Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 06:37 PM GMT

PDB ID	:	4M4Y
Title	:	Crystal structure of a 2009 H1N1 influenza virus hemagglutinin with a stabilization mutation HA2 E47G
Authors	:	Zhu, X.; Wilson, I.A.
Deposited on	:	2013-08-07
Resolution	:	2.20 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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

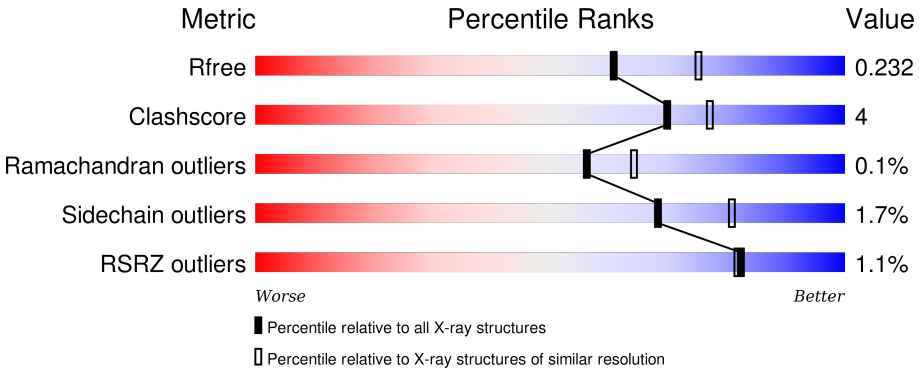
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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.



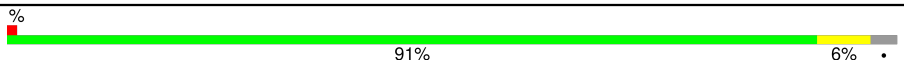
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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 >=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 <=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	85% 13% •
1	C	331	2% 86% 11% • •
1	E	331	% 82% 16% • •
2	B	177	2% 90% 7% • •
2	D	177	% 89% 8% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	177	

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
3	NAG	A	401	-	-	-	X
4	MAN	C	406	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12892 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 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	324	Total	C	N	O	S	0	0	0
			2529	1599	436	483	11			
1	C	323	Total	C	N	O	S	0	0	0
			2523	1596	435	481	11			
1	E	326	Total	C	N	O	S	0	0	0
			2544	1608	438	487	11			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	EXPRESSION TAG	UNP C3W5S1
A	8	ASP	-	EXPRESSION TAG	UNP C3W5S1
A	9	PRO	-	EXPRESSION TAG	UNP C3W5S1
A	10	GLY	-	EXPRESSION TAG	UNP C3W5S1
C	7	ALA	-	EXPRESSION TAG	UNP C3W5S1
C	8	ASP	-	EXPRESSION TAG	UNP C3W5S1
C	9	PRO	-	EXPRESSION TAG	UNP C3W5S1
C	10	GLY	-	EXPRESSION TAG	UNP C3W5S1
E	7	ALA	-	EXPRESSION TAG	UNP C3W5S1
E	8	ASP	-	EXPRESSION TAG	UNP C3W5S1
E	9	PRO	-	EXPRESSION TAG	UNP C3W5S1
E	10	GLY	-	EXPRESSION TAG	UNP C3W5S1

