



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:15 PM GMT

PDB ID : 4HLZ  
Title : Crystal Structure of Fab C179 in Complex with a H2N2 influenza virus hemagglutinin  
Authors : Dreyfus, C.; Wilson, I.A.  
Deposited on : 2012-10-17  
Resolution : 2.90 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

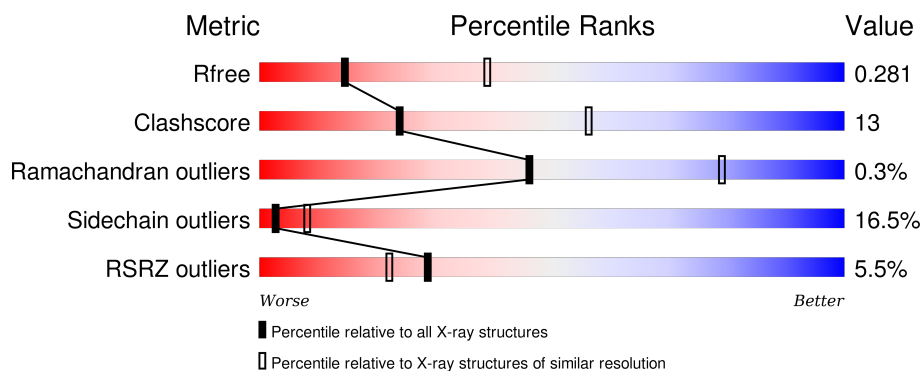
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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))
$R_{free}$	91344	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	327	<div> <div>5%</div> <div>65%</div> <div>29%</div> <div>5%</div> </div>
1	C	327	<div> <div>3%</div> <div>69%</div> <div>28%</div> <div>••</div> </div>
1	E	327	<div> <div>5%</div> <div>66%</div> <div>28%</div> <div>5%</div> </div>
2	B	174	<div> <div>%</div> <div>72%</div> <div>24%</div> <div>••</div> </div>
2	D	174	<div> <div>2%</div> <div>70%</div> <div>25%</div> <div>••</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	174	
3	G	229	
3	I	229	
3	K	229	
4	H	214	
4	J	214	
4	L	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	NAG	A	403	-	-	-	X
7	SO4	F	201	-	-	X	-
7	SO4	H	301	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 21839 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 Hemagglutinin HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	324	Total	C	N	O	S	0	4	0
			2546	1600	441	490	15			
1	C	324	Total	C	N	O	S	0	4	0
			2546	1600	441	490	15			
1	E	324	Total	C	N	O	S	0	5	0
			2551	1605	441	490	15			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	PRO	-	EXPRESSION TAG	UNP C7S226
C	9	PRO	-	EXPRESSION TAG	UNP C7S226
E	9	PRO	-	EXPRESSION TAG	UNP C7S226

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	168	Total	C	N	O	S	0	1	0
			1370	854	235	272	9			
2	D	170	Total	C	N	O	S	0	1	0
			1384	862	237	276	9			
2	F	172	Total	C	N	O	S	0	1	0
			1404	876	240	279	9			

- Molecule 3 is a protein called Fab C179 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	223	Total	C	N	O	S	0	1	0
			1696	1071	283	334	8			
3	I	221	Total	C	N	O	S	0	0	0
			1675	1059	278	330	8			

*Continued on next page...*

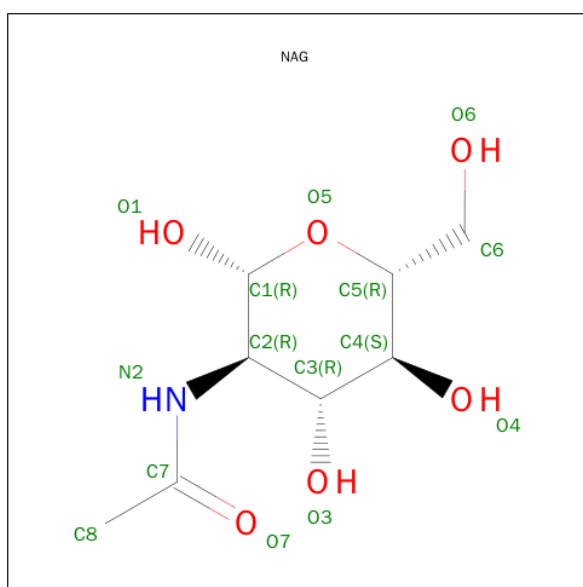
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	K	200	Total	C	N	O	S	0	0	0
			1532	970	253	302	7			

- Molecule 4 is a protein called Fab C179 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	211	Total	C	N	O	S	0	0	0
			1627	1021	270	330	6			
4	J	212	Total	C	N	O	S	0	0	0
			1635	1025	272	332	6			
4	L	205	Total	C	N	O	S	0	0	0
			1580	994	261	319	6			

