



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 08:06 pm GMT

PDB ID : 5E64  
Title : Hemagglutinin-esterase-fusion protein structure of influenza D virus  
Authors : Song, H.; Qi, J.; Shi, Y.; Gao, G.F.  
Deposited on : 2015-10-09  
Resolution : 2.40 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

<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 : trunk28620  
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) : recalc28949

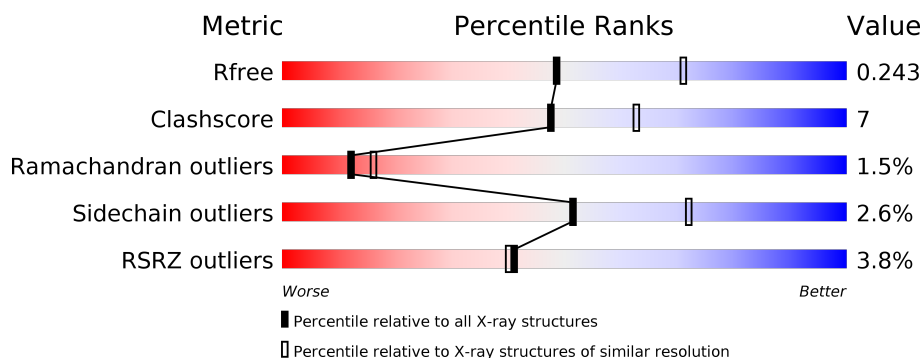
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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	427	<div> <div>86%</div> <div>13%</div> <div>•</div> </div>
1	C	427	<div> <div>%</div> <div>86%</div> <div>12%</div> <div>•</div> </div>
2	B	166	<div> <div>8%</div> <div>68%</div> <div>20%</div> <div>•</div> <div>10%</div> </div>
2	D	166	<div> <div>13%</div> <div>70%</div> <div>14%</div> <div>•</div> <div>•</div> <div>10%</div> </div>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 9653 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.

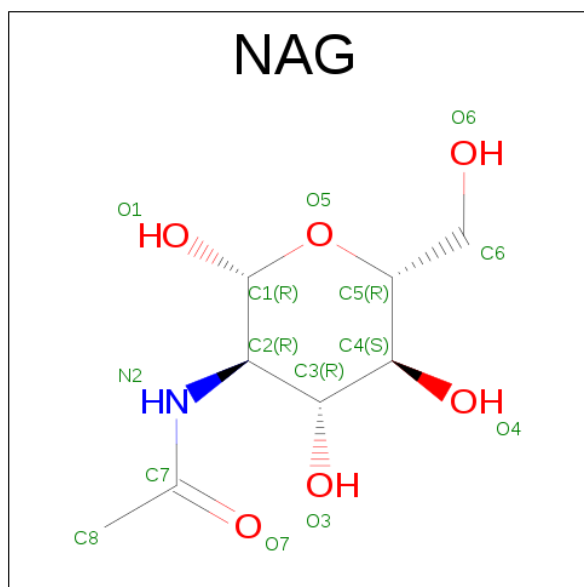
- Molecule 1 is a protein called Hemagglutinin-esterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	424	Total	C	N	O	S	0	0	0
			3249	2043	542	642	22			
1	C	421	Total	C	N	O	S	0	0	0
			3226	2028	537	639	22			

- Molecule 2 is a protein called Hemagglutinin-esterase.

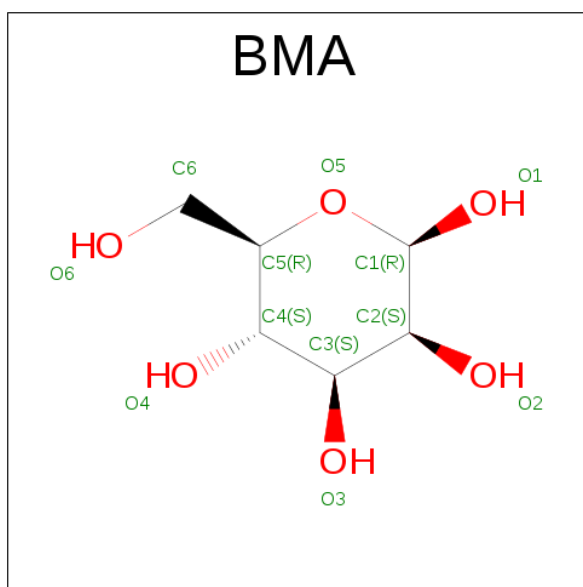
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	149	Total	C	N	O	S	0	0	0
			1110	695	194	217	4			
2	D	149	Total	C	N	O	S	0	0	0
			1110	695	194	217	4			

