



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 24, 2023 – 01:57 AM EDT

PDB ID : 5T08  
Title : Crystal structure of H6 hemagglutinin G225D mutant from Taiwan (2013)  
H6N1 influenza virus  
Authors : Wilson, I.A.; Tzarum, N.; Zhu, X.  
Deposited on : 2016-08-15  
Resolution : 2.19 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

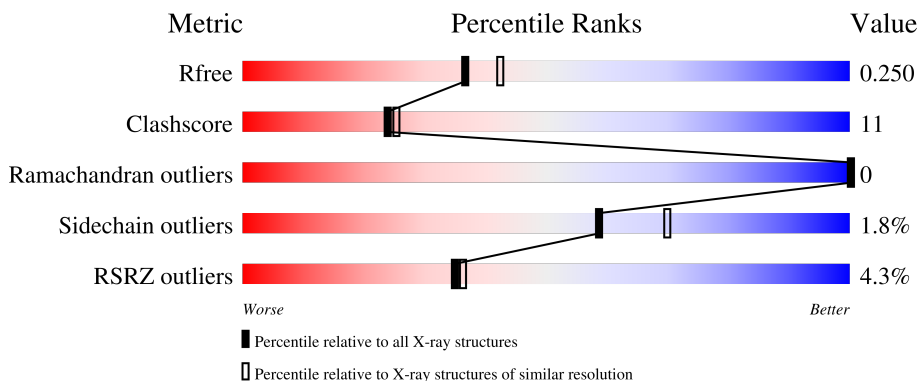
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.19 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.



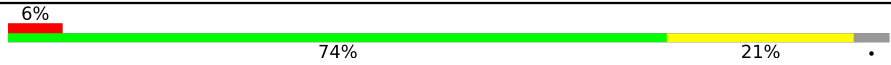
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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

Mol	Chain	Length	Quality of chain
1	A	333	 4% 72% 26% ..
1	C	333	 5% 78% 19% ..
1	E	333	 3% 70% 26% ..
2	B	180	 2% 75% 21% ..
2	D	180	 8% 82% 14% .

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	180	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into three segments: a small red segment on the left labeled '6%', a large green segment in the middle labeled '74%', and a yellow segment on the right labeled '21%'. A small grey dot is visible at the far right end of the bar.</p>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12376 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	330	Total 2604	C 1650	N 443	O 498	S 13	0	0	0
1	C	325	Total 2566	C 1627	N 436	O 490	S 13	0	0	0
1	E	325	Total 2569	C 1629	N 437	O 490	S 13	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	225	ASP	GLY	engineered mutation	UNP A0A0J9X268
C	225	ASP	GLY	engineered mutation	UNP A0A0J9X268
E	225	ASP	GLY	engineered mutation	UNP A0A0J9X268

- Molecule 2 is a protein called Hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	173	Total 1393	C 868	N 244	O 274	S 7	0	0	0
2	D	173	Total 1393	C 868	N 244	O 274	S 7	0	0	0
2	F	172	Total 1385	C 862	N 243	O 273	S 7	0	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

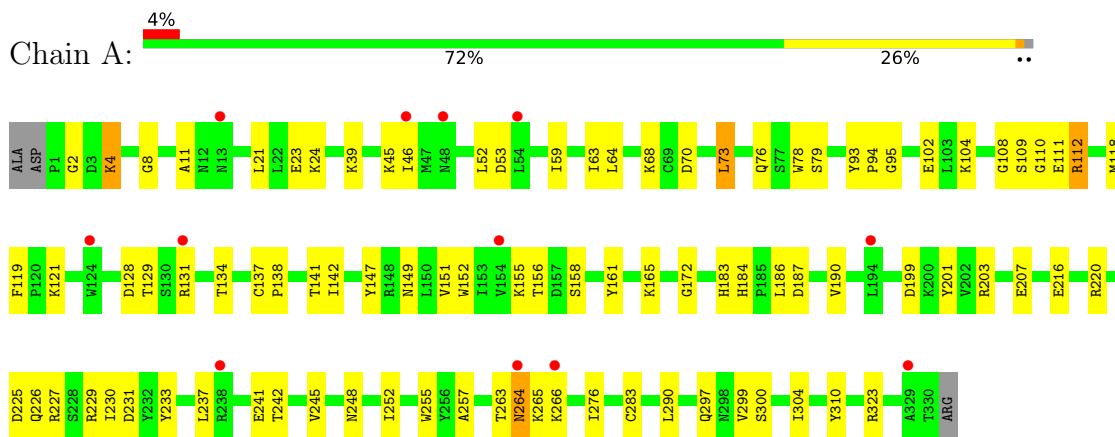
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	71	71	71	0	0
4	B	64	64	64	0	0
4	C	85	85	85	0	0
4	D	43	43	43	0	0
4	E	89	89	89	0	0
4	F	58	58	58	0	0