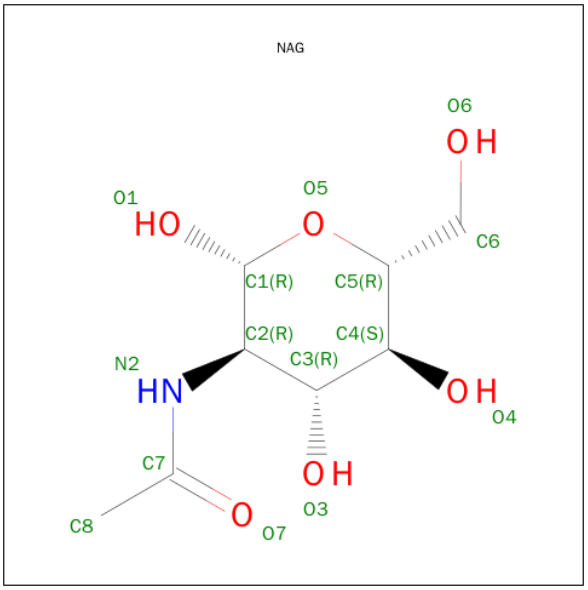
- Molecule 2 is a protein called Hemagglutinin HA2 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	175	Total	C	N	O	S	0	0	0
			1406	881	238	281	6			
2	D	171	Total	C	N	O	S	0	0	0
			1375	863	234	272	6			
2	F	171	Total	C	N	O	S	0	0	0
			1375	863	234	272	6			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	47	GLY	GLU	ENGINEERED MUTATION	UNP C3W5S1
B	175	SER	-	EXPRESSION TAG	UNP C3W5S1
B	176	GLY	-	EXPRESSION TAG	UNP C3W5S1
B	177	ARG	-	EXPRESSION TAG	UNP C3W5S1
D	47	GLY	GLU	ENGINEERED MUTATION	UNP C3W5S1
D	175	SER	-	EXPRESSION TAG	UNP C3W5S1
D	176	GLY	-	EXPRESSION TAG	UNP C3W5S1
D	177	ARG	-	EXPRESSION TAG	UNP C3W5S1
F	47	GLY	GLU	ENGINEERED MUTATION	UNP C3W5S1
F	175	SER	-	EXPRESSION TAG	UNP C3W5S1
F	176	GLY	-	EXPRESSION TAG	UNP C3W5S1
F	177	ARG	-	EXPRESSION TAG	UNP C3W5S1

- Molecule 3 is SUGAR (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	C	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

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

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

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

- Molecule 5 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	E	3	Total C N O 38 22 2 14	0	0

- Molecule 6 is water.

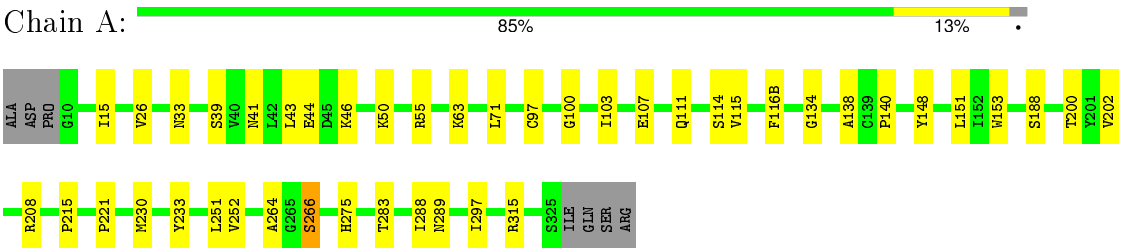
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	173	Total O 173 173	0	0
6	C	181	Total O 181 181	0	0
6	E	205	Total O 205 205	0	0
6	B	118	Total O 118 118	0	0
6	D	120	Total O 120 120	0	0
6	F	143	Total O 143 143	0	0

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.

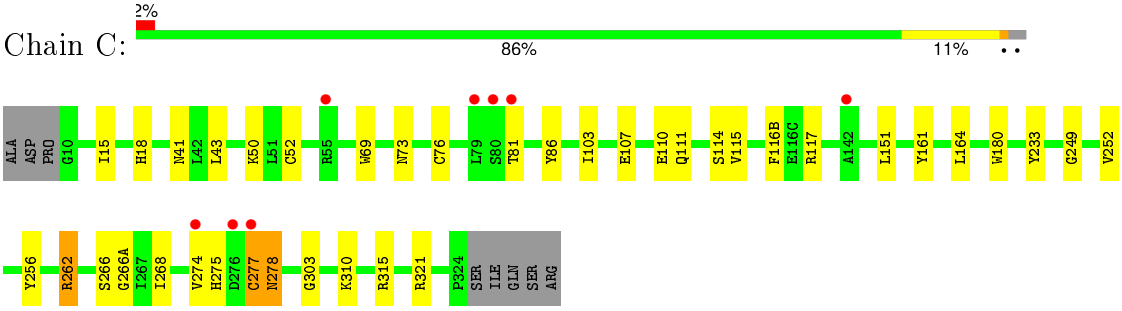
- Molecule 1: Hemagglutinin HA1 subunit

Chain A:



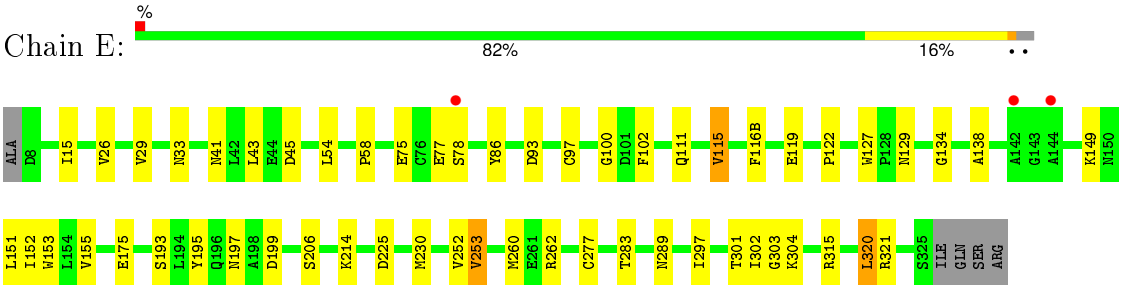
- Molecule 1: Hemagglutinin HA1 subunit

Chain C:



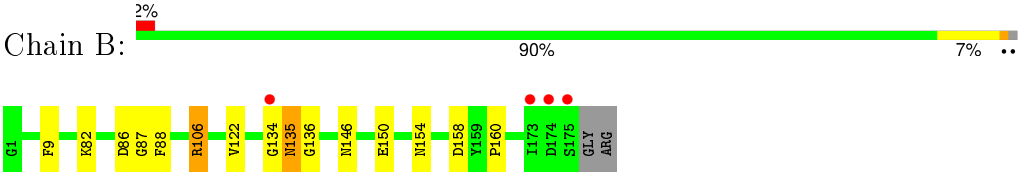
- Molecule 1: Hemagglutinin HA1 subunit

Chain E:



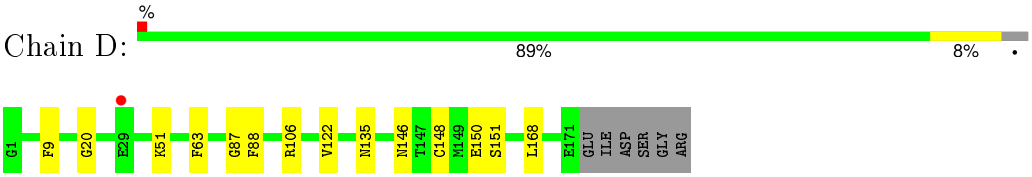
- Molecule 2: Hemagglutinin HA2 subunit

Chain B:



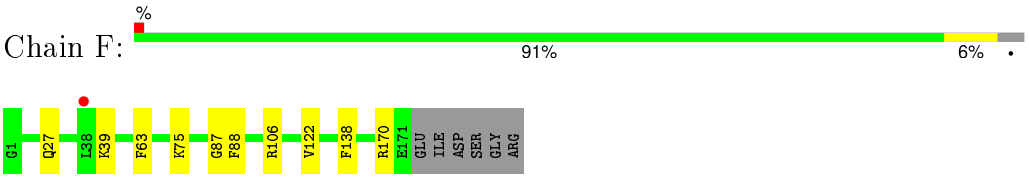
• Molecule 2: Hemagglutinin HA2 subunit

Chain D:



• Molecule 2: Hemagglutinin HA2 subunit

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	71.37Å 132.17Å 201.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.49 – 2.20 48.49 – 2.18	Depositor EDS
% Data completeness (in resolution range)	95.7 (48.49-2.20) 95.2 (48.49-2.18)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 2.18Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.179 , 0.232 0.186 , 0.232	Depositor DCC
R_{free} test set	4709 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	30.2	Xtriage
Anisotropy	0.815	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 44.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 95245 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12892	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.79% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: FUL, 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	0.39	0/2593	0.54	0/3524
1	C	0.39	0/2587	0.58	1/3516 (0.0%)
1	E	0.40	0/2609	0.59	0/3547
2	B	0.42	0/1434	0.55	1/1932 (0.1%)
2	D	0.40	0/1403	0.54	0/1890
2	F	0.41	0/1403	0.55	0/1890
All	All	0.40	0/12029	0.56	2/16299 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	164	LEU	CA-CB-CG	5.57	128.10	115.30
2	B	106	ARG	NE-CZ-NH2	-5.48	117.56	120.30

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	2529	0	2476	24	0
1	C	2523	0	2470	34	0
1	E	2544	0	2487	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1406	0	1326	10	0
2	D	1375	0	1300	10	0
2	F	1375	0	1300	9	0
3	A	42	0	39	1	0
3	C	42	0	39	0	0
3	E	28	0	26	1	0
4	C	50	0	43	1	0
5	E	38	0	34	0	0
6	A	173	0	0	4	0
6	B	118	0	0	2	0
6	C	181	0	0	2	0
6	D	120	0	0	1	0
6	E	205	0	0	3	0
6	F	143	0	0	2	0
All	All	12892	0	11540	105	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 4.