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



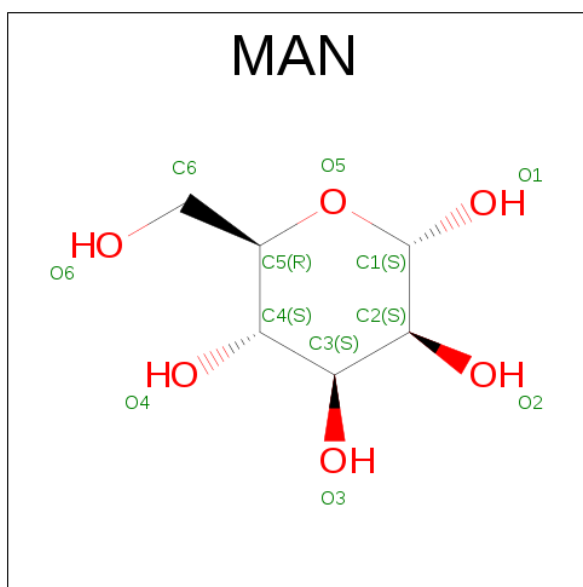
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1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

- Molecule 4 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
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		
4	C	1	Total	C	O	0	0
			11	6	5		

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



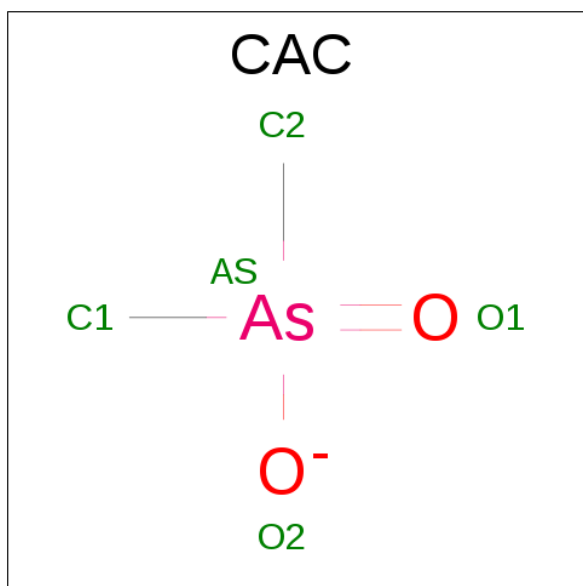
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			11	6	5		

*Continued on next page...*

Continued from previous page...

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

- Molecule 6 is CACODYLATE ION (three-letter code: CAC) (formula:  $\text{C}_2\text{H}_6\text{AsO}_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	As	C	O	0	0
			4	1	2	1		
6	C	1	Total	As	C	O	0	0
			4	1	2	1		

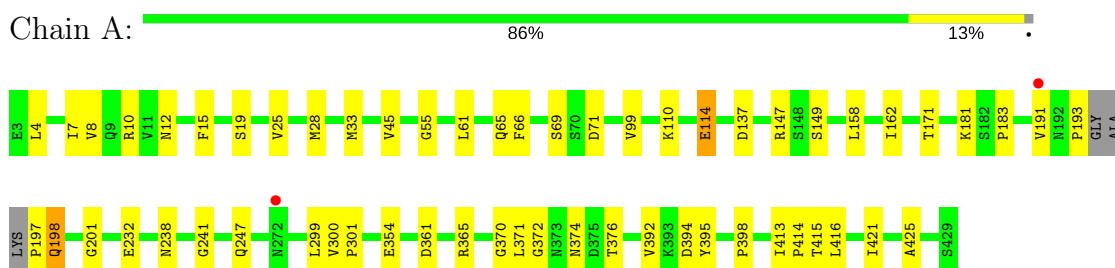
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	264	Total 264	O 264	0	0
7	B	58	Total 58	O 58	0	0
7	C	223	Total 223	O 223	0	0
7	D	21	Total 21	O 21	0	0

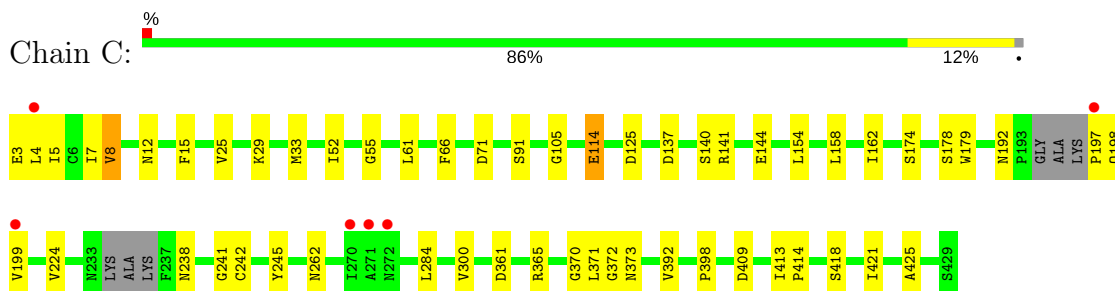
### 3 Residue-property plots

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.