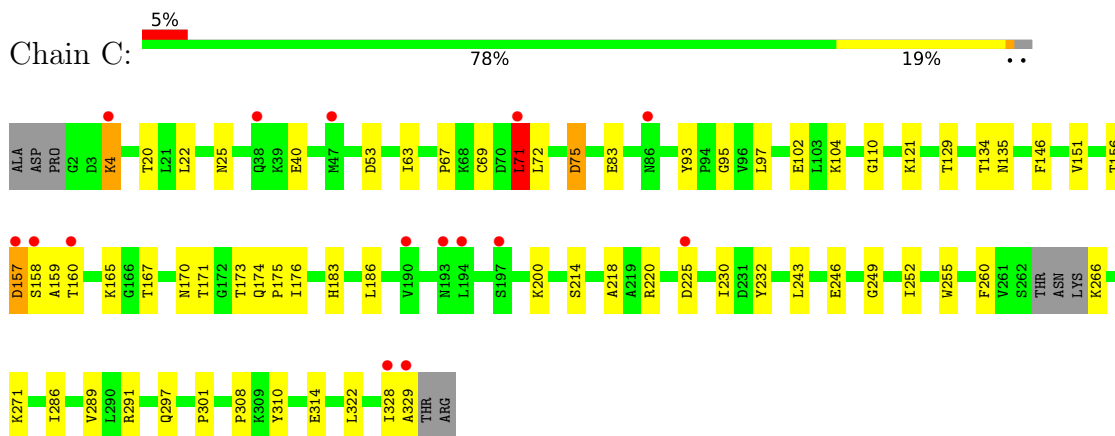
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

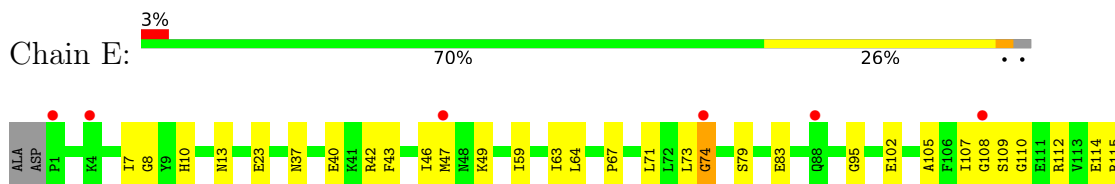
- Molecule 1: Hemagglutinin

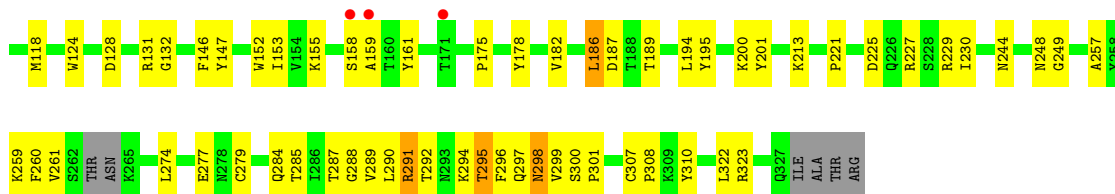


- Molecule 1: Hemagglutinin

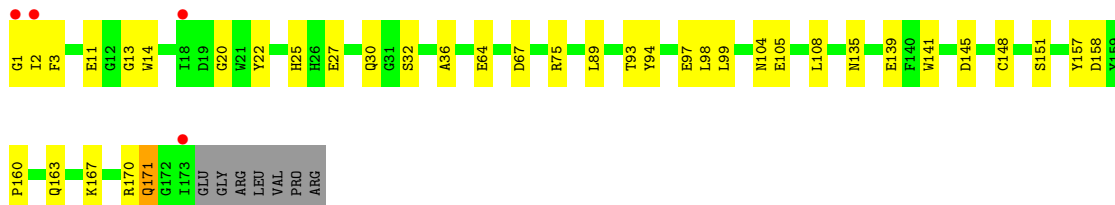
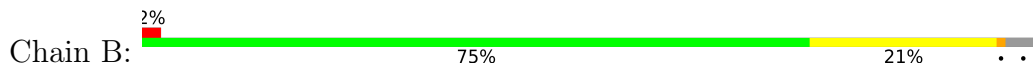


- Molecule 1: Hemagglutinin

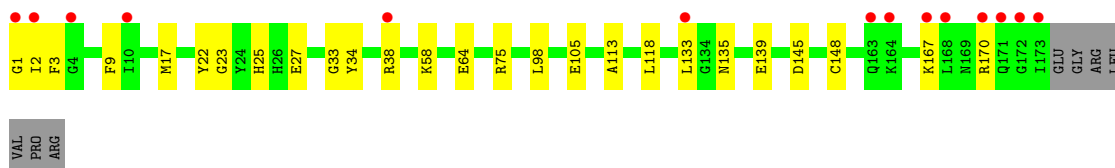
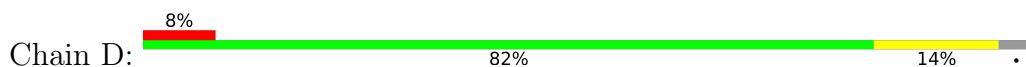




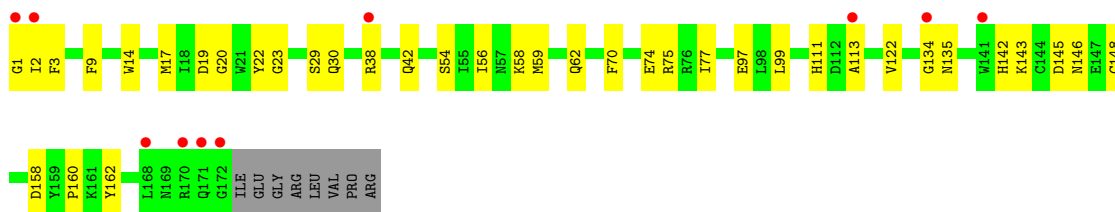
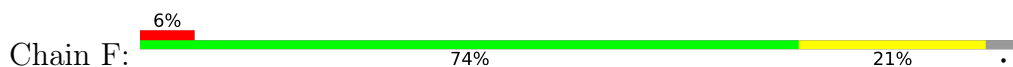
- Molecule 2: Hemagglutinin HA2 chain