All (105) 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:262:ARG:HG3	1:C:262:ARG:HH11	0.96	1.11
1:C:111:GLN:HE22	1:C:262:ARG:NH2	1.60	0.97
1:C:262:ARG:HG3	1:C:262:ARG:NH1	1.74	0.93
1:C:111:GLN:HE22	1:C:262:ARG:HH21	1.11	0.92
1:C:115:VAL:HG11	1:C:116(B):PHE:HB2	1.63	0.80
1:C:111:GLN:NE2	1:C:262:ARG:HH21	1.79	0.80
1:C:111:GLN:NE2	1:C:262:ARG:NH2	2.33	0.76
1:E:115:VAL:HG11	1:E:116(B):PHE:HB2	1.68	0.75
1:C:262:ARG:HH11	1:C:262:ARG:CG	1.86	0.74
1:C:41:ASN:ND2	1:C:43:LEU:O	2.26	0.68
1:A:103:ILE:HG13	1:A:233:TYR:CE2	2.30	0.67
1:A:114:SER:HB2	1:A:266:SER:HB2	1.78	0.64
1:C:117:ARG:HD3	1:C:256:TYR:CD1	2.34	0.63
1:A:115:VAL:HG11	1:A:116(B):PHE:HB2	1.81	0.63
1:C:262:ARG:NH1	1:C:262:ARG:CG	2.53	0.63
1:E:283:THR:HG22	1:E:301:THR:HG22	1.82	0.61
1:A:15:ILE:HD11	2:B:122:VAL:HG21	1.82	0.61
1:E:289:ASN:ND2	6:E:667:HOH:O	2.27	0.60
1:A:140:PRO:HD2	3:A:401:NAG:H83	1.84	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:106:ARG:HD3	6:F:280:HOH:O	2.01	0.59
1:C:110:GLU:HG3	2:F:75:LYS:HE3	1.84	0.59
1:E:77:GLU:HG3	1:E:149:LYS:HE3	1.85	0.59
2:B:134:GLY:O	2:B:136:GLY:N	2.36	0.58
1:A:50:LYS:HD2	1:A:275:HIS:HB2	1.85	0.58
1:E:175:GLU:OE1	1:E:262:ARG:NH1	2.37	0.57
2:F:170:ARG:NH2	6:F:297:HOH:O	2.38	0.57
1:E:33:ASN:ND2	6:E:612:HOH:O	2.33	0.56
2:D:106:ARG:HD3	6:D:267:HOH:O	2.05	0.56
1:C:111:GLN:O	1:C:262:ARG:HD2	2.07	0.55
2:D:88:PHE:CZ	2:F:87:GLY:HA3	2.42	0.55
1:E:199:ASP:OD1	1:E:214:LYS:NZ	2.36	0.54
1:C:151:LEU:HB3	1:C:252:VAL:HG12	1.90	0.53
1:A:151:LEU:HB3	1:A:252:VAL:HG12	1.90	0.53
2:B:106:ARG:HD3	6:B:263:HOH:O	2.10	0.52
1:C:103:ILE:HG13	1:C:233:TYR:CE2	2.45	0.52
1:A:221:PRO:HG2	1:E:206:SER:HA	1.92	0.52
1:C:266:SER:OG	1:C:266(A):GLY:N	2.44	0.50
2:D:9:PHE:O	2:D:135:ASN:HA	2.13	0.49
1:E:29:VAL:HG22	2:D:51:LYS:HG3	1.94	0.49
1:C:52:CYS:HB3	1:C:277:CYS:O	2.13	0.49
1:E:15:ILE:HD11	2:F:122:VAL:HG21	1.95	0.48
1:A:33:ASN:ND2	6:A:586:HOH:O	2.20	0.48
1:E:26:VAL:HG12	1:E:315:ARG:HG2	1.96	0.48
2:B:82:LYS:NZ	2:B:86:ASP:OD2	2.42	0.48
1:E:100:GLY:HA3	1:E:230:MET:O	2.13	0.48
1:A:188:SER:O	6:A:607:HOH:O	2.20	0.48
1:E:152:ILE:HB	1:E:253:VAL:HG12	1.96	0.48
1:E:111:GLN:OE1	1:E:262:ARG:NE	2.47	0.48
1:C:69:TRP:HE1	1:C:81:THR:CG2	2.26	0.47
1:C:111:GLN:NE2	1:C:262:ARG:CZ	2.77	0.47
2:B:9:PHE:O	2:B:135:ASN:HA	2.15	0.47
1:C:15:ILE:HD11	2:D:122:VAL:HG21	1.96	0.47
2:D:146:ASN:O	2:D:150:GLU:HG2	2.15	0.47
1:A:289:ASN:HB3	6:A:672:HOH:O	2.13	0.47
2:B:154:ASN:ND2	6:B:283:HOH:O	2.47	0.47
1:E:41:ASN:ND2	1:E:43:LEU:O	2.48	0.47
1:E:303:GLY:HA2	2:F:63:PHE:CE2	2.49	0.47
