



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 11:55 pm GMT

PDB ID : 4PY8  
Title : Crystal structure of Fab 3.1 in complex with the 1918 influenza virus hemagglutinin  
Authors : Dreyfus, C.  
Deposited on : 2014-03-26  
Resolution : 2.91 Å(reported)

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

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

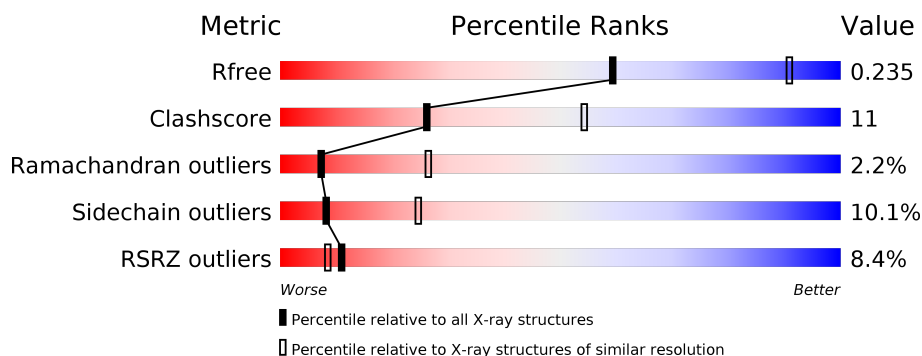
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.91 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	111664	1983 (2.94-2.90)
Clashscore	122126	2200 (2.94-2.90)
Ramachandran outliers	120053	2150 (2.94-2.90)
Sidechain outliers	120020	2152 (2.94-2.90)
RSRZ outliers	108989	1928 (2.94-2.90)

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	331	<div> <div>2%</div> <div>73%</div> <div>25%</div> <div>..</div> </div>
2	B	179	<div> <div>12%</div> <div>84%</div> <div>7%</div> <div>...</div> </div>
3	I	219	<div> <div>13%</div> <div>64%</div> <div>29%</div> <div>6%</div> <div>.</div> </div>
4	J	214	<div> <div>9%</div> <div>66%</div> <div>29%</div> <div>..</div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7222 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 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	328	Total	C	N	O	S	0	0	0
			2528	1595	433	489	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	EXPRESSION TAG	UNP Q9WFX3
A	8	ASP	-	EXPRESSION TAG	UNP Q9WFX3
A	9	PRO	-	EXPRESSION TAG	UNP Q9WFX3
A	10	GLY	-	EXPRESSION TAG	UNP Q9WFX3

- Molecule 2 is a protein called Hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	171	Total	C	N	O	S	0	1	0
			1376	857	237	276	6			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	177	SER	-	EXPRESSION TAG	UNP Q9WFX3
B	178	GLY	-	EXPRESSION TAG	UNP Q9WFX3
B	179	ARG	-	EXPRESSION TAG	UNP Q9WFX3

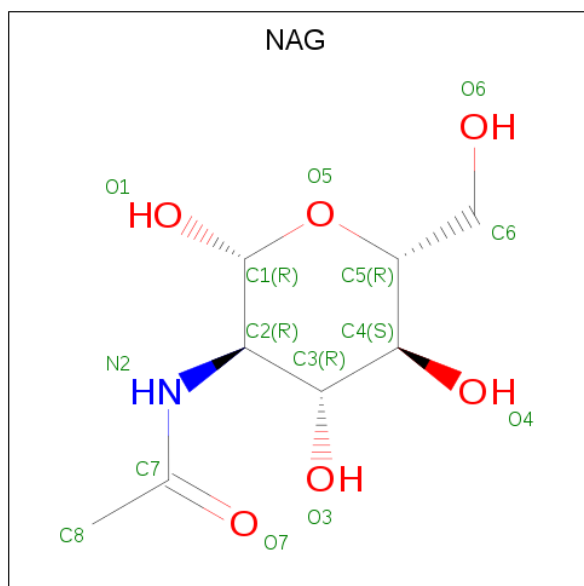
- Molecule 3 is a protein called antibody 3.1 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	219	Total	C	N	O	S	0	0	0
			1644	1038	281	317	8			

- Molecule 4 is a protein called antibody 3.1 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	J	212	Total 1604	C 1005	N 270	O 324	S 5	0	0	0

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



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

- Molecule 6 is MALONATE ION (three-letter code: MLI) (formula:  $C_3H_2O_4$ ).