- Molecule 2: Hemagglutinin HA2 chain



- Molecule 2: Hemagglutinin HA2 chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	185.06Å 98.36Å 133.24Å 90.00° 126.48° 90.00°	Depositor
Resolution (Å)	49.34 – 2.19 49.34 – 2.19	Depositor EDS
% Data completeness (in resolution range)	97.7 (49.34-2.19) 93.4 (49.34-2.19)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.04 (at 2.18Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.215 , 0.250 0.216 , 0.250	Depositor DCC
$R_{free}$ test set	4800 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.2	Xtrriage
Anisotropy	0.685	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 47.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12376	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.49% 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: 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	0.56	3/2666 (0.1%)	0.74	3/3628 (0.1%)
1	C	0.60	4/2626 (0.2%)	0.73	4/3572 (0.1%)
1	E	0.55	0/2630	0.74	4/3576 (0.1%)
2	B	0.54	0/1421	0.66	1/1913 (0.1%)
2	D	0.55	0/1421	0.68	1/1913 (0.1%)
2	F	0.52	0/1413	0.69	1/1902 (0.1%)
All	All	0.56	7/12177 (0.1%)	0.71	14/16504 (0.1%)

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	C	0	1
2	F	0	1
All	All	0	2

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	104	LYS	CE-NZ	6.55	1.65	1.49
1	C	4	LYS	CE-NZ	6.20	1.64	1.49
1	A	4	LYS	CD-CE	6.20	1.66	1.51
1	C	4	LYS	CD-CE	6.15	1.66	1.51
1	C	104	LYS	CE-NZ	6.07	1.64	1.49
1	A	4	LYS	CG-CD	5.12	1.69	1.52
1	C	4	LYS	CG-CD	5.07	1.69	1.52

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	4	LYS	CD-CE-NZ	9.61	133.81	111.70
2	D	1	GLY	N-CA-C	-8.55	91.72	113.10
1	C	4	LYS	CD-CE-NZ	8.54	131.35	111.70
2	B	1	GLY	N-CA-C	-8.26	92.46	113.10
2	F	1	GLY	N-CA-C	-8.13	92.77	113.10
1	E	74	GLY	N-CA-C	7.47	131.77	113.10
1	C	71	LEU	CA-CB-CG	6.64	130.58	115.30
1	E	186	LEU	CA-CB-CG	6.15	129.44	115.30
1	C	71	LEU	C-N-CA	6.03	136.78	121.70
1	A	108	GLY	N-CA-C	-5.82	98.55	113.10
1	A	172	GLY	N-CA-C	5.78	127.55	113.10
1	E	195	TYR	C-N-CA	-5.39	110.97	122.30
1	E	71	LEU	CA-CB-CG	5.14	127.12	115.30
1	C	71	LEU	CA-C-N	5.05	128.31	117.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	71	LEU	Peptide
2	F	134	GLY	Peptide

## 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	2604	0	2552	81	0
1	C	2566	0	2510	60	0
1	E	2569	0	2515	69	0
2	B	1393	0	1309	28	0
2	D	1393	0	1309	27	0
2	F	1385	0	1298	28	0
3	A	28	0	26	0	0
3	E	28	0	26	0	0
4	A	71	0	0	4	0
4	B	64	0	0	1	0
4	C	85	0	0	10	0
4	D	43	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	89	0	0	3	0
4	F	58	0	0	3	0
All	All	12376	0	11545	261	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 (261) 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:C:156:THR:O	1:C:159:ALA:CB	1.72	1.37
1:C:156:THR:O	1:C:159:ALA:HB1	1.32	1.14
1:E:307:CYS:O	2:F:62:GLN:NE2	1.95	0.98
1:C:4:LYS:NZ	2:D:25:HIS:HE1	1.61	0.96
2:B:30:GLN:HE21	2:B:145:ASP:HB2	1.34	0.93
1:C:156:THR:O	1:C:159:ALA:HB2	1.70	0.90
1:A:111:GLU:HB2	1:A:263:THR:HB	1.55	0.88
1:C:157:ASP:C	1:C:159:ALA:HB2	1.94	0.86
1:C:266:LYS:NZ	2:D:64:GLU:OE1	2.10	0.83
1:C:158:SER:N	1:C:159:ALA:HB2	1.93	0.83
1:A:290:LEU:HD11	1:A:297:GLN:HG3	1.61	0.82
1:A:263:THR:OG1	1:A:265:LYS:HG2	1.78	0.82
2:B:30:GLN:NE2	2:B:145:ASP:HB2	1.95	0.82
2:F:17:MET:SD	2:F:23:GLY:HA3	2.20	0.81
1:A:297:GLN:NE2	1:A:300:SER:H	1.81	0.78
2:F:142:HIS:O	4:F:201:HOH:O	2.02	0.77
1:C:53:ASP:OD2	4:C:401:HOH:O	2.03	0.76
1:C:171:THR:N	4:C:405:HOH:O	2.18	0.76
1:A:119:PHE:O	4:A:501:HOH:O	2.02	0.76
1:A:297:GLN:HE21	1:A:299:VAL:N	1.83	0.75
1:A:242:THR:OG1	4:A:502:HOH:O	2.04	0.75
1:C:4:LYS:HZ1	2:D:25:HIS:HE1	1.32	0.75
1:E:285:THR:HG22	1:E:287:THR:H	1.51	0.73
1:E:221:PRO:O	1:E:229:ARG:NH2	2.22	0.72
1:C:286:ILE:O	4:C:402:HOH:O	2.07	0.72
2:F:143:LYS:O	4:F:202:HOH:O	2.07	0.71
1:C:4:LYS:NZ	2:D:25:HIS:CE1	2.52	0.70
1:A:156:THR:HG23	1:A:158:SER:H	1.56	0.70
1:E:186:LEU:HD13	1:E:227:ARG:HB2	1.72	0.70
1:C:160:THR:O	4:C:403:HOH:O	2.09	0.70
1:A:68:LYS:HE2	1:A:138:PRO:HG2	1.72	0.70