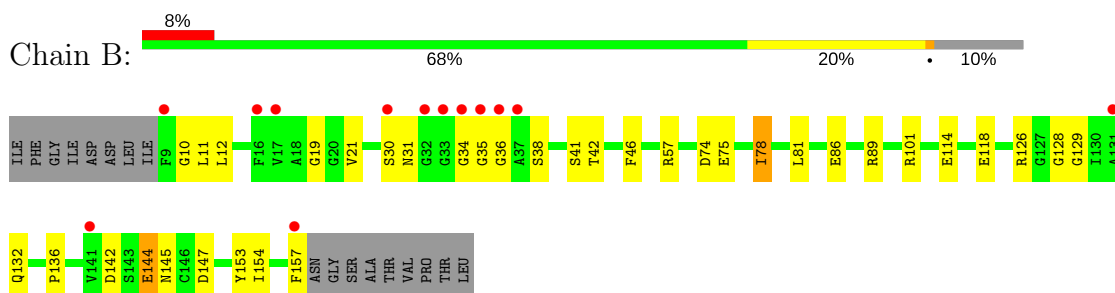
#### • Molecule 1: Hemagglutinin-esterase



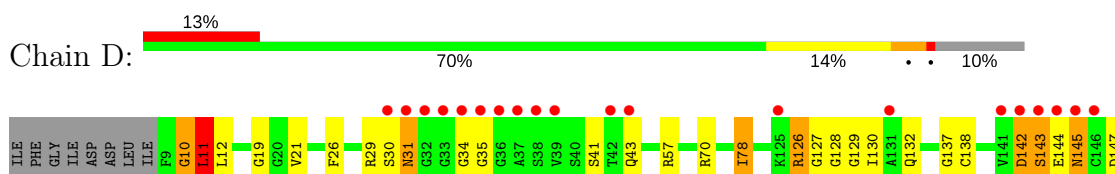
#### • Molecule 1: Hemagglutinin-esterase



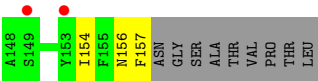
#### • Molecule 2: Hemagglutinin-esterase



#### • Molecule 2: Hemagglutinin-esterase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	165.22Å 165.22Å 165.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.70 – 2.40 49.82 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.70-2.40) 100.0 (49.82-2.40)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.55 (at 2.39Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.9_1692)	Depositor
R, $R_{free}$	0.201 , 0.245 0.201 , 0.243	Depositor DCC
$R_{free}$ test set	2963 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.3	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 41.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.031 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	9653	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.91% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: CAC, 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	A	0.36	0/3325	0.52	0/4507
1	C	0.34	0/3301	0.51	0/4475
2	B	0.34	0/1127	0.51	0/1515
2	D	0.36	0/1127	0.63	2/1515 (0.1%)
All	All	0.35	0/8880	0.53	2/12012 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	142	ASP	CB-CG-OD1	7.83	125.35	118.30
2	D	142	ASP	CB-CG-OD2	-5.55	113.31	118.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3249	0	3114	45	1
1	C	3226	0	3082	48	0
2	B	1110	0	1067	24	0
2	D	1110	0	1067	20	0
3	A	84	0	73	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	42	0	38	0	0
3	C	84	0	74	2	0
3	D	42	0	38	1	0
4	A	22	0	17	0	0
4	C	11	0	9	1	0
5	A	66	0	57	0	0
5	C	33	0	28	0	0
6	A	4	0	0	0	0
6	C	4	0	0	0	0
7	A	264	0	0	12	1
7	B	58	0	0	6	1
7	C	223	0	0	10	1
7	D	21	0	0	1	0
All	All	9653	0	8664	125	2

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 7.

All (125) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:193:PRO:C	1:A:197:PRO:HD2	1.48	1.30
1:C:192:ASN:H	1:C:199:VAL:HG23	1.19	1.07
1:C:192:ASN:CB	1:C:199:VAL:HG21	1.78	1.07
1:C:192:ASN:HB3	1:C:199:VAL:HG21	1.06	1.04
1:C:192:ASN:N	1:C:199:VAL:HG23	1.71	1.02
1:C:192:ASN:HB3	1:C:199:VAL:CG2	1.94	0.97
1:A:114:GLU:OE1	7:A:801:HOH:O	1.87	0.93
1:A:193:PRO:C	1:A:197:PRO:CD	2.38	0.90
1:C:192:ASN:CB	1:C:199:VAL:CG2	2.52	0.86
1:A:415:THR:O	2:B:101:ARG:NH1	2.10	0.85
1:C:192:ASN:N	1:C:199:VAL:CG2	2.40	0.85
2:B:10:GLY:O	7:B:801:HOH:O	1.97	0.81
1:C:192:ASN:H	1:C:199:VAL:CG2	1.93	0.81
1:A:247:GLN:HE22	1:A:299:LEU:H	1.31	0.79
1:C:162:ILE:HD11	1:C:300:VAL:HG21	1.68	0.74
2:B:126:ARG:NH1	2:B:153:TYR:O	2.20	0.74
2:B:30:SER:OG	2:B:31:ASN:N	2.21	0.73
1:A:33:MET:HE1	2:B:101:ARG:HB2	1.69	0.73
1:A:247:GLN:OE1	7:A:803:HOH:O	2.05	0.72
4:C:703:BMA:O3	7:C:802:HOH:O	2.05	0.72