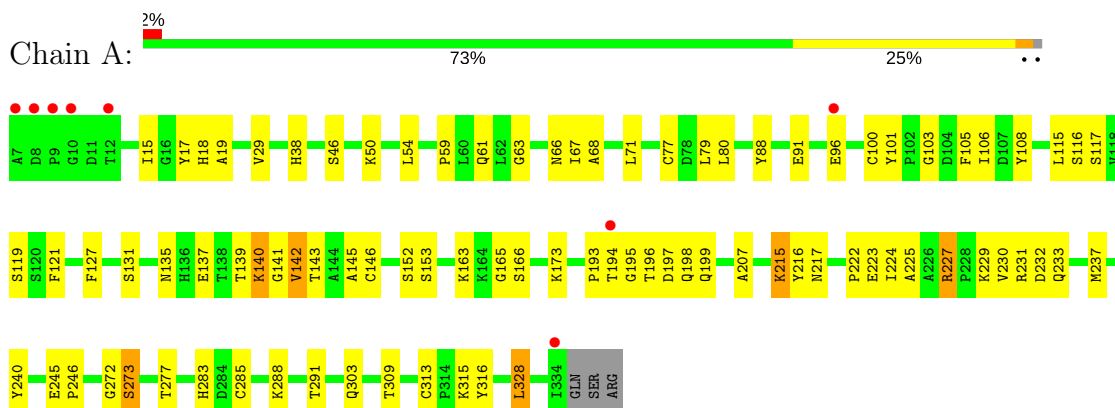


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	I	1	Total	C	O	0	0
			7	3	4		
6	J	1	Total	C	O	0	0
			7	3	4		

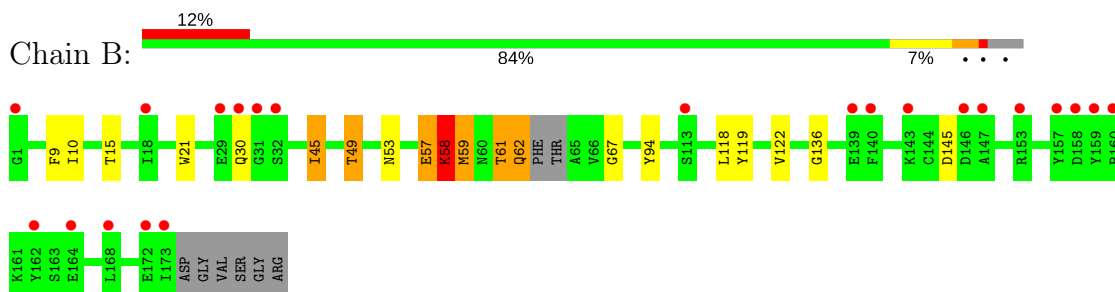
### 3 Residue-property plots [i](#)

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.

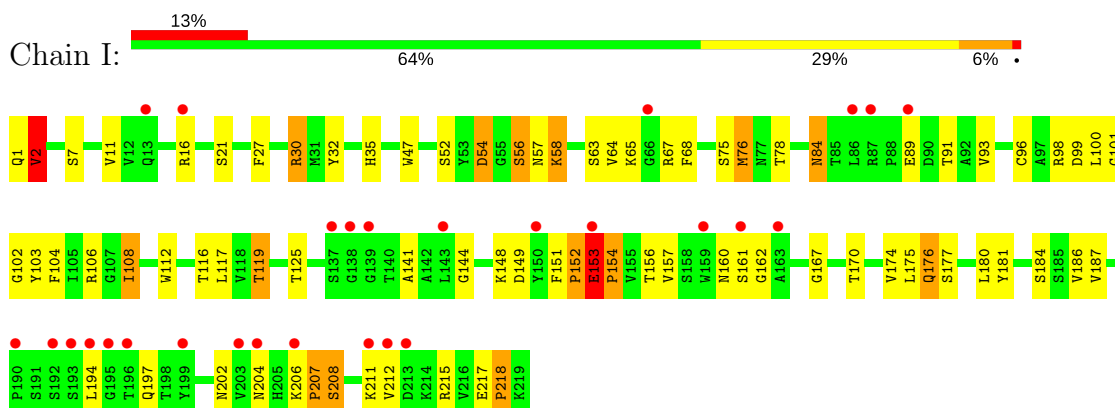
#### • Molecule 1: Hemagglutinin HA1 chain