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:290:LEU:CD1	1:A:297:GLN:HG3	2.22	0.70
1:A:128:ASP:OD2	1:A:131:ARG:HG2	1.91	0.69
1:E:109:SER:HB3	1:E:110:GLY:O	1.90	0.69
1:E:13:ASN:OD1	4:E:501:HOH:O	2.09	0.69
1:C:173:THR:HG23	1:C:174:GLN:HG3	1.73	0.69
2:D:133:LEU:HD11	2:D:139:GLU:HB2	1.75	0.68
1:C:271:LYS:NZ	4:C:406:HOH:O	2.27	0.68
1:A:79:SER:HA	1:A:265:LYS:HE2	1.76	0.67
1:C:20:THR:HG23	2:D:105:GLU:HB2	1.77	0.66
1:C:20:THR:HG22	1:C:22:LEU:H	1.60	0.65
1:C:314:GLU:OE2	4:C:404:HOH:O	2.13	0.65
1:A:183:HIS:HB2	1:A:252:ILE:HD11	1.78	0.65
1:A:199:ASP:O	1:A:248:ASN:ND2	2.29	0.65
2:B:75:ARG:N	1:E:102:GLU:OE2	2.25	0.64
1:A:266:LYS:O	1:A:304:ILE:CD1	2.47	0.63
1:C:4:LYS:HG2	2:D:27:GLU:HB3	1.81	0.62
1:E:285:THR:HB	1:E:288:GLY:O	1.98	0.62
1:A:46:ILE:HD11	1:A:304:ILE:HG23	1.81	0.62
1:E:112:ARG:HB3	1:E:261:VAL:CG1	2.31	0.61
1:E:285:THR:HG23	1:E:300:SER:HB3	1.81	0.61
2:D:34:TYR:OH	4:D:201:HOH:O	2.15	0.61
1:E:107:ILE:O	1:E:109:SER:HA	2.00	0.60
1:A:264:ASN:OD1	1:A:264:ASN:N	2.32	0.60
1:C:183:HIS:HB2	1:C:252:ILE:HD11	1.83	0.60
1:E:118:MET:HE2	1:E:257:ALA:HB2	1.84	0.60
1:E:7:ILE:HD11	2:F:122:VAL:HG21	1.84	0.60
1:A:297:GLN:NE2	1:A:310:TYR:HD1	2.01	0.59
2:B:148:CYS:O	2:B:151:SER:HB3	2.01	0.59
1:A:45:LYS:O	4:A:503:HOH:O	2.17	0.59
1:A:39:LYS:HD3	1:A:299:VAL:HG13	1.85	0.59
1:A:76:GLN:O	1:A:112:ARG:HB2	2.03	0.58
1:C:4:LYS:HZ1	2:D:25:HIS:CE1	2.19	0.58
1:A:141:THR:HG23	1:A:142:ILE:HG13	1.85	0.58
1:A:297:GLN:HE22	1:A:300:SER:H	1.51	0.58
2:D:145:ASP:HB3	2:D:148:CYS:H	1.69	0.57
1:C:158:SER:HB2	1:C:159:ALA:HA	1.86	0.57
1:E:118:MET:HE1	1:E:178:TYR:HB2	1.85	0.57
1:A:266:LYS:O	1:A:304:ILE:HD12	2.04	0.56
1:A:297:GLN:HE21	1:A:299:VAL:H	1.49	0.56
1:A:70:ASP:O	1:A:73:LEU:HB2	2.06	0.56
1:C:20:THR:HG22	1:C:22:LEU:N	2.21	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:297:GLN:HE22	1:A:310:TYR:HD1	1.52	0.56
1:C:4:LYS:HZ2	2:D:25:HIS:CE1	2.24	0.55
1:E:295:THR:HG21	2:F:56:ILE:HG12	1.88	0.55
1:E:201:TYR:HE1	1:E:248:ASN:HD22	1.52	0.55
1:C:328:ILE:O	1:C:329:ALA:HB2	2.06	0.55
1:A:121:LYS:HD3	1:A:129:THR:HB	1.89	0.55
1:E:110:GLY:N	4:E:507:HOH:O	2.38	0.55
1:A:201:TYR:CD1	1:A:248:ASN:HB2	2.42	0.55
1:A:151:VAL:HG23	1:A:255:TRP:HB2	1.88	0.54
2:B:98:LEU:HD21	2:F:99:LEU:HD13	1.88	0.54
1:E:158:SER:O	1:E:159:ALA:HB3	2.06	0.54
2:B:2:ILE:HG23	2:B:3:PHE:N	2.22	0.54
1:A:207:GLU:OE1	1:A:241:GLU:HG2	2.07	0.54
1:A:201:TYR:CE1	1:A:248:ASN:HB2	2.43	0.54
1:C:67:PRO:HG3	1:C:146:PHE:O	2.07	0.54
1:A:266:LYS:HE3	2:B:64:GLU:OE1	2.08	0.54
2:B:11:GLU:HG3	2:B:135:ASN:OD1	2.08	0.54
1:C:69:CYS:O	1:C:72:LEU:HB3	2.08	0.54
2:B:151:SER:OG	2:B:157:TYR:HA	2.08	0.53
1:A:187:ASP:HB3	1:A:190:VAL:H	1.74	0.53
1:C:25:ASN:ND2	4:C:409:HOH:O	2.40	0.53
2:D:167:LYS:HA	2:D:170:ARG:HE	1.74	0.53
1:E:290:LEU:HD21	1:E:299:VAL:HG21	1.91	0.53
1:A:227:ARG:NH2	1:E:244:ASN:HD21	2.07	0.53
1:A:102:GLU:OE2	2:D:75:ARG:N	2.33	0.52
1:E:155:LYS:HB2	1:E:194:LEU:O	2.10	0.52
1:E:46:ILE:HD11	1:E:284:GLN:HB2	1.92	0.52
1:A:227:ARG:HH21	1:E:244:ASN:HD21	1.58	0.52
2:F:2:ILE:HG23	2:F:3:PHE:N	2.25	0.52
2:F:3:PHE:CE2	2:F:113:ALA:HB2	2.44	0.52
1:A:78:TRP:CZ3	1:A:112:ARG:HA	2.44	0.52
1:E:95:GLY:HA3	1:E:230:ILE:O	2.08	0.52
1:E:42:ARG:NH2	1:E:279:CYS:SG	2.83	0.51
1:A:266:LYS:O	1:A:304:ILE:HD11	2.10	0.51
1:C:40:GLU:HB2	1:C:289:VAL:HB	1.92	0.51
2:F:158:ASP:OD1	2:F:160:PRO:HG2	2.11	0.51
1:C:157:ASP:OD1	1:C:157:ASP:N	2.40	0.50
2:D:25:HIS:HD2	2:D:33:GLY:O	1.95	0.50
1:C:121:LYS:HG2	1:C:129:THR:HB	1.93	0.50
1:E:200:LYS:NZ	4:E:509:HOH:O	2.44	0.50
1:E:285:THR:HG22	1:E:287:THR:N	2.24	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:58:LYS:HE3	2:F:97:GLU:HB3	1.94	0.50
2:D:17:MET:SD	2:D:23:GLY:HA3	2.51	0.49
1:E:43:PHE:CE2	1:E:274:LEU:HB2	2.47	0.49
1:A:283:CYS:SG	1:A:290:LEU:HB2	2.52	0.49
1:A:201:TYR:H	1:A:248:ASN:HB3	1.77	0.49
1:A:134:THR:HG23	1:A:137:CYS:H	1.77	0.49
2:D:2:ILE:HG23	2:D:3:PHE:N	2.26	0.49
1:A:118:MET:HE2	1:A:257:ALA:HB2	1.94	0.49
1:E:23:GLU:CD	1:E:323:ARG:HH22	2.15	0.49
2:F:162:TYR:OH	4:F:203:HOH:O	2.18	0.49
1:E:74:GLY:HA2	1:E:115:ARG:HD3	1.94	0.49
1:A:109:SER:N	1:A:110:GLY:HA3	2.27	0.48
1:E:64:LEU:O	1:E:147:TYR:HB3	2.13	0.48
1:E:128:ASP:OD2	1:E:131:ARG:HG3	2.13	0.48
1:E:37:ASN:HA	1:E:298:ASN:ND2	2.28	0.48
1:E:10:HIS:HB2	2:F:20:GLY:O	2.13	0.48
1:E:59:ILE:O	1:E:63:ILE:HG13	2.12	0.48
1:C:297:GLN:HG2	1:C:308:PRO:HG2	1.96	0.47
1:E:49:LYS:HB3	1:E:79:SER:HB3	1.96	0.47
2:F:38:ARG:NH1	2:F:42:GLN:OE1	2.38	0.47