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:714:NAG:O4	7:A:804:HOH:O	2.11	0.68
3:C:708:NAG:O4	7:C:803:HOH:O	2.11	0.67
1:A:137:ASP:OD2	7:A:806:HOH:O	2.13	0.67
1:C:55:GLY:HA2	1:C:114:GLU:HB3	1.75	0.66
2:B:118:GLU:OE2	7:B:802:HOH:O	2.12	0.66
2:B:114:GLU:OE1	7:B:803:HOH:O	2.14	0.65
1:A:19:SER:O	7:A:805:HOH:O	2.13	0.65
2:D:142:ASP:OD1	2:D:143:SER:N	2.30	0.64
1:A:4:LEU:HB3	2:B:30:SER:HB2	1.80	0.64
1:C:242:CYS:O	7:C:804:HOH:O	2.15	0.64
1:C:33:MET:HE2	1:C:414:PRO:HB2	1.79	0.63
1:C:144:GLU:OE1	7:C:805:HOH:O	2.15	0.63
1:C:371:LEU:N	1:C:372:GLY:HA3	2.15	0.62
1:C:114:GLU:HG2	7:C:801:HOH:O	1.99	0.60
2:B:57:ARG:NH2	7:B:806:HOH:O	2.37	0.58
1:C:425:ALA:O	7:C:807:HOH:O	2.17	0.58
1:A:371:LEU:N	1:A:372:GLY:HA3	2.20	0.57
1:A:247:GLN:NE2	1:A:299:LEU:H	2.02	0.57
2:D:156:ASN:O	2:D:157:PHE:HB2	2.04	0.57
2:D:126:ARG:HD3	2:D:154:ILE:HA	1.86	0.57
1:A:361:ASP:O	1:A:365:ARG:HG3	2.07	0.55
2:D:127:GLY:O	2:D:130:ILE:N	2.40	0.55
1:C:137:ASP:O	1:C:141:ARG:NH2	2.40	0.55
1:A:25:VAL:HG13	1:A:421:ILE:HG23	1.89	0.54
1:A:55:GLY:HA2	1:A:114:GLU:HG3	1.89	0.54
1:C:8:VAL:HG13	2:D:26:PHE:HB2	1.90	0.54
1:A:183:PRO:HB3	7:A:933:HOH:O	2.07	0.54
1:A:33:MET:HG3	1:A:416:LEU:HG	1.88	0.54
1:A:7:ILE:O	2:B:10:GLY:HA2	2.09	0.53
1:C:25:VAL:HG13	1:C:421:ILE:HG23	1.90	0.52
1:A:374:ASN:ND2	7:A:817:HOH:O	2.41	0.52
1:A:4:LEU:H	2:B:30:SER:HB3	1.76	0.50
1:C:4:LEU:O	2:D:30:SER:HB2	2.11	0.50
3:C:702:NAG:O6	3:C:708:NAG:H81	2.10	0.50
1:A:395:TYR:CD2	2:B:86:GLU:HG2	2.46	0.50
1:C:125:ASP:HB3	1:C:174:SER:O	2.11	0.50
2:B:31:ASN:OD1	2:B:36:GLY:N	2.33	0.49
2:D:10:GLY:O	2:D:12:LEU:N	2.36	0.49
3:D:701:NAG:H61	3:D:702:NAG:C7	2.43	0.49
1:A:110:LYS:NZ	1:A:149:SER:O	2.37	0.48
2:D:34:GLY:HA2	2:D:35:GLY:HA2	1.66	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:361:ASP:O	1:C:365:ARG:HG3	2.14	0.48
1:A:197:PRO:C	1:A:198:GLN:HG2	2.34	0.48
1:C:7:ILE:O	2:D:10:GLY:HA2	2.13	0.48
1:C:71:ASP:HA	1:C:365:ARG:HG2	1.95	0.48
1:A:398:PRO:HB2	2:B:78:ILE:O	2.14	0.48
2:D:10:GLY:HA3	2:D:138:CYS:SG	2.53	0.48
1:A:370:GLY:C	1:A:372:GLY:HA3	2.33	0.48
1:C:162:ILE:HG23	7:C:877:HOH:O	2.14	0.48
1:C:3:GLU:HG3	2:D:30:SER:HB3	1.96	0.48
1:A:33:MET:CE	1:A:414:PRO:HB2	2.45	0.47
1:C:197:PRO:HA	1:C:198:GLN:HA	1.42	0.47
1:A:232:GLU:HB2	7:A:822:HOH:O	2.14	0.47
1:C:398:PRO:HB2	2:D:78:ILE:O	2.15	0.46