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

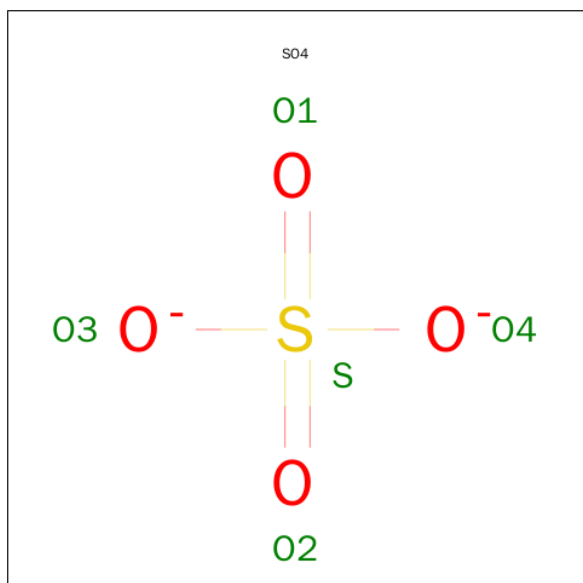


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		
5	F	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	3	Total	C	N	O	0	0
			39	22	2	15		
6	E	3	Total	C	N	O	0	0
			39	22	2	15		

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	O	S	0	0
			5	4	1		
7	D	1	Total	O	S	0	0
			5	4	1		
7	E	1	Total	O	S	0	0
			5	4	1		
7	F	1	Total	O	S	0	0
			5	4	1		
7	H	1	Total	O	S	0	0
			5	4	1		
7	J	1	Total	O	S	0	0
			5	4	1		
7	L	1	Total	O	S	0	0
			5	4	1		

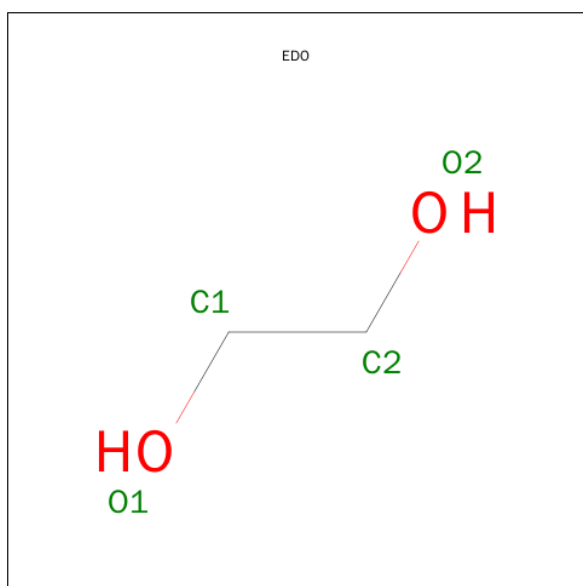
- Molecule 8 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	C	4	Total	C	N	O	0	0
			50	28	2	20		

- Molecule 9 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	C	2	Total	C	N	O	0	0
			28	16	2	10		
9	E	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 10 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).

