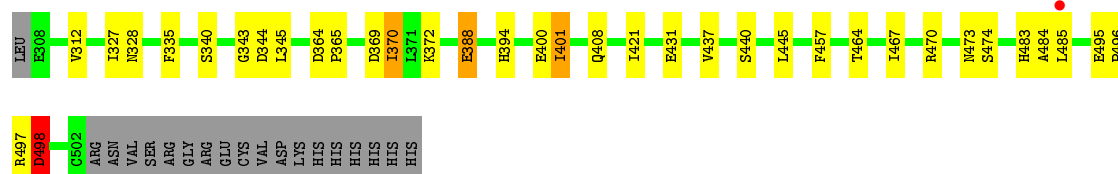




ARG  
ASN  
VAL  
SER  
ARG  
GLY  
ARG  
GLU  
CYS  
VAL  
ASP  
LYS  
HIS  
HIS  
HIS  
HIS  
HIS

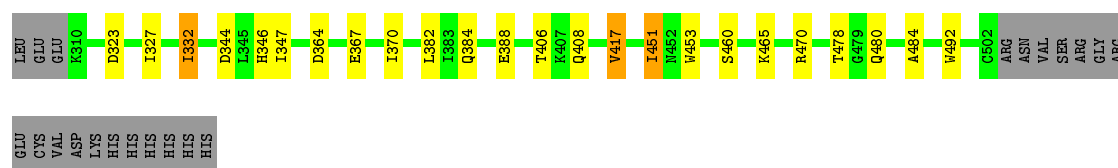
• Molecule 3: Epidermal growth factor receptor

Chain M:  74% 15% 9%



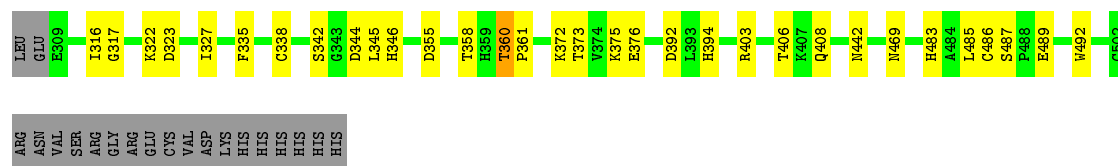
• Molecule 3: Epidermal growth factor receptor

Chain P:  79% 10% 10%




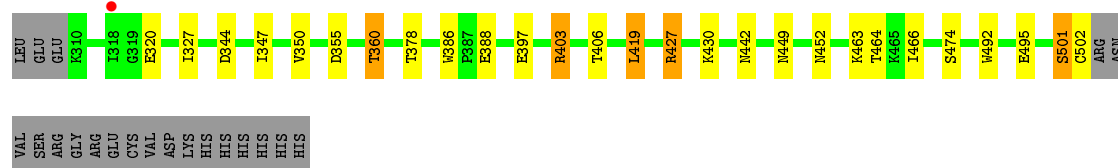
• Molecule 3: Epidermal growth factor receptor

Chain S:  76% 14% 9%



• Molecule 3: Epidermal growth factor receptor

Chain V:  78% 10% 10%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	154.37Å 139.12Å 175.32Å 90.00° 90.02° 90.00°	Depositor
Resolution (Å)	43.07 – 2.58 43.05 – 2.58	Depositor EDS
% Data completeness (in resolution range)	98.9 (43.07-2.58) 98.9 (43.05-2.58)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 2.58Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.227 , 0.291 0.241 , 0.301	Depositor DCC
$R_{free}$ test set	11652 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	37.0	Xtriage
Anisotropy	0.085	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 8.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.345 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	39035	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.88% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, BMA, NAG, MAN

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	D	0.56	0/1641	0.62	0/2235
1	G	0.58	0/1637	0.66	0/2228
1	K	0.56	0/1634	0.65	1/2226 (0.0%)
1	L	0.58	0/1650	0.69	0/2244
1	O	0.56	0/1642	0.67	0/2235
1	R	0.60	0/1646	0.66	0/2240
1	U	0.55	0/1654	0.64	0/2247
1	X	0.58	0/1648	0.69	1/2241 (0.0%)
2	C	0.57	0/1605	0.69	0/2205
2	F	0.59	0/1621	0.69	3/2225 (0.1%)
2	H	0.57	0/1645	0.68	1/2255 (0.0%)
2	J	0.57	0/1633	0.68	1/2239 (0.0%)
2	N	0.56	0/1645	0.66	0/2257
2	Q	0.57	0/1636	0.67	0/2244
2	T	0.55	0/1628	0.69	2/2234 (0.1%)
2	W	0.60	0/1637	0.67	1/2245 (0.0%)
3	A	0.53	0/1480	0.65	0/2011
3	B	0.55	0/1452	0.68	0/1979
3	E	0.56	0/1469	0.66	0/1997
3	I	0.57	0/1482	0.70	0/2011
3	M	0.52	0/1472	0.67	0/2003
3	P	0.58	0/1469	0.67	0/1997
3	S	0.52	0/1454	0.64	0/1978
3	V	0.53	0/1449	0.64	0/1972
All	All	0.56	0/37929	0.67	10/51748 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	L	0	1