2:F:74:GLU:HB3	2:F:77:ILE:HD11	1.94	0.47
1:E:277:GLU:C	1:E:279:CYS:H	2.18	0.47
1:E:67:PRO:HG3	1:E:146:PHE:O	2.14	0.47
1:E:186:LEU:HD13	1:E:227:ARG:CB	2.43	0.47
2:F:70:PHE:CD1	2:F:77:ILE:HD12	2.50	0.47
2:F:38:ARG:NH1	2:F:42:GLN:HB2	2.30	0.47
1:E:59:ILE:HG13	1:E:83:GLU:OE2	2.14	0.46
1:E:37:ASN:HA	1:E:298:ASN:HD21	1.81	0.46
1:C:165:LYS:HG2	1:C:246:GLU:HG3	1.96	0.46
1:A:95:GLY:HA3	1:A:230:ILE:O	2.15	0.46
1:E:161:TYR:CZ	1:E:249:GLY:HA2	2.50	0.46
1:E:322:LEU:HB3	2:F:111:HIS:CG	2.51	0.46
2:D:145:ASP:N	2:D:148:CYS:HB3	2.30	0.46
1:E:187:ASP:HB3	1:E:189:THR:H	1.80	0.46
2:D:3:PHE:CE2	2:D:113:ALA:HB2	2.51	0.46
1:E:285:THR:HG21	1:E:299:VAL:HG12	1.97	0.46
2:F:145:ASP:N	2:F:148:CYS:HB3	2.31	0.46
1:A:265:LYS:HA	4:A:535:HOH:O	2.16	0.45
2:F:2:ILE:HG23	2:F:3:PHE:CD2	2.51	0.45
1:A:4:LYS:NZ	2:B:25:HIS:NE2	2.51	0.45
1:C:200:LYS:HD2	1:C:249:GLY:HA3	1.98	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:9:PHE:O	2:D:135:ASN:HA	2.16	0.45
1:C:175:PRO:HD2	1:C:260:PHE:O	2.16	0.45
1:A:63:ILE:O	1:A:149:ASN:ND2	2.49	0.45
2:B:2:ILE:HG12	2:B:2:ILE:O	2.15	0.45
1:C:156:THR:C	1:C:159:ALA:CB	2.73	0.45
1:E:112:ARG:HB3	1:E:261:VAL:HG11	1.98	0.45
2:B:93:THR:O	2:B:97:GLU:HG3	2.16	0.45
1:C:95:GLY:HA3	1:C:230:ILE:O	2.16	0.45
1:C:158:SER:N	1:C:159:ALA:CB	2.73	0.45
1:E:298:ASN:H	1:E:298:ASN:HD22	1.64	0.45
1:E:132:GLY:HA3	1:E:152:TRP:HB3	1.99	0.45
1:A:78:TRP:NE1	1:A:110:GLY:O	2.50	0.45
1:C:110:GLY:HA3	1:C:260:PHE:CZ	2.52	0.45
1:E:301:PRO:HD3	1:E:310:TYR:CZ	2.51	0.45
1:E:40:GLU:HB2	1:E:289:VAL:HB	1.99	0.44
1:A:46:ILE:CD1	1:A:304:ILE:HG23	2.44	0.44
2:D:25:HIS:CD2	2:D:33:GLY:O	2.70	0.44
2:F:29:SER:OG	2:F:143:LYS:HE2	2.18	0.44
1:A:23:GLU:CD	1:A:323:ARG:HH22	2.21	0.44
1:E:118:MET:HE2	1:E:257:ALA:CB	2.47	0.44
2:D:38:ARG:HG2	2:D:38:ARG:HH11	1.83	0.44
1:E:279:CYS:SG	1:E:291:ARG:NH2	2.91	0.44
1:A:94:PRO:HG3	1:A:226:GLN:HB2	2.00	0.44
1:E:297:GLN:HG2	1:E:308:PRO:HD2	1.98	0.44
1:A:59:ILE:O	1:A:63:ILE:HG13	2.16	0.44
1:A:165:LYS:HA	1:A:245:VAL:O	2.18	0.44
1:A:310:TYR:CD2	2:B:89:LEU:HD13	2.53	0.44
1:C:71:LEU:HD12	1:C:72:LEU:HD13	1.99	0.44
1:C:83:GLU:O	1:C:271:LYS:HA	2.18	0.44
1:A:201:TYR:CD2	1:A:203:ARG:HG3	2.52	0.44
1:A:121:LYS:HE3	1:A:255:TRP:CZ2	2.53	0.43
1:A:186:LEU:HD22	1:A:227:ARG:HB2	1.99	0.43
1:C:200:LYS:O	1:C:214:SER:HB2	2.19	0.43
1:E:118:MET:HE1	1:E:178:TYR:CB	2.48	0.43
1:E:175:PRO:HD2	1:E:260:PHE:O	2.18	0.43
1:A:155:LYS:NZ	1:A:158:SER:O	2.43	0.43
1:C:93:TYR:HD2	1:C:134:THR:HG21	1.83	0.43
1:C:170:ASN:HB3	4:C:405:HOH:O	2.17	0.43
1:E:292:THR:OG1	1:E:294:LYS:HG3	2.18	0.43
1:E:105:ALA:O	1:E:108:GLY:O	2.36	0.43
1:E:124:TRP:CD1	1:E:153:ILE:HD11	2.54	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:231:ASP:HB3	1:A:233:TYR:CE2	2.54	0.43
1:C:4:LYS:CE	2:D:25:HIS:HE1	2.32	0.43
1:C:63:ILE:HD12	1:C:63:ILE:HA	1.90	0.43
1:C:186:LEU:HA	1:C:218:ALA:O	2.18	0.43
2:D:2:ILE:HD13	2:D:2:ILE:HG21	1.77	0.43
1:A:52:LEU:HD13	1:A:76:GLN:HG3	2.00	0.43
1:A:220:ARG:HB3	1:A:229:ARG:NH1	2.34	0.43
1:C:102:GLU:OE2	2:F:75:ARG:N	2.41	0.43
1:E:114:GLU:HB2	1:E:259:LYS:HB2	1.99	0.43
1:A:2:GLY:HA3	2:B:139:GLU:OE1	2.19	0.43
1:C:220:ARG:NE	4:C:421:HOH:O	2.52	0.43
1:A:23:GLU:OE2	1:A:323:ARG:NH2	2.49	0.42
1:A:93:TYR:CE1	1:A:230:ILE:HG13	2.54	0.42
2:B:20:GLY:HA3	2:B:36:ALA:HB1	1.99	0.42
1:A:290:LEU:HD21	1:A:299:VAL:HG21	2.02	0.42
2:B:167:LYS:O	2:B:171:GLN:HG2	2.20	0.42
1:A:64:LEU:O	1:A:147:TYR:HB3	2.19	0.42
1:C:135:ASN:OD1	1:C:135:ASN:N	2.46	0.42
1:A:187:ASP:CB	1:A:190:VAL:H	2.32	0.42
2:B:2:ILE:HG23	2:B:3:PHE:CD1	2.54	0.42
1:A:8:GLY:HA3	2:B:14:TRP:CH2	2.55	0.42
1:A:53:ASP:HB2	1:A:276:ILE:HD12	2.02	0.42
2:B:170:ARG:NH2	4:B:210:HOH:O	2.53	0.42
2:F:19:ASP:OD1	2:F:19:ASP:N	2.53	0.42
2:F:30:GLN:HE22	2:F:146:ASN:H	1.67	0.42
2:B:141:TRP:CE2	2:B:170:ARG:HG3	2.55	0.42
1:C:75:ASP:OD1	1:C:75:ASP:N	2.53	0.42
1:C:151:VAL:HG23	1:C:255:TRP:HB2	2.02	0.42
1:C:97:LEU:HA	1:C:232:TYR:HB2	2.01	0.42
1:E:182:VAL:HG21	1:E:213:LYS:HB3	2.02	0.42
2:B:94:TYR:OH	2:B:98:LEU:HD22	2.20	0.41
1:E:285:THR:CG2	1:E:300:SER:HB3	2.49	0.41
2:B:158:ASP:OD1	2:B:160:PRO:HD2	2.19	0.41
1:A:11:ALA:HB2	2:B:13:GLY:HA3	2.01	0.41
2:B:99:LEU:HD13	2:D:98:LEU:HD21	2.02	0.41
1:C:301:PRO:HD3	1:C:310:TYR:CZ	2.56	0.41
4:C:445:HOH:O	1:E:229:ARG:HD2	2.20	0.41
2:D:167:LYS:HA	2:D:170:ARG:NE	2.33	0.41
1:A:184:HIS:HB3	1:A:216:GLU:O	2.19	0.41
2:B:104:ASN:O	2:B:108:LEU:HG	2.21	0.41
2:F:54:SER:O	2:F:58:LYS:HG2	2.20	0.41