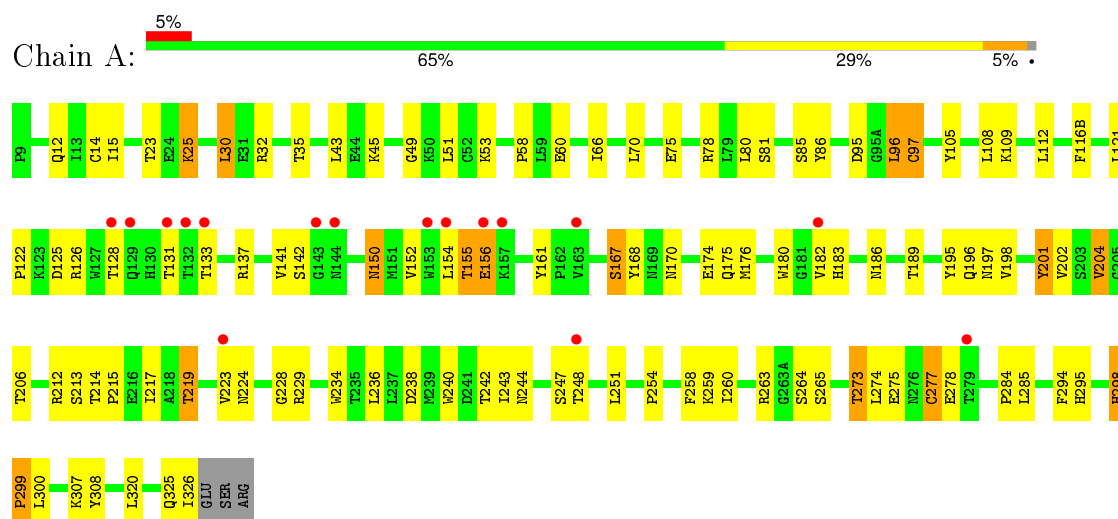


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	J	1	Total	C	O	0	0
			4	2	2		

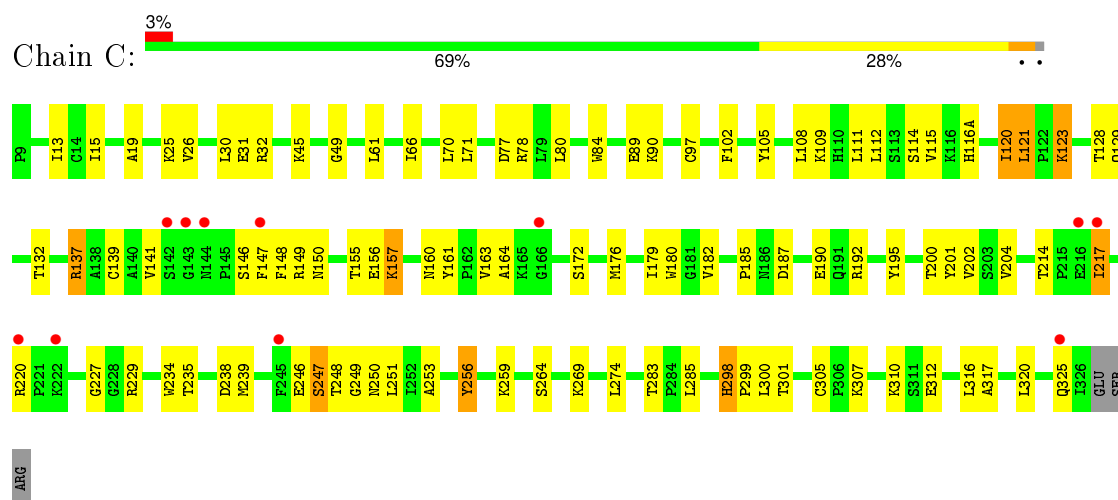
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Hemagglutinin HA1