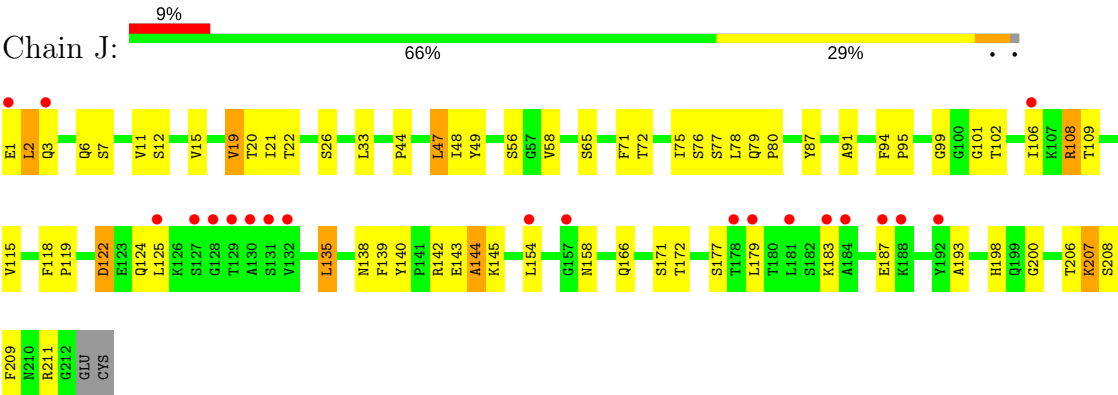
#### • Molecule 2: Hemagglutinin HA2 chain



#### • Molecule 3: antibody 3.1 heavy chain



● Molecule 4: antibody 3.1 light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.06Å 135.06Å 230.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.84 – 2.91 42.84 – 2.91	Depositor EDS
% Data completeness (in resolution range)	99.6 (42.84-2.91) 99.6 (42.84-2.91)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.74 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.193 , 0.242 0.188 , 0.235	Depositor DCC
$R_{free}$ test set	1734 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	80.6	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 54.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.025 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7222	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.15% 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: MLI, NAG

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	0.60	0/2593	0.72	0/3536
2	B	0.60	0/1404	0.68	0/1889
3	I	0.55	0/1684	0.73	0/2298
4	J	0.56	0/1639	0.70	0/2227
All	All	0.58	0/7320	0.71	0/9950

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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	2528	0	2433	62	0
2	B	1376	0	1284	17	0
3	I	1644	0	1592	50	0
4	J	1604	0	1558	38	0
5	A	56	0	51	2	0
6	I	7	0	2	0	0
6	J	7	0	2	1	0
All	All	7222	0	6922	158	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 11.