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	W	18	LEU	CA-CB-CG	6.22	129.60	115.30
1	X	21	LEU	CA-CB-CG	6.14	129.42	115.30
2	F	186	LEU	CA-CB-CG	5.96	129.00	115.30
2	H	186	LEU	CA-CB-CG	5.81	128.67	115.30
2	T	186	LEU	CA-CB-CG	5.73	128.48	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	L	51	ALA	Peptide

## 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	D	1605	0	1535	20	0
1	G	1601	0	1535	20	0
1	K	1598	0	1519	12	0
1	L	1614	0	1552	22	0
1	O	1606	0	1537	11	0
1	R	1610	0	1541	28	0
1	U	1618	0	1570	15	0
1	X	1612	0	1552	18	0
2	C	1563	0	1485	34	0
2	F	1579	0	1504	15	0
2	H	1603	0	1541	20	0
2	J	1591	0	1525	19	0
2	N	1603	0	1528	11	0
2	Q	1594	0	1521	27	0
2	T	1586	0	1516	20	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	W	1595	0	1530	8	0
3	A	1452	0	1387	20	0
3	B	1424	0	1344	17	0
3	E	1441	0	1376	26	0
3	I	1454	0	1409	21	0
3	M	1444	0	1372	23	0
3	P	1441	0	1383	19	0
3	S	1427	0	1347	21	0
3	V	1422	0	1356	14	0
4	A	84	0	75	1	0
4	B	84	0	75	0	0
4	E	84	0	75	0	0
4	I	98	0	86	1	0
4	M	98	0	85	1	0
4	P	84	0	75	0	0
4	S	84	0	75	0	0
4	V	84	0	75	1	0
5	A	11	0	9	0	0
5	B	11	0	10	0	0
5	E	11	0	9	0	0
5	I	22	0	19	1	0
5	M	33	0	29	0	0
5	P	11	0	10	0	0
5	S	11	0	10	0	0
5	V	11	0	10	0	0
6	A	11	0	10	0	0
6	E	11	0	10	0	0
6	I	11	0	10	2	0
6	M	22	0	19	0	0
7	A	5	0	0	1	0
7	E	5	0	0	0	0
7	G	5	0	0	0	0
7	K	5	0	0	0	0
7	O	5	0	0	0	0
7	U	5	0	0	0	0
7	V	5	0	0	1	0
7	X	5	0	0	0	0
8	A	30	0	0	3	0
8	B	22	0	0	2	0
8	C	45	0	0	3	0
8	D	55	0	0	1	0
8	E	30	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	F	40	0	0	3	0
8	G	50	0	0	0	0
8	H	47	0	0	0	0
8	I	41	0	0	4	0
8	J	44	0	0	2	0
8	K	63	0	0	0	0
8	L	56	0	0	1	0
8	M	35	0	0	2	0
8	N	40	0	0	0	0
8	O	57	0	0	0	0
8	P	41	0	0	3	0
8	Q	51	0	0	3	0
8	R	62	0	0	2	0
8	S	23	0	0	2	0
8	T	41	0	0	1	0
8	U	42	0	0	1	0
8	V	24	0	0	1	0
8	W	50	0	0	1	0
8	X	47	0	0	2	0
All	All	39035	0	36241	453	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 453 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)
2:F:113:GLN:HG3	8:F:248:HOH:O	1.36	1.21
1:R:24:ARG:HG2	1:R:24:ARG:HH11	1.09	1.11
3:E:327:ILE:HD11	3:E:345:LEU:HD22	1.35	1.08
2:H:1:GLN:HE21	2:H:1:GLN:HA	1.14	1.07
1:X:142:ARG:HH11	1:X:142:ARG:HG2	1.18	1.06

There are no symmetry-related clashes.



## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	211/213 (99%)	204 (97%)	6 (3%)	1 (0%)	32	56
1	G	211/213 (99%)	201 (95%)	10 (5%)	0	100	100
1	K	211/213 (99%)	201 (95%)	10 (5%)	0	100	100
1	L	211/213 (99%)	203 (96%)	5 (2%)	3 (1%)	13	25
1	O	211/213 (99%)	199 (94%)	11 (5%)	1 (0%)	32	56
1	R	211/213 (99%)	205 (97%)	5 (2%)	1 (0%)	32	56
1	U	211/213 (99%)	200 (95%)	10 (5%)	1 (0%)	32	56
1	X	211/213 (99%)	200 (95%)	9 (4%)	2 (1%)	20	39
2	C	209/223 (94%)	199 (95%)	8 (4%)	2 (1%)	18	36
2	F	211/223 (95%)	197 (93%)	11 (5%)	3 (1%)	13	25
2	H	213/223 (96%)	199 (93%)	12 (6%)	2 (1%)	20	39
2	J	213/223 (96%)	198 (93%)	11 (5%)	4 (2%)	9	17
2	N	214/223 (96%)	203 (95%)	11 (5%)	0	100	100
2	Q	213/223 (96%)	198 (93%)	14 (7%)	1 (0%)	32	56
2	T	212/223 (95%)	201 (95%)	10 (5%)	1 (0%)	32	56
2	W	212/223 (95%)	205 (97%)	7 (3%)	0	100	100
3	A	193/214 (90%)	188 (97%)	4 (2%)	1 (0%)	32	56
3	B	192/214 (90%)	184 (96%)	7 (4%)	1 (0%)	32	56
3	E	192/214 (90%)	185 (96%)	7 (4%)	0	100	100
3	I	191/214 (89%)	181 (95%)	8 (4%)	2 (1%)	18	36
3	M	193/214 (90%)	177 (92%)	14 (7%)	2 (1%)	18	36
3	P	191/214 (89%)	178 (93%)	12 (6%)	1 (0%)	32	56
3	S	192/214 (90%)	176 (92%)	16 (8%)	0	100	100
3	V	191/214 (89%)	179 (94%)	12 (6%)	0	100	100
All	All	4920/5200 (95%)	4661 (95%)	230 (5%)	29 (1%)	28	52

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	51	ALA
1	L	52	SER
2	C	200	GLN
3	B	501	SER
2	F	164	SER

### 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	D	177/181 (98%)	170 (96%)	7 (4%)	36	62
1	G	176/181 (97%)	165 (94%)	11 (6%)	21	40
1	K	176/181 (97%)	164 (93%)	12 (7%)	18	35
1	L	179/181 (99%)	172 (96%)	7 (4%)	37	63
1	O	177/181 (98%)	168 (95%)	9 (5%)	28	51
1	R	178/181 (98%)	165 (93%)	13 (7%)	16	32
1	U	180/181 (99%)	168 (93%)	12 (7%)	19	36
1	X	179/181 (99%)	169 (94%)	10 (6%)	25	46
2	C	174/192 (91%)	156 (90%)	18 (10%)	8	15
2	F	177/192 (92%)	167 (94%)	10 (6%)	25	46
2	H	181/192 (94%)	168 (93%)	13 (7%)	17	32
2	J	178/192 (93%)	165 (93%)	13 (7%)	16	32
2	N	180/192 (94%)	167 (93%)	13 (7%)	17	32
2	Q	178/192 (93%)	165 (93%)	13 (7%)	16	32
2	T	178/192 (93%)	162 (91%)	16 (9%)	11	21
2	W	180/192 (94%)	169 (94%)	11 (6%)	22	42
3	A	158/188 (84%)	154 (98%)	4 (2%)	53	77
3	B	153/188 (81%)	146 (95%)	7 (5%)	31	56
3	E	157/188 (84%)	153 (98%)	4 (2%)	53	77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	I	161/188 (86%)	149 (92%)	12 (8%)	16	30
3	M	156/188 (83%)	146 (94%)	10 (6%)	20	39
3	P	158/188 (84%)	151 (96%)	7 (4%)	33	58
3	S	153/188 (81%)	145 (95%)	8 (5%)	27	50
3	V	155/188 (82%)	138 (89%)	17 (11%)	7	12
All	All	4099/4488 (91%)	3842 (94%)	257 (6%)	21	40

5 of 257 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	501	SER
3	M	394	HIS
2	W	199	THR
1	O	73	LEU
2	N	67	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 84 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	124	GLN
1	O	90	GLN
1	X	90	GLN
1	K	160	GLN
2	J	200	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