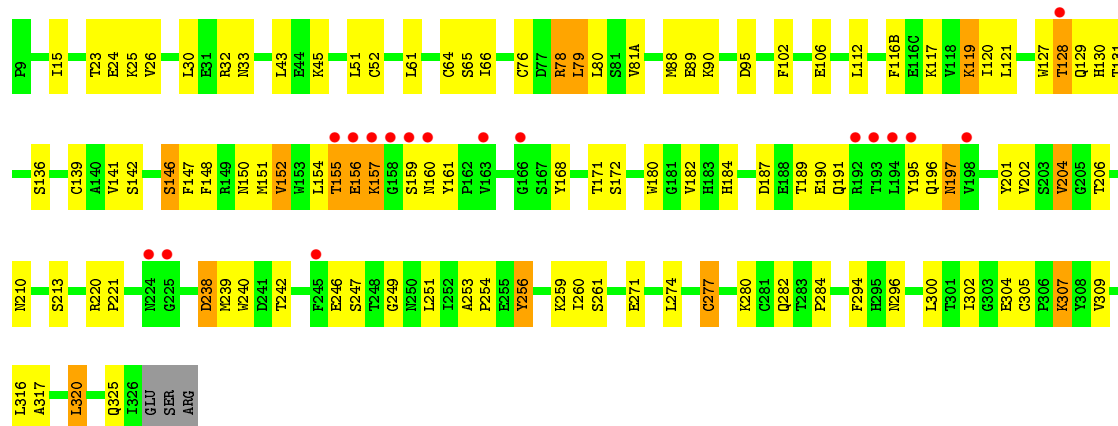
#### • Molecule 1: Hemagglutinin HA1



#### • Molecule 1: Hemagglutinin HA1







• Molecule 2: Hemagglutinin HA2



• Molecule 2: Hemagglutinin HA2

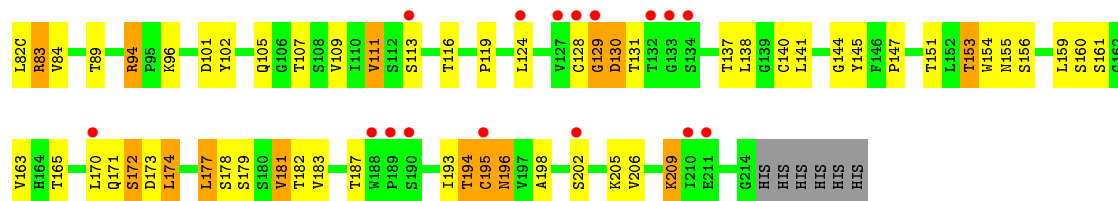


• Molecule 2: Hemagglutinin HA2

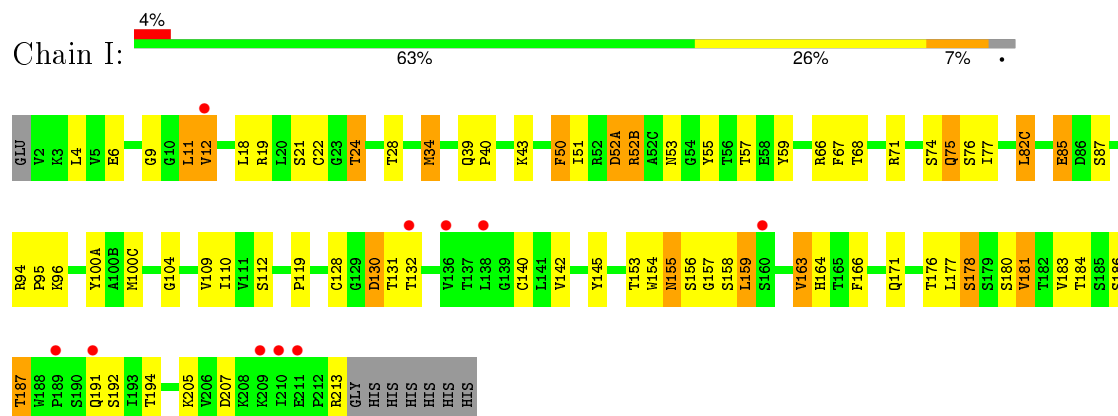


• Molecule 3: Fab C179 heavy chain

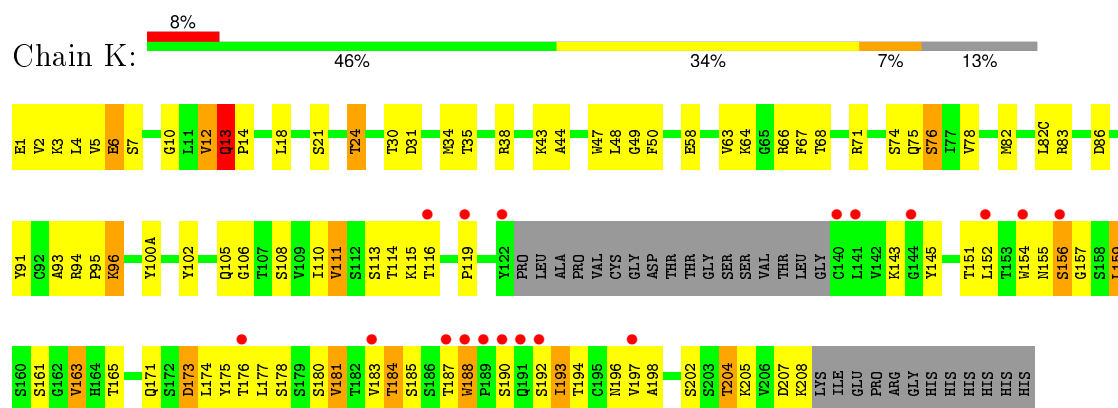




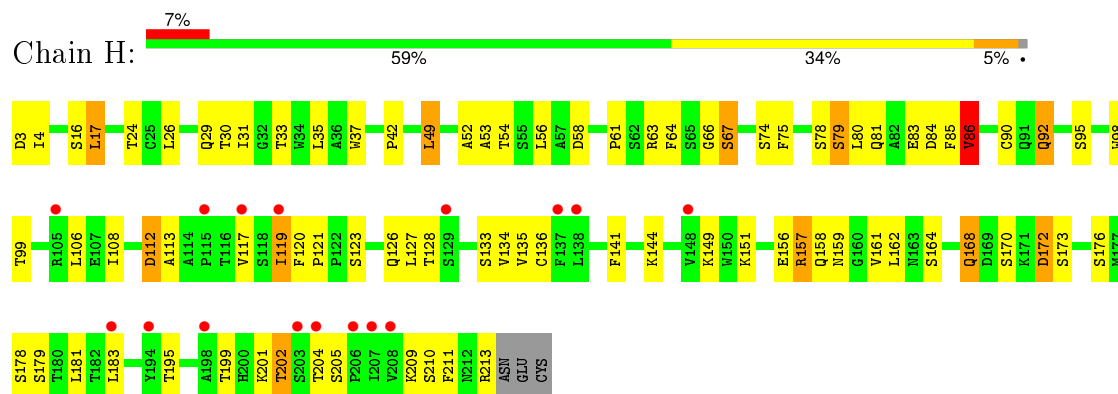
• Molecule 3: Fab C179 heavy chain



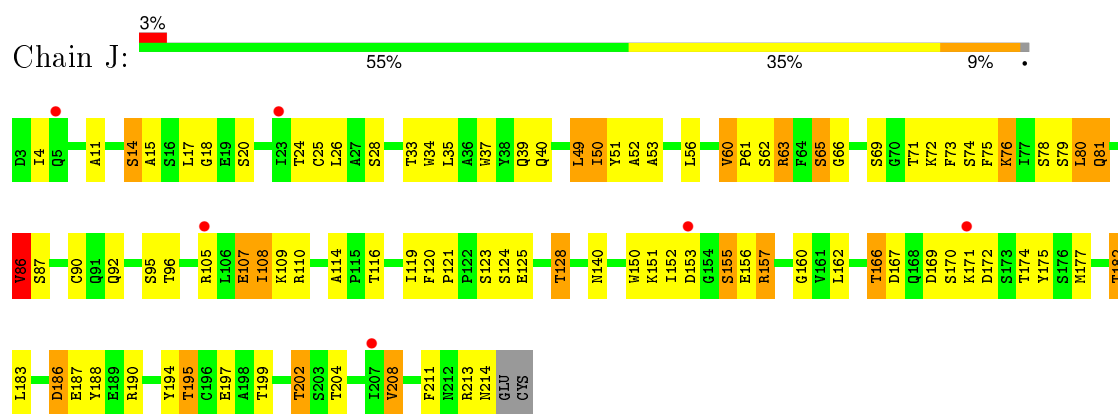
• Molecule 3: Fab C179 heavy chain



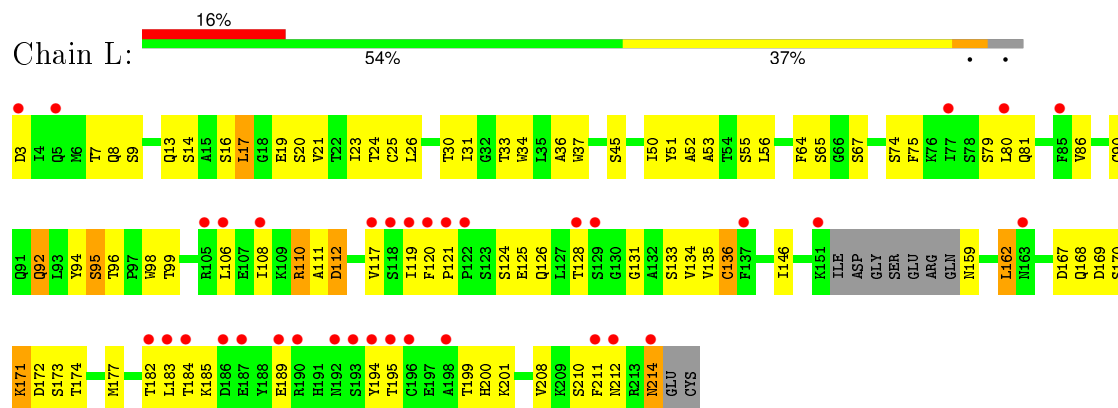
• Molecule 4: Fab C179 light chain



• Molecule 4: Fab C179 light chain



• Molecule 4: Fab C179 light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	136.58Å 150.78Å 217.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.55 – 2.90 39.55 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.3 (39.55-2.90) 98.6 (39.55-2.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.19 (at 2.90Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, $R_{free}$	0.228 , 0.284 0.231 , 0.281	Depositor DCC
$R_{free}$ test set	5001 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	86.1	Xtriage
Anisotropy	0.063	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 72.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 98674 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	21839	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	101.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.21% 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, 0.375 respectively for untwinned datasets, and 0.333, 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, EDO, 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	A	0.47	1/2612 (0.0%)	0.66	1/3545 (0.0%)
1	C	0.46	0/2612	0.64	1/3545 (0.0%)
1	E	0.48	0/2620	0.66	0/3556
2	B	0.57	0/1399	0.69	0/1876
2	D	0.56	0/1413	0.68	0/1895
2	F	0.54	0/1435	0.67	0/1926
3	G	0.47	0/1742	0.67	1/2378 (0.0%)
3	I	0.46	0/1718	0.71	2/2347 (0.1%)
3	K	0.47	0/1571	0.66	1/2143 (0.0%)
4	H	0.44	0/1666	0.61	0/2264
4	J	0.40	0/1674	0.62	0/2275
4	L	0.42	0/1618	0.63	1/2199 (0.0%)
All	All	0.48	1/22080 (0.0%)	0.66	7/29949 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	299	PRO	N-CD	5.35	1.55	1.47