All (158) 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:153:GLU:HB3	3:I:154:PRO:HD3	1.20	1.14
1:A:117:SER:HA	1:A:273:SER:HB2	1.36	1.02
1:A:141:GLY:HA2	1:A:142:VAL:HB	1.39	0.99
3:I:153:GLU:HB3	3:I:154:PRO:CD	1.96	0.95
3:I:54:ASP:OD2	3:I:56:SER:HB2	1.66	0.94
1:A:225:ALA:HB1	1:A:227:ARG:HH21	1.35	0.91
1:A:231:ARG:NH2	5:A:402:NAG:O3	2.05	0.90
4:J:143:GLU:OE2	4:J:143:GLU:HA	1.73	0.89
2:B:58:LYS:HG3	2:B:59:MET:N	1.87	0.88
1:A:137:GLU:OE2	1:A:139:THR:HG22	1.73	0.88
1:A:141:GLY:CA	1:A:142:VAL:HB	2.06	0.84
3:I:57:ASN:HA	3:I:58:LYS:HB2	1.60	0.84
1:A:116:SER:O	1:A:273:SER:HB2	1.82	0.80
1:A:315:LYS:HD2	2:B:62:GLN:OE1	1.81	0.80
1:A:195:GLY:HA2	1:A:196:THR:HB	1.63	0.79
3:I:144:GLY:HA3	3:I:186:VAL:HG12	1.65	0.78
1:A:224:ILE:HG13	1:A:224:ILE:O	1.84	0.77
1:A:117:SER:HA	1:A:273:SER:CB	2.15	0.76
3:I:102:GLY:HA2	3:I:104:PHE:N	2.00	0.76
1:A:195:GLY:HA2	1:A:197:ASP:H	1.50	0.76
1:A:195:GLY:CA	1:A:197:ASP:H	2.03	0.72
2:B:61:THR:O	2:B:62:GLN:HG3	1.90	0.71
3:I:76:MET:CE	3:I:76:MET:HA	2.20	0.71
3:I:76:MET:HE3	3:I:76:MET:HA	1.71	0.71
3:I:64:VAL:HG21	3:I:68:PHE:CG	2.28	0.69
3:I:153:GLU:CB	3:I:154:PRO:HD3	2.12	0.69
1:A:142:VAL:HG13	1:A:152:SER:HB3	1.74	0.68
1:A:225:ALA:CB	1:A:227:ARG:HH21	2.04	0.68
1:A:195:GLY:CA	1:A:197:ASP:N	2.58	0.67
3:I:35:HIS:HD2	3:I:47:TRP:HE1	1.45	0.65
1:A:71:LEU:HD11	1:A:115:LEU:HD11	1.79	0.65
3:I:57:ASN:HA	3:I:58:LYS:CB	2.27	0.64
1:A:101:TYR:CD2	1:A:237:MET:HB2	2.34	0.62
4:J:108:ARG:HD3	4:J:109:THR:O	1.99	0.62
4:J:122:ASP:HA	4:J:125:LEU:HD23	1.81	0.62
1:A:61:GLN:HE21	1:A:63:GLY:H	1.44	0.62
3:I:217:GLU:HB2	3:I:218:PRO:HD2	1.80	0.62
1:A:50:LYS:HG2	1:A:283:HIS:HD2	1.66	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:J:144:ALA:HA	4:J:145:LYS:HB3	1.82	0.61
1:A:106:ILE:HG13	1:A:240:TYR:CE2	2.36	0.61
4:J:108:ARG:HD2	4:J:140:TYR:CB	2.31	0.61
1:A:193:PRO:O	1:A:194:THR:C	2.40	0.60
1:A:225:ALA:HB1	1:A:227:ARG:NH2	2.13	0.60
3:I:57:ASN:CA	3:I:58:LYS:HB2	2.31	0.59
1:A:103:GLY:HA3	1:A:237:MET:O	2.02	0.59
4:J:91:ALA:O	6:J:301:MLI:O6	2.20	0.59
3:I:1:GLN:HG2	3:I:2:VAL:H	1.69	0.58
3:I:100:LEU:HD23	3:I:106:ARG:HD3	1.84	0.58
4:J:183:LYS:O	4:J:187:GLU:HG2	2.04	0.58
3:I:1:GLN:NE2	3:I:98:ARG:HH22	2.03	0.57
3:I:67:ARG:O	3:I:84:ASN:ND2	2.37	0.57