1:C:278:ASN:OD1	1:C:278:ASN:N	2.47	0.46
1:A:97:CYS:HB2	1:A:138:ALA:O	2.15	0.46
1:E:45:ASP:C	1:E:297:ILE:HD11	2.35	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:50:LYS:HD3	1:C:275:HIS:CG	2.51	0.46
1:A:107:GLU:O	1:A:111:GLN:HG2	2.15	0.46
1:A:71:LEU:O	1:A:148:TYR:HB3	2.16	0.46
1:A:288:ILE:HG21	1:A:297:ILE:HG13	1.98	0.45
1:E:116(B):PHE:CE1	1:E:260:MET:HE2	2.52	0.45
1:A:39:SER:OG	1:A:315:ARG:NE	2.47	0.45
1:C:321:ARG:NH1	6:C:650:HOH:O	2.49	0.45
1:C:73:ASN:HB3	1:C:76:CYS:SG	2.57	0.44
1:C:111:GLN:NE2	1:C:262:ARG:NE	2.66	0.44
1:C:114:SER:HB2	1:C:266:SER:HB3	2.00	0.44
1:A:44:GLU:OE2	1:A:46:LYS:HG2	2.17	0.44
1:E:195:TYR:O	1:E:197:ASN:N	2.51	0.44
1:E:320:LEU:HD12	1:E:321:ARG:O	2.18	0.44
2:B:88:PHE:CZ	2:D:87:GLY:HA3	2.53	0.43
1:C:86:TYR:CE2	1:C:268:ILE:HD12	2.53	0.43
1:E:75:GLU:HG2	3:E:401:NAG:H82	1.98	0.43
1:E:127:TRP:CH2	1:E:253:VAL:HG21	2.53	0.43
1:A:200:THR:OG1	1:A:215:PRO:HG2	2.18	0.43
1:E:119:GLU:CD	1:E:122:PRO:HA	2.38	0.43
2:D:148:CYS:O	2:D:151:SER:OG	2.24	0.43
1:E:151:LEU:HB3	1:E:252:VAL:HG12	2.00	0.43
1:C:310:LYS:HA	1:C:310:LYS:HD2	1.80	0.43
1:A:26:VAL:HG12	1:A:315:ARG:HG2	2.00	0.43
1:C:161:TYR:CZ	1:C:249:GLY:HA2	2.54	0.43
1:E:15:ILE:HD13	2:F:138:PHE:HE1	1.84	0.42
2:B:87:GLY:HA3	2:F:88:PHE:CZ	2.55	0.42
1:E:129:ASN:ND2	6:E:664:HOH:O	2.28	0.42
1:E:58:PRO:HB3	1:E:86:TYR:CZ	2.55	0.42
2:F:39:LYS:HE3	2:F:39:LYS:HB2	1.84	0.42
1:C:111:GLN:NE2	1:C:262:ARG:HE	2.18	0.41
1:A:202:VAL:HG11	1:A:251:LEU:HD13	2.02	0.41
1:A:41:ASN:ND2	1:A:43:LEU:O	2.53	0.41
1:C:107:GLU:O	1:C:111:GLN:HG2	2.20	0.41
1:C:303:GLY:HA2	2:D:63:PHE:CE2	2.55	0.41
4:C:405:BMA:O6	6:C:627:HOH:O	2.22	0.41
1:E:54:LEU:HD21	1:E:302:ILE:HG22	2.03	0.41
1:A:55:ARG:NH2	6:A:669:HOH:O	2.53	0.41
1:C:180:TRP:CE2	1:C:233:TYR:HB2	2.56	0.41
1:A:134:GLY:HA3	1:A:153:TRP:HB3	2.03	0.41
2:B:146:ASN:O	2:B:150:GLU:HG2	2.21	0.41
1:E:97:CYS:HB2	1:E:138:ALA:O	2.21	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:134:GLY:HA3	1:E:153:TRP:HB3	2.03	0.41
1:C:18:HIS:HB2	2:D:20:GLY:O	2.20	0.40
1:A:100:GLY:HA3	1:A:230:MET:O	2.20	0.40
2:B:158:ASP:OD1	2:B:160:PRO:HD2	2.21	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	322/331 (97%)	311 (97%)	10 (3%)	1 (0%)	46 50
1	C	321/331 (97%)	313 (98%)	8 (2%)	0	100 100
1	E	324/331 (98%)	313 (97%)	11 (3%)	0	100 100
2	B	173/177 (98%)	167 (96%)	5 (3%)	1 (1%)	30 29
2	D	169/177 (96%)	166 (98%)	3 (2%)	0	100 100
2	F	169/177 (96%)	166 (98%)	3 (2%)	0	100 100
All	All	1478/1524 (97%)	1436 (97%)	40 (3%)	2 (0%)	56 64