1:C:262:ASN:ND2	7:C:816:HOH:O	2.38	0.46
2:B:142:ASP:HB3	2:B:144:GLU:CD	2.36	0.46
1:A:114:GLU:HB3	7:A:870:HOH:O	2.14	0.46
1:A:162:ILE:HD11	1:A:300:VAL:HG21	1.96	0.46
1:C:409:ASP:HB3	7:C:856:HOH:O	2.15	0.46
1:C:61:LEU:HA	1:C:66:PHE:CD1	2.51	0.46
1:C:4:LEU:H	2:D:30:SER:CB	2.29	0.46
2:B:128:GLY:HA2	2:B:129:GLY:HA2	1.51	0.46
1:A:71:ASP:HA	1:A:365:ARG:HG2	1.98	0.45
2:D:128:GLY:HA2	2:D:129:GLY:HA2	1.48	0.45
1:C:137:ASP:HA	1:C:141:ARG:HH22	1.82	0.45
2:D:30:SER:OG	2:D:31:ASN:ND2	2.50	0.45
1:A:394:ASP:OD1	7:A:807:HOH:O	2.21	0.45
1:A:61:LEU:HA	1:A:66:PHE:CD1	2.52	0.45
2:B:34:GLY:HA2	2:B:35:GLY:HA2	1.66	0.45
2:D:156:ASN:O	2:D:157:PHE:CB	2.65	0.45
1:C:12:ASN:H	1:C:15:PHE:HD1	1.62	0.45
1:C:238:ASN:HA	1:C:241:GLY:O	2.16	0.45
1:A:10:ARG:HD3	7:A:845:HOH:O	2.17	0.44
2:B:89:ARG:HD3	7:B:854:HOH:O	2.16	0.44
1:C:140:SER:HB2	1:C:141:ARG:NH2	2.32	0.44
1:A:12:ASN:H	1:A:15:PHE:HD1	1.66	0.44
1:A:238:ASN:HA	1:A:241:GLY:O	2.18	0.44
1:A:65:GLN:HB2	1:A:69:SER:HB3	2.00	0.44
1:A:191:VAL:HG22	1:A:201:GLY:HA3	2.01	0.43
1:C:418:SER:HA	2:D:57:ARG:HD3	2.00	0.43
2:B:74:ASP:OD1	2:B:75:GLU:N	2.52	0.43
1:C:224:VAL:HG21	1:C:284:LEU:HD12	2.00	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:70:ARG:NH2	7:D:802:HOH:O	2.24	0.43
1:A:354:GLU:O	1:C:29:LYS:HE2	2.18	0.43
1:C:370:GLY:C	1:C:372:GLY:HA3	2.39	0.43
7:A:805:HOH:O	2:B:101:ARG:NE	2.48	0.43
1:A:395:TYR:CG	2:B:81:LEU:HD21	2.54	0.43
1:A:28:MET:SD	3:A:713:NAG:H5	2.58	0.42
1:A:300:VAL:HA	1:A:301:PRO:HD3	1.91	0.42
1:C:413:ILE:HA	1:C:414:PRO:HD3	1.88	0.42
2:B:75:GLU:OE2	7:B:804:HOH:O	2.22	0.42
2:D:11:LEU:HD22	2:D:137:GLY:HA3	2.02	0.42
1:C:245:TYR:O	1:C:262:ASN:HB2	2.20	0.42
1:A:413:ILE:HA	1:A:414:PRO:HD3	1.94	0.41
1:A:197:PRO:HB2	1:A:198:GLN:H	1.72	0.41
1:C:178:SER:HB2	1:C:179:TRP:H	1.72	0.41
1:C:3:GLU:OE2	2:D:29:ARG:HB3	2.21	0.41
1:C:373:ASN:HB3	7:C:863:HOH:O	2.20	0.41
1:A:181:LYS:HD3	1:A:181:LYS:HA	1.95	0.41
2:B:126:ARG:NH2	2:B:154:ILE:O	2.55	0.40
1:A:45:VAL:O	1:A:376:THR:HB	2.21	0.40
1:C:52:ILE:CD1	1:C:105:GLY:HA3	2.51	0.40
1:C:3:GLU:HG2	1:C:5:ILE:HG12	2.03	0.40
2:B:12:LEU:HG	2:B:136:PRO:HG2	2.04	0.40
1:C:91:SER:HB3	1:C:154:LEU:HD22	2.03	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:1005:HOH:O	7:C:1000:HOH:O[3_544]	2.12	0.08
1:A:425:ALA:O	7:B:802:HOH:O[5_555]	2.12	0.08