1:A:15:ILE:HG23	2:B:118:LEU:HD23	1.85	0.57
3:I:108:ILE:HD12	4:J:49:TYR:HB2	1.86	0.57
1:A:91:GLU:O	1:A:277:THR:HA	2.05	0.56
1:A:116:SER:O	1:A:273:SER:CB	2.52	0.56
3:I:151:PHE:CE2	3:I:152:PRO:HB3	2.40	0.56
1:A:96:GLU:OE1	5:A:402:NAG:H62	2.05	0.56
3:I:64:VAL:HG21	3:I:68:PHE:CD2	2.41	0.55
4:J:6:GLN:HE21	4:J:99:GLY:HA3	1.71	0.55
4:J:198:HIS:CD2	4:J:200:GLY:H	2.24	0.55
1:A:195:GLY:HA2	1:A:196:THR:CB	2.32	0.54
1:A:309:THR:HB	1:A:313:CYS:SG	2.49	0.53
3:I:206:LYS:CB	3:I:207:PRO:HD3	2.38	0.53
4:J:138:ASN:C	4:J:172:THR:HG21	2.29	0.53
4:J:6:GLN:NE2	4:J:101:GLY:H	2.07	0.53
3:I:27:PHE:O	3:I:30:ARG:HD2	2.08	0.53
3:I:93:VAL:HG12	3:I:117:LEU:HD13	1.90	0.52
1:A:146:CYS:O	1:A:153:SER:HB3	2.10	0.52
3:I:204:ASN:HD22	3:I:211:LYS:HG3	1.74	0.52
3:I:148:LYS:HG3	3:I:149:ASP:N	2.24	0.52
1:A:38:HIS:CD2	2:B:21:TRP:HE1	2.28	0.52
1:A:291:THR:HG22	1:A:309:THR:HG22	1.92	0.52
3:I:152:PRO:HD2	3:I:207:PRO:HB2	1.91	0.51
4:J:139:PHE:N	4:J:172:THR:HG22	2.25	0.51
1:A:15:ILE:HD11	2:B:122:VAL:HG21	1.92	0.51
1:A:195:GLY:HA3	1:A:197:ASP:N	2.25	0.50
3:I:151:PHE:CD2	3:I:152:PRO:HB3	2.47	0.50
4:J:108:ARG:HD2	4:J:140:TYR:HB2	1.92	0.50
2:B:61:THR:C	2:B:62:GLN:HE21	2.14	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:117:SER:CA	1:A:273:SER:HB2	2.25	0.50
4:J:115:VAL:HG22	4:J:207:LYS:HG3	1.94	0.50
3:I:176:GLN:H	3:I:176:GLN:HE21	1.59	0.50
2:B:53:ASN:O	2:B:57:GLU:HB3	2.12	0.50
3:I:64:VAL:HG21	3:I:68:PHE:CD1	2.47	0.49
3:I:160:ASN:C	3:I:162:GLY:H	2.16	0.49
4:J:19:VAL:HG13	4:J:75:ILE:HB	1.94	0.49
4:J:6:GLN:HE21	4:J:99:GLY:CA	2.26	0.48
1:A:145:ALA:O	1:A:231:ARG:NH1	2.45	0.48
1:A:17:TYR:HB2	1:A:328:LEU:HD11	1.95	0.48
3:I:112:TRP:CE3	4:J:44:PRO:HD2	2.49	0.48
1:A:230:VAL:HG12	1:A:231:ARG:HG3	1.96	0.48
4:J:2:LEU:O	4:J:26:SER:OG	2.16	0.48
3:I:167:GLY:O	3:I:187:VAL:HG23	2.13	0.48
2:B:119:TYR:CE1	2:B:136:GLY:HA2	2.49	0.47
4:J:144:ALA:HA	4:J:145:LYS:CB	2.44	0.47
1:A:71:LEU:HD11	1:A:115:LEU:CD1	2.43	0.47
2:B:58:LYS:HG3	2:B:59:MET:H	1.71	0.47
4:J:193:ALA:HB2	4:J:208:SER:HB3	1.96	0.47
4:J:198:HIS:HD2	4:J:200:GLY:H	1.62	0.47
1:A:315:LYS:CD	2:B:62:GLN:OE1	2.59	0.47
1:A:127:PHE:CE2	1:A:173:LYS:HB3	2.50	0.47
4:J:6:GLN:HE22	4:J:87:TYR:HA	1.79	0.47
3:I:206:LYS:HB2	3:I:207:PRO:HD3	1.97	0.47
1:A:101:TYR:HD2	1:A:237:MET:HB2	1.79	0.46