74 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
7	SO4	A	1	-	4,4,4	0.17	0	6,6,6	0.12	0
4	NAG	A	3281	3,4	14,14,15	0.57	0	15,19,21	1.32	2 (13%)
4	NAG	A	3282	5,4	14,14,15	0.54	0	15,19,21	1.06	1 (6%)
5	BMA	A	3283	4,6	11,11,12	0.98	0	13,15,17	1.35	2 (15%)
6	MAN	A	3284	5	11,11,12	1.46	2 (18%)	13,15,17	2.63	6 (46%)
4	NAG	A	3371	3	14,14,15	0.60	0	15,19,21	1.22	2 (13%)
4	NAG	A	3891	3	14,14,15	0.49	0	15,19,21	1.84	4 (26%)
4	NAG	A	4201	3,4	14,14,15	0.63	0	15,19,21	1.17	2 (13%)
4	NAG	A	4202	4	14,14,15	0.56	0	15,19,21	0.75	0
4	NAG	B	3281	3,4	14,14,15	0.49	0	15,19,21	1.20	2 (13%)
4	NAG	B	3282	5,4	14,14,15	0.49	0	15,19,21	1.28	1 (6%)
5	BMA	B	3283	4	11,11,12	0.66	0	13,15,17	1.01	0
4	NAG	B	3371	3	14,14,15	0.57	0	15,19,21	1.20	1 (6%)
4	NAG	B	3891	3	14,14,15	0.59	0	15,19,21	1.74	4 (26%)
4	NAG	B	4201	3,4	14,14,15	0.56	0	15,19,21	1.23	2 (13%)
4	NAG	B	4202	4	14,14,15	0.74	1 (7%)	15,19,21	1.50	3 (20%)
7	SO4	E	1	-	4,4,4	0.17	0	6,6,6	0.46	0
4	NAG	E	3281	3,4	14,14,15	0.67	0	15,19,21	1.26	1 (6%)
4	NAG	E	3282	5,4	14,14,15	0.62	0	15,19,21	1.06	1 (6%)
5	BMA	E	3283	4,6	11,11,12	0.67	0	13,15,17	0.74	0
6	MAN	E	3284	5	11,11,12	1.30	1 (9%)	13,15,17	3.23	7 (53%)
4	NAG	E	3371	3	14,14,15	0.54	0	15,19,21	1.22	2 (13%)
4	NAG	E	3891	3	14,14,15	0.59	0	15,19,21	1.25	2 (13%)
4	NAG	E	4201	3,4	14,14,15	0.51	0	15,19,21	1.04	1 (6%)
4	NAG	E	4202	4	14,14,15	0.40	0	15,19,21	1.51	2 (13%)
7	SO4	G	214	-	4,4,4	0.17	0	6,6,6	0.26	0
4	NAG	I	3281	3,4	14,14,15	0.58	0	15,19,21	1.44	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	I	3282	5,4	14,14,15	0.64	0	15,19,21	1.24	1 (6%)
5	BMA	I	3283	4,6	11,11,12	0.82	0	13,15,17	2.41	4 (30%)
6	MAN	I	3284	5	11,11,12	0.65	0	13,15,17	0.52	0
4	NAG	I	3371	3	14,14,15	0.76	1 (7%)	15,19,21	1.04	1 (6%)
4	NAG	I	3891	3,4	14,14,15	0.52	0	15,19,21	1.51	3 (20%)
4	NAG	I	3892	5,4	14,14,15	0.64	0	15,19,21	1.46	3 (20%)
5	BMA	I	3893	4	11,11,12	0.62	0	13,15,17	1.21	2 (15%)
4	NAG	I	4201	3,4	14,14,15	0.57	0	15,19,21	2.05	4 (26%)
4	NAG	I	4202	4	14,14,15	0.47	0	15,19,21	0.97	2 (13%)
7	SO4	K	214	-	4,4,4	0.17	0	6,6,6	0.24	0
4	NAG	M	3281	3,4	14,14,15	0.59	0	15,19,21	1.01	1 (6%)
4	NAG	M	3282	5,4	14,14,15	0.71	0	15,19,21	1.13	1 (6%)
5	BMA	M	3283	4,6	11,11,12	0.67	0	13,15,17	1.15	1 (7%)
6	MAN	M	3284	5,6	11,11,12	1.48	2 (18%)	13,15,17	2.63	6 (46%)
6	MAN	M	3285	6	11,11,12	1.53	2 (18%)	13,15,17	1.40	1 (7%)
4	NAG	M	3371	3	14,14,15	0.70	1 (7%)	15,19,21	2.21	3 (20%)
4	NAG	M	3891	3,4	14,14,15	0.57	0	15,19,21	1.38	2 (13%)