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:71:LEU:CD1	1:C:72:LEU:HD13	2.50	0.41
1:E:201:TYR:CD1	1:E:248:ASN:HB2	2.56	0.41
1:A:263:THR:HG23	1:A:265:LYS:O	2.21	0.41
2:B:20:GLY:HA3	2:B:36:ALA:CB	2.50	0.41
1:E:112:ARG:HB3	1:E:261:VAL:HG12	2.02	0.41
1:E:296:PHE:HZ	2:F:59:MET:HG3	1.86	0.41
1:A:111:GLU:O	1:A:112:ARG:HB3	2.20	0.41
1:A:152:TRP:CD1	1:A:152:TRP:C	2.94	0.41
1:C:156:THR:C	1:C:159:ALA:HB2	2.38	0.41
2:F:9:PHE:O	2:F:135:ASN:HA	2.21	0.41
1:A:186:LEU:CD2	1:A:227:ARG:HB2	2.50	0.41
1:A:201:TYR:HD2	1:A:203:ARG:HG3	1.86	0.41
1:A:237:LEU:HD12	1:A:241:GLU:HB3	2.02	0.41
2:B:27:GLU:HG3	2:B:32:SER:HB3	2.02	0.41
1:C:4:LYS:HG2	2:D:27:GLU:CB	2.49	0.41
1:C:167:THR:HA	1:C:243:LEU:O	2.20	0.41
1:E:201:TYR:CE1	1:E:248:ASN:HB2	2.56	0.41
1:A:21:LEU:HB2	2:B:105:GLU:OE2	2.20	0.41
1:A:78:TRP:CE2	1:A:110:GLY:O	2.73	0.41
1:C:218:ALA:O	1:C:220:ARG:NH1	2.51	0.41
1:E:118:MET:HE3	1:E:118:MET:HB3	1.89	0.41
1:E:8:GLY:HA3	2:F:14:TRP:CH2	2.55	0.40
1:A:63:ILE:HG13	1:A:63:ILE:H	1.57	0.40
1:C:170:ASN:ND2	1:C:176:ILE:HD12	2.36	0.40
1:E:73:LEU:HD11	1:E:115:ARG:CZ	2.51	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	328/333 (98%)	323 (98%)	5 (2%)	0	100 100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	321/333 (96%)	315 (98%)	6 (2%)	0	100	100
1	E	321/333 (96%)	312 (97%)	9 (3%)	0	100	100
2	B	171/180 (95%)	169 (99%)	2 (1%)	0	100	100
2	D	171/180 (95%)	168 (98%)	3 (2%)	0	100	100
2	F	170/180 (94%)	165 (97%)	5 (3%)	0	100	100
All	All	1482/1539 (96%)	1452 (98%)	30 (2%)	0	100	100