1:A:195:GLY:HA2	1:A:197:ASP:N	2.22	0.46
1:A:193:PRO:O	1:A:195:GLY:N	2.48	0.46
4:J:19:VAL:CG1	4:J:78:LEU:HG	2.45	0.46
1:A:229:LYS:HA	1:A:233:GLN:O	2.16	0.46
2:B:9:PHE:CD1	2:B:10:ILE:HG13	2.51	0.46
3:I:102:GLY:HA2	3:I:103:TYR:C	2.35	0.46
4:J:108:ARG:HG2	4:J:171:SER:HB2	1.98	0.46
4:J:33:LEU:HD22	4:J:71:PHE:CG	2.51	0.46
1:A:61:GLN:NE2	1:A:63:GLY:H	2.14	0.46
4:J:22:THR:HG22	4:J:72:THR:HG22	1.96	0.46
3:I:141:ALA:HB3	3:I:194:LEU:HD11	1.98	0.45
3:I:176:GLN:NE2	3:I:180:LEU:O	2.49	0.45
3:I:206:LYS:O	3:I:208:SER:N	2.49	0.45
4:J:209:PHE:CD1	4:J:209:PHE:C	2.90	0.45
3:I:206:LYS:C	3:I:208:SER:H	2.20	0.45
1:A:77:CYS:HB3	1:A:80:LEU:HD12	1.97	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:J:158:ASN:O	4:J:179:LEU:HD12	2.17	0.44
1:A:195:GLY:HA3	1:A:198:GLN:H	1.81	0.44
3:I:151:PHE:HA	3:I:152:PRO:HA	1.74	0.44
1:A:207:ALA:O	1:A:222:PRO:HD3	2.17	0.44
1:A:100:CYS:O	1:A:231:ARG:HD3	2.18	0.44
2:B:30:GLN:HE22	2:B:145:ASP:HB2	1.83	0.44
3:I:91:THR:HG23	3:I:119:THR:HA	2.00	0.44
4:J:118:PHE:HA	4:J:119:PRO:HD2	1.86	0.43
1:A:245:GLU:HG3	1:A:246:PRO:HD2	2.00	0.43
3:I:56:SER:HB3	3:I:57:ASN:OD1	2.19	0.43
1:A:66:ASN:O	1:A:67:ILE:C	2.55	0.43
3:I:52:SER:HB3	3:I:57:ASN:O	2.19	0.43
2:B:45:ILE:O	2:B:49:THR:OG1	2.33	0.43
3:I:7:SER:HB3	3:I:21:SER:HB2	2.00	0.43
1:A:119:SER:HB3	1:A:272:GLY:HA3	2.00	0.42
3:I:35:HIS:HE1	3:I:99:ASP:OD1	2.02	0.42
4:J:21:ILE:HG12	4:J:102:THR:HG21	2.00	0.42
1:A:59:PRO:HB3	1:A:88:TYR:CZ	2.54	0.42
3:I:175:LEU:HD13	3:I:181:TYR:CZ	2.54	0.42
4:J:115:VAL:HA	4:J:135:LEU:O	2.19	0.42
4:J:138:ASN:C	4:J:172:THR:CG2	2.87	0.42
1:A:68:ALA:HB2	1:A:108:TYR:CE1	2.55	0.42
4:J:6:GLN:HE21	4:J:99:GLY:C	2.22	0.42
2:B:59:MET:H	2:B:59:MET:HG2	1.56	0.42
4:J:94:PHE:HA	4:J:95:PRO:C	2.39	0.42
1:A:140:LYS:HA	1:A:141:GLY:HA2	1.69	0.42
1:A:19:ALA:O	2:B:15:THR:HA	2.20	0.42
3:I:1:GLN:HE22	3:I:98:ARG:HH22	1.68	0.42
4:J:79:GLN:HG3	4:J:80:PRO:HD2	2.01	0.41
3:I:32:TYR:CD1	3:I:98:ARG:HD3	2.55	0.41
1:A:215:LYS:HG2	1:A:216:TYR:N	2.35	0.41
3:I:35:HIS:O	3:I:96:CYS:HA	2.21	0.41
4:J:47:LEU:HA	4:J:58:VAL:HG21	2.03	0.41
1:A:163:LYS:HG3	1:A:163:LYS:O	2.20	0.41
3:I:101:GLY:O	3:I:102:GLY:C	2.59	0.41
4:J:108:ARG:NH2	4:J:109:THR:O	2.43	0.40
1:A:303:GLN:O	1:A:316:TYR:HA	2.21	0.40
1:A:18:HIS:HB2	2:B:21:TRP:HA	2.04	0.40