4	NAG	M	3892	5,4	14,14,15	0.57	0	15,19,21	1.16	1 (6%)
5	BMA	M	3893	4	11,11,12	0.72	0	13,15,17	0.95	0
4	NAG	M	4201	3,4	14,14,15	0.66	0	15,19,21	1.47	3 (20%)
4	NAG	M	4202	5,4	14,14,15	0.53	0	15,19,21	1.00	1 (6%)
5	BMA	M	4203	4	11,11,12	0.78	0	13,15,17	1.05	1 (7%)
7	SO4	O	214	-	4,4,4	0.14	0	6,6,6	0.15	0
4	NAG	P	3281	3,4	14,14,15	0.63	0	15,19,21	1.06	2 (13%)
4	NAG	P	3282	5,4	14,14,15	0.69	0	15,19,21	1.76	3 (20%)
5	BMA	P	3283	4	11,11,12	0.82	0	13,15,17	0.86	0
4	NAG	P	3371	3	14,14,15	0.57	0	15,19,21	1.02	1 (6%)
4	NAG	P	3891	3	14,14,15	0.62	0	15,19,21	2.11	1 (6%)
4	NAG	P	4201	3,4	14,14,15	0.53	0	15,19,21	1.51	2 (13%)
4	NAG	P	4202	4	14,14,15	0.52	0	15,19,21	1.89	4 (26%)
4	NAG	S	3281	3,4	14,14,15	0.59	0	15,19,21	1.06	1 (6%)
4	NAG	S	3282	5,4	14,14,15	0.62	0	15,19,21	1.58	3 (20%)
5	BMA	S	3283	4	11,11,12	0.62	0	13,15,17	1.01	2 (15%)
4	NAG	S	3371	3	14,14,15	0.70	1 (7%)	15,19,21	1.59	4 (26%)
4	NAG	S	3891	3	14,14,15	0.66	0	15,19,21	1.36	1 (6%)
4	NAG	S	4201	3,4	14,14,15	0.64	0	15,19,21	2.15	4 (26%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	S	4202	4	14,14,15	0.58	0	15,19,21	1.55	3 (20%)
7	SO4	U	214	-	4,4,4	0.19	0	6,6,6	0.11	0
7	SO4	V	1	-	4,4,4	0.18	0	6,6,6	0.16	0
4	NAG	V	3281	3,4	14,14,15	0.43	0	15,19,21	1.28	3 (20%)
4	NAG	V	3282	5,4	14,14,15	0.68	0	15,19,21	1.16	2 (13%)
5	BMA	V	3283	4	11,11,12	0.65	0	13,15,17	0.97	1 (7%)
4	NAG	V	3371	3	14,14,15	0.50	0	15,19,21	1.30	2 (13%)
4	NAG	V	3891	3	14,14,15	0.46	0	15,19,21	2.03	4 (26%)
4	NAG	V	4201	3,4	14,14,15	0.60	0	15,19,21	1.75	3 (20%)
4	NAG	V	4202	4	14,14,15	0.42	0	15,19,21	1.67	2 (13%)
7	SO4	X	214	-	4,4,4	0.16	0	6,6,6	0.18	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
7	SO4	A	1	-	-	0/0/0/0	0/0/0/0
4	NAG	A	3281	3,4	-	0/6/23/26	0/1/1/1
4	NAG	A	3282	5,4	-	0/6/23/26	0/1/1/1
5	BMA	A	3283	4,6	-	0/2/19/22	0/1/1/1
6	MAN	A	3284	5	-	0/2/19/22	0/1/1/1
4	NAG	A	3371	3	-	2/6/23/26	0/1/1/1
4	NAG	A	3891	3	-	0/6/23/26	0/1/1/1
4	NAG	A	4201	3,4	-	0/6/23/26	0/1/1/1
4	NAG	A	4202	4	-	0/6/23/26	0/1/1/1
4	NAG	B	3281	3,4	-	0/6/23/26	0/1/1/1
4	NAG	B	3282	5,4	-	0/6/23/26	0/1/1/1
5	BMA	B	3283	4	-	0/2/19/22	0/1/1/1
4	NAG	B	3371	3	-	0/6/23/26	0/1/1/1
4	NAG	B	3891	3	-	0/6/23/26	0/1/1/1
4	NAG	B	4201	3,4	-	0/6/23/26	0/1/1/1
4	NAG	B	4202	4	-	0/6/23/26	0/1/1/1
7	SO4	E	1	-	-	0/0/0/0	0/0/0/0
4	NAG	E	3281	3,4	-	0/6/23/26	0/1/1/1
4	NAG	E	3282	5,4	-	0/6/23/26	0/1/1/1
5	BMA	E	3283	4,6	-	0/2/19/22	0/1/1/1
6	MAN	E	3284	5	-	0/2/19/22	1/1/1/1
4	NAG	E	3371	3	-	1/6/23/26	0/1/1/1
4	NAG	E	3891	3	-	0/6/23/26	0/1/1/1