## 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	420/427 (98%)	401 (96%)	19 (4%)	0	100	100
1	C	415/427 (97%)	398 (96%)	17 (4%)	0	100	100
2	B	147/166 (89%)	120 (82%)	20 (14%)	7 (5%)	2	1
2	D	147/166 (89%)	123 (84%)	14 (10%)	10 (7%)	1	0
All	All	1129/1186 (95%)	1042 (92%)	70 (6%)	17 (2%)	12	16

All (17) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	11	LEU
2	D	43	GLN
2	B	147	ASP
2	D	31	ASN
2	D	41	SER
2	D	145	ASN
2	B	21	VAL
2	B	41	SER
2	B	46	PHE
2	B	19	GLY
2	B	132	GLN
2	D	132	GLN
2	D	143	SER
2	B	38	SER
2	D	21	VAL
2	D	10	GLY
2	D	19	GLY

### 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	360/361 (100%)	352 (98%)	8 (2%)	57	76
1	C	358/361 (99%)	354 (99%)	4 (1%)	78	90

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	109/123 (89%)	103 (94%)	6 (6%)	25	40
2	D	109/123 (89%)	103 (94%)	6 (6%)	25	40
All	All	936/968 (97%)	912 (97%)	24 (3%)	51	72

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	8	VAL
1	A	99	VAL
1	A	114	GLU
1	A	147	ARG
1	A	158	LEU
1	A	171	THR
1	A	198	GLN
1	A	392	VAL
2	B	11	LEU
2	B	42	THR
2	B	78	ILE
2	B	144	GLU
2	B	145	ASN
2	B	157	PHE
1	C	8	VAL
1	C	114	GLU
1	C	158	LEU
1	C	392	VAL
2	D	11	LEU
2	D	78	ILE
2	D	126	ARG
2	D	144	GLU
2	D	145	ASN
2	D	147	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	198	GLN
1	A	247	GLN
2	B	43	GLN
2	D	31	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

32 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	$\# Z  > 2$
3	NAG	A	701	1,3	14,14,15	0.39	0	15,19,21	1.14	2 (13%)
3	NAG	A	702	3,4	14,14,15	0.40	0	15,19,21	1.21	2 (13%)
4	BMA	A	703	3,5	11,11,12	0.45	0	13,15,17	0.77	0
5	MAN	A	704	5,4	11,11,12	0.30	0	13,15,17	0.75	0
5	MAN	A	705	5	11,11,12	0.36	0	13,15,17	0.64	0
5	MAN	A	706	5	11,11,12	0.32	0	13,15,17	0.90	0
3	NAG	A	707	1,3	14,14,15	0.37	0	15,19,21	0.79	0
3	NAG	A	708	3,4	14,14,15	0.43	0	15,19,21	1.62	5 (33%)
4	BMA	A	709	3,5	11,11,12	0.48	0	13,15,17	2.67	6 (46%)
5	MAN	A	710	5,4	11,11,12	0.37	0	13,15,17	1.96	2 (15%)
5	MAN	A	711	5	11,11,12	0.30	0	13,15,17	0.76	0
5	MAN	A	712	4	11,11,12	0.52	0	13,15,17	0.96	1 (7%)
3	NAG	A	713	1,3	14,14,15	0.33	0	15,19,21	1.21	1 (6%)
3	NAG	A	714	3	14,14,15	0.28	0	15,19,21	1.09	1 (6%)
6	CAC	A	715	-	0,3,4	0.00	-	0,3,6	0.00	-
3	NAG	B	701	3,2	14,14,15	0.37	0	15,19,21	1.06	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	B	702	3	14,14,15	0.30	0	15,19,21	1.27	1 (6%)
3	NAG	B	703	2	14,14,15	0.29	0	15,19,21	1.10	1 (6%)
3	NAG	C	701	1,3	14,14,15	0.32	0	15,19,21	0.97	1 (6%)
3	NAG	C	702	3,4	14,14,15	0.63	0	15,19,21	3.12	3 (20%)
4	BMA	C	703	3,5	11,11,12	0.79	0	13,15,17	3.15	2 (15%)
5	MAN	C	704	5,4	11,11,12	0.31	0	13,15,17	1.29	2 (15%)
5	MAN	C	705	5	11,11,12	0.29	0	13,15,17	0.77	0
5	MAN	C	706	5	11,11,12	0.42	0	13,15,17	0.61	0
3	NAG	C	707	1,3	14,14,15	0.39	0	15,19,21	1.04	1 (6%)
3	NAG	C	708	3	14,14,15	0.35	0	15,19,21	0.76	0
3	NAG	C	709	1,3	14,14,15	0.35	0	15,19,21	1.49	1 (6%)
3	NAG	C	710	3	14,14,15	0.29	0	15,19,21	1.03	1 (6%)
6	CAC	C	711	-	0,3,4	0.00	-	0,3,6	0.00	-
3	NAG	D	701	3,2	14,14,15	0.29	0	15,19,21	0.80	1 (6%)
3	NAG	D	702	3	14,14,15	0.30	0	15,19,21	1.04	1 (6%)
3	NAG	D	703	2	14,14,15	0.32	0	15,19,21	1.06	1 (6%)