There are no Ramachandran outliers to report.

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

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	290/292 (99%)	284 (98%)	6 (2%)	53	64
1	C	285/292 (98%)	280 (98%)	5 (2%)	59	70
1	E	286/292 (98%)	281 (98%)	5 (2%)	60	72
2	B	147/153 (96%)	143 (97%)	4 (3%)	44	54
2	D	147/153 (96%)	145 (99%)	2 (1%)	67	78
2	F	146/153 (95%)	145 (99%)	1 (1%)	84	91
All	All	1301/1335 (98%)	1278 (98%)	23 (2%)	59	70

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	LYS
1	A	73	LEU
1	A	112	ARG
1	A	161	TYR
1	A	225	ASP
1	A	264	ASN
2	B	22	TYR
2	B	67	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	163	GLN
2	B	171	GLN
1	C	75	ASP
1	C	157	ASP
1	C	225	ASP
1	C	291	ARG
1	C	322	LEU
2	D	22	TYR
2	D	118	LEU
1	E	47	MET
1	E	225	ASP
1	E	291	ARG
1	E	295	THR
1	E	298	ASN
2	F	22	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	210	ASN
1	A	297	GLN
2	B	30	GLN
1	C	278	ASN
2	D	25	HIS
1	E	244	ASN
1	E	298	ASN
2	F	30	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 monosaccharides in this entry.