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	9	GLY	N-CA-C	-9.83	88.53	113.10
1	C	298	HIS	C-N-CD	6.03	141.07	128.40
3	K	13	GLN	C-N-CD	5.64	140.25	128.40
1	A	298	HIS	C-N-CD	5.56	140.07	128.40
4	L	3	ASP	CB-CG-OD2	5.25	123.02	118.30

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	2546	0	2499	63	0
1	C	2546	0	2500	58	0
1	E	2551	0	2510	68	0
2	B	1370	0	1282	21	0
2	D	1384	0	1294	32	0
2	F	1404	0	1311	45	0
3	G	1696	0	1661	53	0
3	I	1675	0	1636	57	0
3	K	1532	0	1486	64	0
4	H	1627	0	1561	41	0
4	J	1635	0	1567	60	0
4	L	1580	0	1516	47	0
5	A	28	0	26	0	0
5	B	14	0	13	0	0
5	E	14	0	13	0	0
5	F	14	0	13	0	0
6	A	39	0	34	0	0
6	E	39	0	34	0	0
7	B	5	0	0	0	0
7	D	5	0	0	1	0
7	E	5	0	0	0	0
7	F	5	0	0	3	0
7	H	5	0	0	0	0
7	J	5	0	0	1	0
7	L	5	0	0	1	0
8	C	50	0	43	0	0
9	C	28	0	25	0	0
9	E	28	0	25	0	0
10	J	4	0	6	1	0
All	All	21839	0	21055	560	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 13.

The worst 5 of 560 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)
3:I:154:TRP:CZ3	3:I:163:VAL:HG21	1.64	1.32
3:I:154:TRP:HZ3	3:I:163:VAL:CG2	1.49	1.26
4:H:121:PRO:HG3	4:H:211:PHE:CE1	1.71	1.24
3:I:154:TRP:CZ3	3:I:163:VAL:CG2	2.20	1.23
4:H:121:PRO:HG3	4:H:211:PHE:CZ	1.88	1.09

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	325/327 (99%)	316 (97%)	9 (3%)	0	100	100
1	C	325/327 (99%)	315 (97%)	10 (3%)	0	100	100
1	E	326/327 (100%)	309 (95%)	17 (5%)	0	100	100
2	B	165/174 (95%)	157 (95%)	8 (5%)	0	100	100
2	D	167/174 (96%)	153 (92%)	14 (8%)	0	100	100
2	F	171/174 (98%)	162 (95%)	9 (5%)	0	100	100
3	G	222/229 (97%)	204 (92%)	16 (7%)	2 (1%)	21	57
3	I	219/229 (96%)	207 (94%)	11 (5%)	1 (0%)	34	71
3	K	196/229 (86%)	177 (90%)	16 (8%)	3 (2%)	13	42
4	H	209/214 (98%)	200 (96%)	8 (4%)	1 (0%)	34	71
4	J	210/214 (98%)	195 (93%)	14 (7%)	1 (0%)	34	71
4	L	201/214 (94%)	190 (94%)	10 (5%)	1 (0%)	34	71
All	All	2736/2832 (97%)	2585 (94%)	142 (5%)	9 (0%)	46	79

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	K	12	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	L	86	VAL
3	G	130	ASP
4	J	86	VAL
4	H	86	VAL

### 5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	288/288 (100%)	245 (85%)	43 (15%)	4	11
1	C	288/288 (100%)	260 (90%)	28 (10%)	10	30
1	E	289/288 (100%)	249 (86%)	40 (14%)	4	13
2	B	147/151 (97%)	128 (87%)	19 (13%)	5	16
2	D	148/151 (98%)	129 (87%)	19 (13%)	5	16
2	F	150/151 (99%)	136 (91%)	14 (9%)	11	32
3	G	192/197 (98%)	153 (80%)	39 (20%)	1	4
3	I	190/197 (96%)	153 (80%)	37 (20%)	2	5
3	K	173/197 (88%)	139 (80%)	34 (20%)	1	5
4	H	184/187 (98%)	142 (77%)	42 (23%)	1	3
4	J	185/187 (99%)	140 (76%)	45 (24%)	1	2
4	L	179/187 (96%)	139 (78%)	40 (22%)	1	3
All	All	2413/2469 (98%)	2013 (83%)	400 (17%)	3	8

5 of 400 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	131	THR
4	H	127	LEU
4	L	30	THR
3	G	160	SER
4	H	16	SER



Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
3	G	155	ASN
3	I	191	GLN
4	H	40	GLN
3	G	39	GLN
4	H	212	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