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
3	NAG	A	701	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	702	3,4	-	0/6/23/26	0/1/1/1
4	BMA	A	703	3,5	-	0/2/19/22	0/1/1/1
5	MAN	A	704	5,4	-	0/2/19/22	0/1/1/1
5	MAN	A	705	5	-	0/2/19/22	0/1/1/1
5	MAN	A	706	5	-	0/2/19/22	0/1/1/1
3	NAG	A	707	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	708	3,4	-	0/6/23/26	0/1/1/1
4	BMA	A	709	3,5	-	0/2/19/22	0/1/1/1
5	MAN	A	710	5,4	-	0/2/19/22	0/1/1/1
5	MAN	A	711	5	-	0/2/19/22	0/1/1/1
5	MAN	A	712	4	-	0/2/19/22	0/1/1/1
3	NAG	A	713	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	714	3	-	0/6/23/26	0/1/1/1
6	CAC	A	715	-	-	0/0/0/0	0/0/0/0
3	NAG	B	701	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	702	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
3	NAG	B	703	2	-	0/6/23/26	0/1/1/1
3	NAG	C	701	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	702	3,4	-	0/6/23/26	0/1/1/1
4	BMA	C	703	3,5	-	0/2/19/22	0/1/1/1
5	MAN	C	704	5,4	-	0/2/19/22	0/1/1/1
5	MAN	C	705	5	-	0/2/19/22	0/1/1/1
5	MAN	C	706	5	-	0/2/19/22	0/1/1/1
3	NAG	C	707	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	708	3	-	0/6/23/26	0/1/1/1
3	NAG	C	709	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	710	3	-	0/6/23/26	0/1/1/1
6	CAC	C	711	-	-	0/0/0/0	0/0/0/0
3	NAG	D	701	3,2	-	0/6/23/26	0/1/1/1
3	NAG	D	702	3	-	0/6/23/26	0/1/1/1
3	NAG	D	703	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	702	NAG	O4-C4-C3	-11.13	86.14	110.36
4	C	703	BMA	O5-C1-C2	-10.46	94.39	110.79
4	A	709	BMA	C2-C3-C4	-4.24	103.47	110.88
4	A	709	BMA	O3-C3-C2	-3.46	103.73	110.02
3	A	708	NAG	O3-C3-C2	-2.98	103.00	109.39
3	A	708	NAG	O4-C4-C5	-2.84	102.14	109.28
3	A	702	NAG	O4-C4-C5	-2.73	102.41	109.28
3	A	708	NAG	O5-C1-C2	-2.53	107.95	111.47
4	A	709	BMA	O6-C6-C5	-2.44	103.12	111.34
4	A	709	BMA	O2-C2-C3	-2.26	105.73	110.17
5	A	712	MAN	C1-O5-C5	-2.11	109.26	112.17
3	A	701	NAG	C2-N2-C7	-2.06	119.94	122.94
4	A	709	BMA	C1-O5-C5	-2.05	109.34	112.17
3	C	702	NAG	C3-C4-C5	2.01	113.75	110.22
3	D	701	NAG	C1-O5-C5	2.01	114.93	112.17
3	A	702	NAG	C1-O5-C5	2.02	114.95	112.17
3	A	708	NAG	O4-C4-C3	2.06	114.83	110.36
5	C	704	MAN	C1-C2-C3	2.08	112.29	109.65
3	A	708	NAG	C4-C3-C2	2.12	114.12	111.02
3	C	710	NAG	C1-O5-C5	2.26	115.28	112.17
3	A	701	NAG	C1-O5-C5	2.48	115.59	112.17
5	C	704	MAN	C1-O5-C5	2.77	115.98	112.17

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	701	NAG	C1-O5-C5	2.79	116.01	112.17
3	C	707	NAG	C1-O5-C5	2.83	116.07	112.17
4	C	703	BMA	O2-C2-C1	2.87	115.02	109.18
5	A	710	MAN	C1-C2-C3	3.00	113.46	109.65
3	B	702	NAG	C1-O5-C5	3.01	116.32	112.17
3	B	701	NAG	C1-O5-C5	3.13	116.48	112.17
3	C	702	NAG	C1-O5-C5	3.21	116.59	112.17
3	B	703	NAG	C1-O5-C5	3.22	116.61	112.17
3	A	714	NAG	C1-O5-C5	3.24	116.64	112.17
3	D	703	NAG	C1-O5-C5	3.28	116.69	112.17
3	A	713	NAG	C1-O5-C5	3.51	117.00	112.17
3	D	702	NAG	C1-O5-C5	3.53	117.03	112.17
3	C	709	NAG	C1-O5-C5	5.10	119.19	112.17
5	A	710	MAN	C1-O5-C5	5.16	119.28	112.17
4	A	709	BMA	C1-C2-C3	5.99	117.25	109.65