## 5.6 Ligand geometry

4 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	401	1	14,14,15	0.61	0	17,19,21	1.29	2 (11%)
3	NAG	E	401	1	14,14,15	0.49	0	17,19,21	1.48	3 (17%)
3	NAG	E	402	1	14,14,15	0.75	0	17,19,21	1.72	4 (23%)
3	NAG	A	402	1	14,14,15	0.65	0	17,19,21	1.81	5 (29%)

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	401	1	-	4/6/23/26	0/1/1/1
3	NAG	E	401	1	-	4/6/23/26	0/1/1/1
3	NAG	E	402	1	-	4/6/23/26	0/1/1/1
3	NAG	A	402	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	402	NAG	C1-C2-N2	4.25	117.74	110.49
3	A	402	NAG	C6-C5-C4	-3.87	103.94	113.00
3	E	402	NAG	O5-C5-C6	3.73	113.05	107.20
3	E	402	NAG	C6-C5-C4	-3.70	104.33	113.00
3	E	401	NAG	C2-N2-C7	3.47	127.84	122.90
3	A	401	NAG	O5-C5-C6	-3.24	102.13	107.20
3	E	402	NAG	O3-C3-C2	-3.08	103.09	109.47
3	A	402	NAG	C1-O5-C5	2.79	115.97	112.19
3	E	401	NAG	C1-O5-C5	2.77	115.94	112.19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	401	NAG	C1-C2-N2	-2.75	105.79	110.49
3	E	402	NAG	C1-O5-C5	2.73	115.89	112.19
3	A	402	NAG	O5-C5-C6	-2.52	103.25	107.20
3	A	401	NAG	O6-C6-C5	-2.20	103.73	111.29
3	A	402	NAG	C2-N2-C7	-2.19	119.78	122.90

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	401	NAG	C8-C7-N2-C2
3	E	401	NAG	O7-C7-N2-C2
3	E	402	NAG	C8-C7-N2-C2
3	E	402	NAG	O7-C7-N2-C2
3	A	401	NAG	O5-C5-C6-O6
3	A	401	NAG	C4-C5-C6-O6
3	E	402	NAG	O5-C5-C6-O6
3	A	401	NAG	C1-C2-N2-C7
3	E	402	NAG	C4-C5-C6-O6
3	E	401	NAG	O5-C5-C6-O6
3	E	401	NAG	C4-C5-C6-O6
3	A	401	NAG	C8-C7-N2-C2

There are no ring outliers.

No monomer is involved in short contacts.

## 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	330/333 (99%)	0.32	12 (3%) 42 43	30, 49, 66, 86	0
1	C	325/333 (97%)	0.22	15 (4%) 32 33	31, 43, 58, 87	0
1	E	325/333 (97%)	0.18	9 (2%) 53 54	31, 43, 56, 80	0
2	B	173/180 (96%)	0.43	4 (2%) 60 61	31, 40, 60, 77	0
2	D	173/180 (96%)	0.60	14 (8%) 12 12	31, 42, 61, 87	0
2	F	172/180 (95%)	0.39	10 (5%) 23 24	32, 42, 57, 85	0
All	All	1498/1539 (97%)	0.32	64 (4%) 35 36	30, 43, 61, 87	0

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	158	SER	9.0
2	D	173	ILE	8.7
2	D	1	GLY	7.0
2	F	1	GLY	6.2
2	B	1	GLY	6.2
1	A	154	VAL	5.6
1	C	329	ALA	5.1
1	A	131	ARG	4.7
2	B	173	ILE	4.6
1	E	159	ALA	4.2
2	F	171	GLN	4.2
1	E	1	PRO	4.1
2	D	171	GLN	4.0
1	A	329	ALA	3.8
2	D	168	LEU	3.7
2	D	163	GLN	3.6
1	E	108	GLY	3.6
2	F	2	ILE	3.6
1	C	157	ASP	3.4

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
2	D	170	ARG	3.3
2	F	141	TRP	2.9
1	E	158	SER	2.9
1	C	38	GLN	2.9
2	D	172	GLY	2.8
2	D	164	LYS	2.8
1	C	47	MET	2.7
2	F	170	ARG	2.7
1	A	13	ASN	2.7
2	D	2	ILE	2.6
1	A	194	LEU	2.6
1	A	46	ILE	2.5
2	B	2	ILE	2.5
2	B	18	ILE	2.5
1	C	4	LYS	2.5
2	F	38	ARG	2.4
1	A	54	LEU	2.4
2	F	172	GLY	2.4
1	C	194	LEU	2.4
2	D	10	ILE	2.4
1	E	74	GLY	2.4
1	C	193	ASN	2.4
2	D	4	GLY	2.4
2	D	167	LYS	2.3
2	F	134	GLY	2.3
1	C	160	THR	2.2
1	E	171	THR	2.2
2	F	168	LEU	2.2
2	F	113	ALA	2.2
1	A	266	LYS	2.2
1	C	86	ASN	2.2
1	C	225	ASP	2.2
1	E	47	MET	2.1
1	C	328	ILE	2.1
2	D	133	LEU	2.1
1	E	4	LYS	2.1
1	C	71	LEU	2.1
1	C	190	VAL	2.1
1	A	48	ASN	2.1
1	A	124	TRP	2.1
1	A	238	ARG	2.0
1	C	197	SER	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	38	ARG	2.0
1	A	264	ASN	2.0
1	E	88	GLN	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 monosaccharides in this entry.

## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	A	401	14/15	0.74	0.21	56,75,78,83	0
3	NAG	E	401	14/15	0.78	0.39	80,87,91,95	0
3	NAG	E	402	14/15	0.82	0.17	59,68,74,74	0
3	NAG	A	402	14/15	0.83	0.11	65,68,73,76	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.