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	A	326/331 (98%)	293 (90%)	29 (9%)	4 (1%)	14	42
2	B	168/179 (94%)	158 (94%)	7 (4%)	3 (2%)	9	31
3	I	217/219 (99%)	184 (85%)	23 (11%)	10 (5%)	2	9
4	J	210/214 (98%)	193 (92%)	14 (7%)	3 (1%)	12	38
All	All	921/943 (98%)	828 (90%)	73 (8%)	20 (2%)	7	26

All (20) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	61	THR
3	I	58	LYS
3	I	152	PRO
3	I	153	GLU
1	A	140	LYS
1	A	142	VAL
1	A	165	GLY
3	I	154	PRO
3	I	161	SER
2	B	58	LYS
2	B	67	GLY
3	I	65	LYS
1	A	166	SER
3	I	208	SER
4	J	3	GLN
4	J	144	ALA
3	I	2	VAL
4	J	211	ARG
3	I	218	PRO
3	I	207	PRO

### 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	278/285 (98%)	259 (93%)	19 (7%)	17	44
2	B	145/152 (95%)	138 (95%)	7 (5%)	28	61
3	I	179/184 (97%)	151 (84%)	28 (16%)	3	8
4	J	181/185 (98%)	156 (86%)	25 (14%)	4	11
All	All	783/806 (97%)	704 (90%)	79 (10%)	8	24

All (79) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	VAL
1	A	46	SER
1	A	54	LEU
1	A	79	LEU
1	A	105	PHE
1	A	121	PHE
1	A	131	SER
1	A	135	ASN
1	A	143	THR
1	A	199	GLN
1	A	215	LYS
1	A	217	ASN
1	A	223	GLU
1	A	227	ARG
1	A	232	ASP
1	A	273	SER
1	A	285	CYS
1	A	288	LYS
1	A	328	LEU
2	B	45	ILE
2	B	49	THR
2	B	57	GLU
2	B	58	LYS
2	B	59	MET

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	62	GLN
2	B	94	TYR
3	I	2	VAL
3	I	11	VAL
3	I	16	ARG
3	I	30	ARG
3	I	54	ASP
3	I	56	SER
3	I	63	SER
3	I	75	SER
3	I	76	MET
3	I	78	THR
3	I	84	ASN
3	I	89	GLU
3	I	108	ILE
3	I	116	THR
3	I	119	THR
3	I	125	THR
3	I	153	GLU
3	I	156	THR
3	I	157	VAL
3	I	170	THR
3	I	174	VAL
3	I	176	GLN
3	I	177	SER
3	I	184	SER
3	I	197	GLN
3	I	202	ASN
3	I	212	VAL
3	I	215	ARG
4	J	1	GLU
4	J	2	LEU
4	J	7	SER
4	J	11	VAL
4	J	12	SER
4	J	15	VAL
4	J	19	VAL
4	J	20	THR
4	J	47	LEU
4	J	48	ILE
4	J	56	SER
4	J	65	SER

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
4	J	76	SER
4	J	77	SER
4	J	106	ILE
4	J	108	ARG
4	J	122	ASP
4	J	124	GLN
4	J	135	LEU
4	J	142	ARG
4	J	154	LEU
4	J	166	GLN
4	J	177	SER
4	J	206	THR
4	J	207	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (20) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	38	HIS
1	A	61	GLN
1	A	94	ASN
1	A	283	HIS
2	B	38	GLN
3	I	1	GLN
3	I	13	GLN
3	I	35	HIS
3	I	84	ASN
3	I	176	GLN
3	I	197	GLN
3	I	204	ASN
3	I	209	ASN
4	J	6	GLN
4	J	37	GLN
4	J	137	ASN
4	J	138	ASN
4	J	166	GLN
4	J	198	HIS
4	J	199	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

6 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	A	401	1	14,14,15	0.67	0	17,19,21	1.40	2 (11%)
5	NAG	A	402	1,5	14,14,15	0.53	0	17,19,21	1.25	2 (11%)
5	NAG	A	403	5	14,14,15	0.80	1 (7%)	17,19,21	1.14	1 (5%)
5	NAG	A	404	1	14,14,15	0.68	0	17,19,21	1.49	3 (17%)
6	MLI	I	301	-	0,6,6	0.00	-	0,7,7	0.00	-
6	MLI	J	301	-	0,6,6	0.00	-	0,7,7	0.00	-

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	402	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	403	5	-	0/6/23/26	0/1/1/1
5	NAG	A	404	1	-	0/6/23/26	0/1/1/1
6	MLI	I	301	-	-	0/0/4/4	0/0/0/0
6	MLI	J	301	-	-	0/0/4/4	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	403	NAG	C1-C2	2.10	1.55	1.52