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	713	NAG	1	0
3	A	714	NAG	1	0
3	C	702	NAG	1	0
4	C	703	BMA	1	0
3	C	708	NAG	2	0
3	D	701	NAG	1	0
3	D	702	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	A	424/427 (99%)	-0.54	2 (0%) 90 89	14, 25, 48, 95	0
1	C	421/427 (98%)	-0.44	6 (1%) 75 74	13, 27, 55, 83	0
2	B	149/166 (89%)	0.17	13 (8%) 11 10	10, 33, 98, 123	0
2	D	149/166 (89%)	0.68	22 (14%) 3 2	22, 47, 122, 148	0
All	All	1143/1186 (96%)	-0.25	43 (3%) 41 40	10, 29, 77, 148	0

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	32	GLY	10.5
2	D	37	ALA	10.2
2	D	33	GLY	8.7
2	D	35	GLY	8.5
2	D	34	GLY	7.1
2	D	145	ASN	7.1
2	B	37	ALA	7.0
2	D	38	SER	6.7
2	B	34	GLY	6.2
2	D	30	SER	5.3
2	B	32	GLY	5.0
2	B	33	GLY	4.7
2	D	141	VAL	4.7
2	D	143	SER	4.4
2	B	141	VAL	4.3
2	D	36	GLY	4.3
2	B	35	GLY	4.2
2	B	36	GLY	3.7
1	C	197	PRO	3.6
2	B	9	PHE	3.3
2	D	43	GLN	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	146	CYS	3.2
1	C	272	ASN	3.1
1	C	199	VAL	3.0
2	B	30	SER	2.9
2	D	149	SER	2.9
2	D	142	ASP	2.9
2	B	16	PHE	2.8
1	A	191	VAL	2.7
1	C	271	ALA	2.6
2	D	131	ALA	2.6
2	D	42	THR	2.6
1	C	270	ILE	2.6
1	C	4	LEU	2.5
2	B	157	PHE	2.5
2	D	125	LYS	2.4
2	D	39	VAL	2.4
2	B	17	VAL	2.4
2	D	31	ASN	2.3
2	D	153	TYR	2.2
2	D	144	GLU	2.1
2	B	131	ALA	2.1
1	A	272	ASN	2.0

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

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

## 6.3 Carbohydrates [i](#)

There are no 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. 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( $\text{\AA}^2$ )	Q<0.9
5	MAN	C	706	11/12	0.91	0.14	1.62	27,31,33,34	0
6	CAC	A	715	4/5	0.99	0.13	0.43	35,37,40,54	0
3	NAG	B	702	14/15	0.95	0.13	0.26	29,33,43,44	0
3	NAG	D	702	14/15	0.90	0.15	0.17	33,45,51,54	0
3	NAG	B	701	14/15	0.96	0.12	-0.18	16,20,22,23	0
6	CAC	C	711	4/5	0.99	0.11	-0.51	26,30,38,49	0
3	NAG	C	709	14/15	0.94	0.13	-0.68	38,44,53,54	0
5	MAN	A	706	11/12	0.93	0.10	-0.72	21,25,29,34	0
3	NAG	D	701	14/15	0.97	0.11	-1.06	23,27,29,30	0
3	NAG	A	713	14/15	0.93	0.10	-1.16	41,46,48,49	0
3	NAG	C	710	14/15	0.87	0.17	-	55,61,78,80	0
3	NAG	A	707	14/15	0.98	0.09	-	19,25,28,32	0
3	NAG	C	702	14/15	0.91	0.16	-	24,29,34,35	0
5	MAN	C	705	11/12	0.94	0.09	-	34,39,46,53	0
3	NAG	A	701	14/15	0.96	0.09	-	19,24,31,36	0
3	NAG	A	708	14/15	0.94	0.12	-	21,27,33,33	0
5	MAN	A	710	11/12	0.94	0.10	-	36,37,46,47	0
3	NAG	A	702	14/15	0.97	0.11	-	20,25,27,29	0
3	NAG	C	708	14/15	0.92	0.14	-	30,37,41,43	0
3	NAG	C	701	14/15	0.95	0.12	-	22,26,30,31	0
5	MAN	A	711	11/12	0.92	0.22	-	46,54,59,65	0
3	NAG	D	703	14/15	0.76	0.35	-	86,91,97,101	0
5	MAN	C	704	11/12	0.98	0.08	-	19,25,33,35	0
3	NAG	A	714	14/15	0.91	0.18	-	51,54,59,60	0
5	MAN	A	712	11/12	0.92	0.11	-	29,33,36,42	0
4	BMA	C	703	11/12	0.96	0.10	-	31,32,38,40	0
5	MAN	A	705	11/12	0.96	0.09	-	20,26,28,28	0
4	BMA	A	703	11/12	0.96	0.13	-	20,23,26,31	0
5	MAN	A	704	11/12	0.97	0.09	-	19,21,22,23	0
4	BMA	A	709	11/12	0.91	0.19	-	34,39,47,49	0
3	NAG	B	703	14/15	0.82	0.24	-	45,58,64,65	0
3	NAG	C	707	14/15	0.95	0.12	-	21,24,29,36	0

## 6.5 Other polymers

There are no such residues in this entry.