14 carbohydrates 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	$\# Z  > 2$
6	NAG	A	402	1,6	14,14,15	0.51	0	15,19,21	1.04	0
6	NAG	A	403	6	14,14,15	0.55	0	15,19,21	1.27	3 (20%)
6	BMA	A	404	6	11,11,12	0.65	0	14,15,17	0.64	0
8	NAG	C	401	1,8	14,14,15	0.49	0	15,19,21	1.24	1 (6%)
8	NAG	C	402	8	14,14,15	0.48	0	15,19,21	0.91	0
8	BMA	C	403	8	11,11,12	0.56	0	14,15,17	1.23	2 (14%)
8	MAN	C	404	8	11,11,12	0.53	0	14,15,17	0.65	0
9	NAG	C	405	1,9	14,14,15	0.43	0	15,19,21	0.79	0
9	NAG	C	406	9	14,14,15	0.42	0	15,19,21	0.74	0
6	NAG	E	403	1,6	14,14,15	0.46	0	15,19,21	1.58	2 (13%)
6	NAG	E	404	6	14,14,15	0.48	0	15,19,21	1.27	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	BMA	E	405	6	11,11,12	0.53	0	14,15,17	0.89	0
9	NAG	E	406	9,1	14,14,15	0.52	0	15,19,21	0.80	0
9	NAG	E	407	9	14,14,15	0.44	0	15,19,21	0.73	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
6	NAG	A	402	1,6	-	0/6/23/26	0/1/1/1
6	NAG	A	403	6	-	0/6/23/26	0/1/1/1
6	BMA	A	404	6	-	0/2/19/22	0/1/1/1
8	NAG	C	401	1,8	-	0/6/23/26	0/1/1/1
8	NAG	C	402	8	-	0/6/23/26	0/1/1/1
8	BMA	C	403	8	-	0/2/19/22	0/1/1/1
8	MAN	C	404	8	-	0/2/19/22	0/1/1/1
9	NAG	C	405	1,9	-	0/6/23/26	0/1/1/1
9	NAG	C	406	9	-	0/6/23/26	0/1/1/1
6	NAG	E	403	1,6	-	0/6/23/26	0/1/1/1
6	NAG	E	404	6	-	0/6/23/26	0/1/1/1
6	BMA	E	405	6	-	0/2/19/22	0/1/1/1
9	NAG	E	406	9,1	-	0/6/23/26	0/1/1/1
9	NAG	E	407	9	-	0/6/23/26	0/1/1/1

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(°)
8	C	401	NAG	C2-N2-C7	-4.23	117.61	123.04
6	E	404	NAG	C4-C3-C2	-2.61	107.18	111.23
8	C	403	BMA	C1-O5-C5	-2.57	108.99	112.25
6	A	403	NAG	O7-C7-C8	-2.16	118.09	122.06
6	A	403	NAG	O7-C7-N2	2.07	126.09	121.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry