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	E	4201	3,4	-	0/6/23/26	0/1/1/1
4	NAG	E	4202	4	-	0/6/23/26	0/1/1/1
7	SO4	G	214	-	-	0/0/0/0	0/0/0/0
4	NAG	I	3281	3,4	-	0/6/23/26	0/1/1/1
4	NAG	I	3282	5,4	-	0/6/23/26	0/1/1/1
5	BMA	I	3283	4,6	-	0/2/19/22	0/1/1/1
6	MAN	I	3284	5	1/1/4/5	0/2/19/22	0/1/1/1
4	NAG	I	3371	3	-	0/6/23/26	0/1/1/1
4	NAG	I	3891	3,4	-	0/6/23/26	0/1/1/1
4	NAG	I	3892	5,4	-	0/6/23/26	0/1/1/1
5	BMA	I	3893	4	-	0/2/19/22	0/1/1/1
4	NAG	I	4201	3,4	-	0/6/23/26	0/1/1/1
4	NAG	I	4202	4	-	0/6/23/26	0/1/1/1
7	SO4	K	214	-	-	0/0/0/0	0/0/0/0
4	NAG	M	3281	3,4	-	0/6/23/26	0/1/1/1
4	NAG	M	3282	5,4	-	0/6/23/26	0/1/1/1
5	BMA	M	3283	4,6	-	0/2/19/22	0/1/1/1
6	MAN	M	3284	5,6	-	0/2/19/22	0/1/1/1
6	MAN	M	3285	6	-	0/2/19/22	1/1/1/1
4	NAG	M	3371	3	-	0/6/23/26	0/1/1/1
4	NAG	M	3891	3,4	-	0/6/23/26	0/1/1/1
4	NAG	M	3892	5,4	-	0/6/23/26	0/1/1/1
5	BMA	M	3893	4	-	0/2/19/22	0/1/1/1
4	NAG	M	4201	3,4	-	0/6/23/26	0/1/1/1
4	NAG	M	4202	5,4	-	0/6/23/26	0/1/1/1
5	BMA	M	4203	4	-	0/2/19/22	0/1/1/1
7	SO4	O	214	-	-	0/0/0/0	0/0/0/0
4	NAG	P	3281	3,4	-	0/6/23/26	0/1/1/1
4	NAG	P	3282	5,4	-	0/6/23/26	0/1/1/1
5	BMA	P	3283	4	-	0/2/19/22	0/1/1/1
4	NAG	P	3371	3	-	0/6/23/26	0/1/1/1
4	NAG	P	3891	3	-	0/6/23/26	0/1/1/1
4	NAG	P	4201	3,4	-	0/6/23/26	0/1/1/1
4	NAG	P	4202	4	-	0/6/23/26	0/1/1/1
4	NAG	S	3281	3,4	-	0/6/23/26	0/1/1/1
4	NAG	S	3282	5,4	-	0/6/23/26	0/1/1/1
5	BMA	S	3283	4	-	0/2/19/22	0/1/1/1
4	NAG	S	3371	3	-	0/6/23/26	0/1/1/1
4	NAG	S	3891	3	-	0/6/23/26	0/1/1/1
4	NAG	S	4201	3,4	-	0/6/23/26	0/1/1/1
4	NAG	S	4202	4	-	0/6/23/26	0/1/1/1
7	SO4	U	214	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	SO4	V	1	-	-	0/0/0/0	0/0/0/0
4	NAG	V	3281	3,4	-	0/6/23/26	0/1/1/1
4	NAG	V	3282	5,4	-	0/6/23/26	0/1/1/1
5	BMA	V	3283	4	-	0/2/19/22	0/1/1/1
4	NAG	V	3371	3	-	0/6/23/26	0/1/1/1
4	NAG	V	3891	3	-	0/6/23/26	0/1/1/1
4	NAG	V	4201	3,4	-	0/6/23/26	0/1/1/1
4	NAG	V	4202	4	-	0/6/23/26	0/1/1/1
7	SO4	X	214	-	-	0/0/0/0	0/0/0/0