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	404	NAG	C4-C3-C2	-2.88	106.80	111.02
5	A	404	NAG	O5-C1-C2	-2.03	108.72	111.52
5	A	402	NAG	C4-C3-C2	2.31	114.41	111.02
5	A	401	NAG	C4-C3-C2	2.48	114.65	111.02
5	A	403	NAG	C4-C3-C2	2.97	115.36	111.02
5	A	402	NAG	C1-O5-C5	3.07	116.41	112.19
5	A	404	NAG	C1-O5-C5	3.11	116.46	112.19
5	A	401	NAG	C1-O5-C5	3.18	116.57	112.19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	402	NAG	2	0
6	J	301	MLI	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	328/331 (99%)	0.25	8 (2%) 59 56	58, 74, 96, 145	0
2	B	171/179 (95%)	0.77	22 (12%) 3 3	63, 85, 136, 168	0
3	I	219/219 (100%)	0.65	28 (12%) 3 3	65, 102, 172, 195	0
4	J	212/214 (99%)	0.48	20 (9%) 8 6	60, 89, 149, 173	0
All	All	930/943 (98%)	0.49	78 (8%) 11 8	58, 82, 150, 195	0

All (78) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	195	GLY	8.3
1	A	9	PRO	7.4
3	I	196	THR	6.8
1	A	7	ALA	6.0
3	I	199	TYR	5.2
3	I	137	SER	5.2
1	A	8	ASP	4.7
4	J	129	THR	4.7
4	J	184	ALA	4.7
1	A	10	GLY	4.5
1	A	334	ILE	4.2
2	B	143	LYS	4.0
3	I	194	LEU	4.0
2	B	158	ASP	3.8
3	I	204	ASN	3.7
3	I	213	ASP	3.7
4	J	192	TYR	3.6
2	B	172	GLU	3.5
2	B	31	GLY	3.5
3	I	212	VAL	3.5
2	B	140	PHE	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	J	154	LEU	3.3
3	I	192	SER	3.2
3	I	203	VAL	3.1
2	B	173	ILE	3.1
2	B	29	GLU	3.1
4	J	183	LYS	3.0
4	J	181	LEU	3.0
3	I	139	GLY	3.0
2	B	168	LEU	3.0
2	B	139	GLU	2.9
2	B	164	GLU	2.9
3	I	138	GLY	2.9
3	I	211	LYS	2.8
2	B	160	PRO	2.7
3	I	16	ARG	2.7
4	J	125	LEU	2.7
4	J	130	ALA	2.7
3	I	163	ALA	2.6
3	I	143	LEU	2.5
2	B	147	ALA	2.5
4	J	187	GLU	2.5
4	J	178	THR	2.5
2	B	32	SER	2.5
4	J	128	GLY	2.5
3	I	190	PRO	2.4
4	J	131	SER	2.4
4	J	3	GLN	2.4
3	I	193	SER	2.4
2	B	113	SER	2.3
3	I	87	ARG	2.3
2	B	18	ILE	2.2
3	I	206	LYS	2.2
2	B	157	TYR	2.2
4	J	1	GLU	2.2
2	B	162	TYR	2.2
3	I	150	TYR	2.2
1	A	96	GLU	2.2
2	B	1	GLY	2.2
4	J	188	LYS	2.2
3	I	13	GLN	2.2
3	I	66	GLY	2.2
1	A	12	THR	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
4	J	127	SER	2.1
2	B	146	ASP	2.1
3	I	86	LEU	2.1
4	J	179	LEU	2.1
3	I	153	GLU	2.1
4	J	132	VAL	2.1
3	I	89	GLU	2.1
2	B	30	GLN	2.1
3	I	161	SER	2.1
4	J	157	GLY	2.0
2	B	159	TYR	2.0
3	I	159	TRP	2.0
4	J	106	ILE	2.0
1	A	194	THR	2.0
2	B	153	ARG	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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NAG	A	403	14/15	0.81	0.33	101,103,104,104	0
5	NAG	A	401	14/15	0.85	0.23	88,97,99,100	0
5	NAG	A	404	14/15	0.86	0.20	84,91,97,98	0
5	NAG	A	402	14/15	0.91	0.17	80,85,91,96	0
6	MLI	J	301	7/7	0.93	0.19	95,96,97,97	0
6	MLI	I	301	7/7	0.97	0.36	95,97,97,98	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.