13 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$
5	NAG	A	401	1	14,14,15	0.48	0	15,19,21	1.11	1 (6%)
5	NAG	A	405	1	14,14,15	0.57	0	15,19,21	1.06	1 (6%)
7	SO4	B	201	-	4,4,4	0.34	0	6,6,6	0.36	0
5	NAG	B	202	2	14,14,15	0.73	0	15,19,21	1.27	2 (13%)
7	SO4	D	201	-	4,4,4	0.24	0	6,6,6	0.22	0
5	NAG	E	401	1	14,14,15	0.38	0	15,19,21	1.37	2 (13%)
7	SO4	E	402	-	4,4,4	0.18	0	6,6,6	0.14	0
7	SO4	F	201	-	4,4,4	0.26	0	6,6,6	0.25	0
5	NAG	F	202	2	14,14,15	0.54	0	15,19,21	0.90	1 (6%)
7	SO4	H	301	-	4,4,4	0.17	0	6,6,6	0.13	0
7	SO4	J	301	-	4,4,4	0.10	0	6,6,6	0.13	0
10	EDO	J	302	-	3,3,3	0.60	0	2,2,2	0.24	0
7	SO4	L	301	-	4,4,4	0.24	0	6,6,6	0.29	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
5	NAG	A	401	1	-	0/6/23/26	0/1/1/1
5	NAG	A	405	1	-	0/6/23/26	0/1/1/1
7	SO4	B	201	-	-	0/0/0/0	0/0/0/0
5	NAG	B	202	2	-	0/6/23/26	0/1/1/1
7	SO4	D	201	-	-	0/0/0/0	0/0/0/0
5	NAG	E	401	1	-	0/6/23/26	0/1/1/1
7	SO4	E	402	-	-	0/0/0/0	0/0/0/0
7	SO4	F	201	-	-	0/0/0/0	0/0/0/0
5	NAG	F	202	2	-	0/6/23/26	0/1/1/1
7	SO4	H	301	-	-	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	J	301	-	-	0/0/0/0	0/0/0/0
10	EDO	J	302	-	-	0/1/1/1	0/0/0/0
7	SO4	L	301	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	401	NAG	C2-N2-C7	-2.79	119.45	123.04
5	F	202	NAG	C2-N2-C7	-2.62	119.67	123.04
5	B	202	NAG	C1-O5-C5	2.63	115.58	112.25
5	B	202	NAG	C2-N2-C7	2.86	126.72	123.04
5	A	405	NAG	C1-O5-C5	3.13	116.22	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	201	SO4	1	0
7	F	201	SO4	3	0
7	J	301	SO4	1	0
10	J	302	EDO	1	0
7	L	301	SO4	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	324/327 (99%)	0.07	16 (4%)	33 27	44, 95, 147, 177	2 (0%)
1	C	324/327 (99%)	0.05	11 (3%)	49 41	49, 100, 147, 169	2 (0%)
1	E	324/327 (99%)	0.20	17 (5%)	31 24	49, 105, 153, 223	2 (0%)
2	B	168/174 (96%)	0.05	1 (0%)	90 89	45, 67, 108, 138	0
2	D	170/174 (97%)	0.12	3 (1%)	71 68	44, 79, 145, 169	0
2	F	172/174 (98%)	0.11	1 (0%)	90 89	44, 82, 126, 176	0
3	G	223/229 (97%)	0.30	17 (7%)	17 11	30, 100, 151, 246	0
3	I	221/229 (96%)	0.25	10 (4%)	37 31	45, 95, 166, 190	0
3	K	200/229 (87%)	0.48	18 (9%)	12 7	32, 89, 195, 245	0
4	H	211/214 (98%)	0.45	16 (7%)	17 11	42, 108, 155, 185	0
4	J	212/214 (99%)	0.15	6 (2%)	56 50	49, 122, 158, 213	0
4	L	205/214 (95%)	0.75	35 (17%)	2 1	46, 120, 185, 268	0
All	All	2754/2832 (97%)	0.24	151 (5%)	29 22	30, 97, 158, 268	6 (0%)

The worst 5 of 151 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	K	189	PRO	13.6
4	L	194	TYR	7.6
4	L	121	PRO	6.6
4	L	117	VAL	6.1
3	K	188	TRP	5.5

### 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 [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. 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
6	NAG	A	403	14/15	0.87	0.28	2.25	97,113,135,138	0
8	NAG	C	401	14/15	0.89	0.26	1.39	94,109,131,134	0
9	NAG	C	405	14/15	0.86	0.17	-	134,152,166,178	0
8	BMA	C	403	11/12	0.87	0.30	-	103,119,135,148	0
6	NAG	A	402	14/15	0.86	0.16	-	90,122,136,143	0
6	NAG	E	404	14/15	0.88	0.25	-	89,125,140,155	0
6	NAG	E	403	14/15	0.89	0.30	-	78,117,137,139	0
9	NAG	E	407	14/15	0.78	0.35	-	143,165,173,175	0
9	NAG	C	406	14/15	0.75	0.32	-	135,182,189,192	0
8	NAG	C	402	14/15	0.92	0.38	-	82,112,130,131	0
9	NAG	E	406	14/15	0.93	0.16	-	108,135,153,164	0
6	BMA	A	404	11/12	0.80	0.32	-	105,121,129,135	0
8	MAN	C	404	11/12	0.86	0.28	-	145,159,166,169	0
6	BMA	E	405	11/12	0.70	0.33	-	143,165,176,185	0

### 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. 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
7	SO4	H	301	5/5	0.95	0.24	2.59	84,88,102,104	0
7	SO4	D	201	5/5	0.96	0.22	1.73	80,90,96,101	0
7	SO4	F	201	5/5	0.96	0.26	1.34	84,89,95,99	0
7	SO4	B	201	5/5	0.99	0.21	0.55	54,69,74,74	0
5	NAG	F	202	14/15	0.86	0.17	0.17	100,114,121,125	0
5	NAG	B	202	14/15	0.85	0.15	-0.44	103,114,125,125	0
7	SO4	J	301	5/5	0.92	0.17	-0.73	110,123,135,151	0
7	SO4	L	301	5/5	0.92	0.16	-1.65	89,106,111,123	0
5	NAG	A	401	14/15	0.86	0.19	-	129,142,146,150	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	E	402	5/5	0.88	0.16	-	107,124,128,130	0
10	EDO	J	302	4/4	0.89	0.08	-	72,74,76,81	0
5	NAG	A	405	14/15	0.88	0.32	-	113,129,153,159	0
5	NAG	E	401	14/15	0.89	0.20	-	92,106,121,124	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.