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	E	3284	MAN	O5-C1	-2.32	1.39	1.43
6	A	3284	MAN	C6-C5	2.03	1.58	1.51
6	M	3284	MAN	C6-C5	2.05	1.58	1.51
4	S	3371	NAG	C1-C2	2.15	1.55	1.52
4	B	4202	NAG	C1-C2	2.18	1.55	1.52

The worst 5 of 139 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	E	3284	MAN	C2-C3-C4	-7.83	97.22	110.88
4	P	3282	NAG	O5-C1-C2	-4.52	105.18	111.47
4	S	4201	NAG	O5-C1-C2	-4.28	105.52	111.47
6	E	3284	MAN	O2-C2-C1	-4.27	100.49	109.18
4	I	4201	NAG	O5-C1-C2	-4.22	105.60	111.47

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	I	3284	MAN	C1

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	E	3371	NAG	O7-C7-N2-C2
4	A	3371	NAG	O7-C7-N2-C2
4	A	3371	NAG	C8-C7-N2-C2

All (2) ring outliers are listed below:



Mol	Chain	Res	Type	Atoms
6	M	3285	MAN	C1-C2-C3-C4-C5-O5
6	E	3284	MAN	C1-C2-C3-C4-C5-O5

8 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1	SO4	1	0
4	A	3891	NAG	1	0
5	I	3283	BMA	1	0
6	I	3284	MAN	2	0
4	I	3891	NAG	1	0
4	M	4201	NAG	1	0
7	V	1	SO4	1	0
4	V	3281	NAG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	D	213/213 (100%)	-0.14	1 (0%) 90 90	15, 26, 31, 33	0
1	G	213/213 (100%)	-0.08	0 100 100	17, 26, 32, 34	0
1	K	213/213 (100%)	-0.20	0 100 100	18, 27, 33, 35	0
1	L	213/213 (100%)	-0.10	1 (0%) 90 90	18, 26, 32, 35	0
1	O	213/213 (100%)	-0.16	0 100 100	18, 26, 33, 36	0
1	R	213/213 (100%)	-0.10	1 (0%) 90 90	16, 25, 31, 35	0
1	U	213/213 (100%)	-0.15	0 100 100	18, 26, 31, 34	0
1	X	213/213 (100%)	-0.11	0 100 100	15, 26, 32, 33	0
2	C	213/223 (95%)	0.17	4 (1%) 67 63	20, 27, 39, 41	0
2	F	215/223 (96%)	0.04	2 (0%) 84 82	20, 28, 39, 47	0
2	H	217/223 (97%)	0.07	2 (0%) 84 82	17, 27, 40, 44	0
2	J	217/223 (97%)	0.16	3 (1%) 75 72	19, 28, 38, 42	0
2	N	218/223 (97%)	-0.01	4 (1%) 69 65	18, 28, 39, 49	0
2	Q	217/223 (97%)	-0.07	0 100 100	19, 28, 39, 42	0
2	T	216/223 (96%)	0.10	5 (2%) 61 56	19, 28, 40, 43	0
2	W	216/223 (96%)	0.01	3 (1%) 75 72	19, 27, 39, 44	0
3	A	195/214 (91%)	-0.12	0 100 100	18, 28, 38, 45	0
3	B	194/214 (90%)	-0.09	1 (0%) 90 90	19, 29, 40, 45	0
3	E	194/214 (90%)	-0.03	0 100 100	15, 28, 40, 46	0
3	I	193/214 (90%)	-0.08	1 (0%) 90 90	17, 26, 40, 48	0
3	M	195/214 (91%)	-0.03	1 (0%) 90 90	16, 27, 42, 47	0
3	P	193/214 (90%)	-0.08	0 100 100	18, 27, 38, 47	0
3	S	194/214 (90%)	-0.13	0 100 100	17, 29, 40, 44	0
3	V	193/214 (90%)	-0.13	1 (0%) 90 90	19, 27, 39, 42	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
All	All	4984/5200 (95%)	-0.05	30 (0%) 89 88	15, 27, 38, 49	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	135	SER	3.1
2	N	198	GLY	3.1
3	M	485	LEU	2.9
2	C	1	GLN	2.9
2	J	198	GLY	2.8

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	V	3281	14/15	0.96	0.18	1.05	22,25,29,33	0
4	NAG	B	3281	14/15	0.95	0.16	0.43	28,31,33,37	0
4	NAG	V	4201	14/15	0.93	0.14	0.35	32,38,39,41	0
7	SO4	E	1	5/5	0.90	0.15	-0.29	56,57,58,58	0
4	NAG	I	3281	14/15	0.97	0.13	-0.55	22,25,27,30	0
4	NAG	P	3281	14/15	0.98	0.13	-0.79	18,22,26,31	0
4	NAG	A	3281	14/15	0.97	0.13	-0.91	27,31,34,38	0
4	NAG	S	3281	14/15	0.96	0.12	-0.94	25,27,30,33	0
4	NAG	E	3281	14/15	0.97	0.10	-1.48	17,22,27,29	0
7	SO4	V	1	5/5	0.97	0.11	-1.62	59,60,60,60	0
4	NAG	M	3281	14/15	0.96	0.12	-1.67	18,27,28,30	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	SO4	A	1	5/5	0.95	0.13	-1.75	50,50,51,51	0
4	NAG	S	4202	14/15	0.83	0.22	-	47,53,56,56	0
4	NAG	P	3371	14/15	0.86	0.16	-	39,42,44,44	0
4	NAG	S	3371	14/15	0.83	0.27	-	40,41,43,43	0
4	NAG	P	4201	14/15	0.94	0.14	-	32,38,42,46	0
7	SO4	U	214	5/5	0.94	0.14	-	76,77,77,78	0
4	NAG	M	4201	14/15	0.93	0.15	-	35,39,43,48	0
4	NAG	S	3891	14/15	0.80	0.18	-	44,46,50,51	0
4	NAG	E	3891	14/15	0.79	0.21	-	40,43,46,46	0
4	NAG	V	3371	14/15	0.75	0.33	-	45,48,52,52	0
4	NAG	B	3282	14/15	0.92	0.17	-	38,41,44,48	0
4	NAG	I	4201	14/15	0.94	0.15	-	33,38,43,46	0
6	MAN	M	3284	11/12	0.66	0.27	-	23,31,33,35	0
4	NAG	E	4201	14/15	0.93	0.14	-	32,36,39,43	0
4	NAG	P	3282	14/15	0.93	0.13	-	36,41,42,47	0
5	BMA	I	3283	11/12	0.81	0.25	-	54,59,62,63	0
4	NAG	A	3891	14/15	0.82	0.17	-	40,42,44,44	0
4	NAG	B	3891	14/15	0.80	0.20	-	44,46,51,52	0
6	MAN	M	3285	11/12	0.80	0.39	-	26,33,34,36	0
5	BMA	A	3283	11/12	0.87	0.14	-	49,50,53,56	0
5	BMA	P	3283	11/12	0.67	0.20	-	50,52,53,54	0
4	NAG	A	3282	14/15	0.94	0.14	-	37,40,43,47	0
7	SO4	G	214	5/5	0.94	0.11	-	63,63,63,64	0
5	BMA	S	3283	11/12	0.73	0.22	-	50,52,53,53	0
4	NAG	M	3892	14/15	0.92	0.21	-	44,45,46,47	0
6	MAN	I	3284	11/12	0.63	0.33	-	23,31,33,35	0
4	NAG	I	4202	14/15	0.88	0.14	-	49,51,53,53	0
6	MAN	E	3284	11/12	0.89	0.17	-	23,31,33,35	0
4	NAG	V	4202	14/15	0.90	0.17	-	42,45,47,47	0
5	BMA	I	3893	11/12	0.76	0.22	-	43,44,45,45	0
4	NAG	I	3891	14/15	0.88	0.17	-	33,36,37,40	0
4	NAG	V	3282	14/15	0.92	0.15	-	34,37,40,42	0
4	NAG	S	3282	14/15	0.93	0.15	-	37,41,43,47	0
5	BMA	M	4203	11/12	0.75	0.24	-	62,64,64,65	0
5	BMA	B	3283	11/12	0.82	0.17	-	49,50,51,52	0
4	NAG	B	4201	14/15	0.94	0.14	-	34,37,40,45	0
4	NAG	B	3371	14/15	0.77	0.27	-	47,50,52,52	0
5	BMA	E	3283	11/12	0.87	0.23	-	45,46,48,51	0
4	NAG	A	3371	14/15	0.82	0.19	-	45,46,47,47	0
7	SO4	O	214	5/5	0.96	0.11	-	70,70,70,71	0
4	NAG	M	3282	14/15	0.92	0.15	-	26,33,36,41	0
4	NAG	A	4201	14/15	0.95	0.17	-	37,40,42,45	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	BMA	V	3283	11/12	0.90	0.17	-	44,45,47,48	0
4	NAG	M	3371	14/15	0.87	0.21	-	42,45,47,47	0
7	SO4	K	214	5/5	0.91	0.18	-	67,68,68,69	0
4	NAG	M	4202	14/15	0.92	0.15	-	50,56,59,60	0
4	NAG	E	3282	14/15	0.96	0.14	-	32,34,37,41	0
4	NAG	I	3892	14/15	0.93	0.23	-	43,44,45,45	0
4	NAG	M	3891	14/15	0.89	0.15	-	35,40,42,43	0
4	NAG	E	4202	14/15	0.80	0.16	-	47,50,52,52	0
4	NAG	P	4202	14/15	0.86	0.14	-	47,50,50,50	0
4	NAG	I	3371	14/15	0.78	0.20	-	43,47,48,49	0
7	SO4	X	214	5/5	0.88	0.16	-	69,70,70,70	0
4	NAG	A	4202	14/15	0.85	0.24	-	49,52,52,52	0
5	BMA	M	3893	11/12	0.85	0.18	-	43,45,46,46	0
5	BMA	M	3283	11/12	0.78	0.24	-	45,48,50,55	0
6	MAN	A	3284	11/12	0.71	0.27	-	23,31,33,35	0
4	NAG	I	3282	14/15	0.93	0.17	-	32,36,41,47	0
4	NAG	B	4202	14/15	0.85	0.25	-	49,52,54,54	0
4	NAG	P	3891	14/15	0.78	0.20	-	41,44,47,49	0
4	NAG	V	3891	14/15	0.84	0.15	-	40,43,45,45	0
4	NAG	S	4201	14/15	0.93	0.17	-	36,41,44,47	0
4	NAG	E	3371	14/15	0.86	0.13	-	43,45,47,48	0

## 6.5 Other polymers

There are no such residues in this entry.