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	135	ASN
1	A	264	ALA

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	284/290 (98%)	280 (99%)	4 (1%)	74 85
1	C	283/290 (98%)	278 (98%)	5 (2%)	66 79
1	E	286/290 (99%)	275 (96%)	11 (4%)	40 49
2	B	150/151 (99%)	150 (100%)	0	100 100
2	D	146/151 (97%)	145 (99%)	1 (1%)	88 94
2	F	146/151 (97%)	145 (99%)	1 (1%)	88 94
All	All	1295/1323 (98%)	1273 (98%)	22 (2%)	68 81

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	63	LYS
1	A	208	ARG
1	A	266	SER
1	A	283	THR
1	C	262	ARG
1	C	274	VAL
1	C	277	CYS
1	C	278	ASN
1	C	315	ARG
1	E	78	SER
1	E	93	ASP
1	E	102	PHE
1	E	115	VAL
1	E	155	VAL
1	E	193	SER
1	E	225	ASP
1	E	253	VAL
1	E	277	CYS
1	E	304	LYS
1	E	320	LEU
2	D	168	LEU
2	F	27	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	133	ASN
1	C	111	GLN
2	B	43	ASN
2	B	95	ASN
2	F	79	ASN

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)

7 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	$\#\|Z\| > 2$
4	NAG	C	403	1,4	14,14,15	0.27	0	15,19,21	0.18	0
4	NAG	C	404	4	14,14,15	0.48	0	15,19,21	0.30	0
4	BMA	C	405	4	11,11,12	0.62	0	14,15,17	1.26	2 (14%)
4	MAN	C	406	4	11,11,12	0.90	1 (9%)	14,15,17	1.28	2 (14%)
5	NAG	E	403	1,5	14,14,15	0.30	0	15,19,21	0.71	1 (6%)
5	NAG	E	404	5	14,14,15	0.25	0	15,19,21	0.43	0
5	FUL	E	405	5	10,10,11	1.56	2 (20%)	14,14,16	1.63	2 (14%)

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
4	NAG	C	403	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	404	4	-	0/6/23/26	0/1/1/1
4	BMA	C	405	4	-	0/2/19/22	0/1/1/1
4	MAN	C	406	4	-	0/2/19/22	0/1/1/1
5	NAG	E	403	1,5	-	0/6/23/26	0/1/1/1
5	NAG	E	404	5	-	0/6/23/26	0/1/1/1
5	FUL	E	405	5	-	0/0/17/20	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	405	FUL	O5-C1	-2.46	1.39	1.43
4	C	406	MAN	C2-C3	2.17	1.55	1.52
5	E	405	FUL	C4-C5	3.48	1.60	1.52

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	405	FUL	C1-C2-C3	-4.49	104.23	109.54
4	C	405	BMA	O2-C2-C3	-2.35	105.39	110.12
5	E	403	NAG	O3-C3-C2	-2.17	104.81	109.11
4	C	406	MAN	O5-C1-C2	2.05	114.18	110.86
5	E	405	FUL	O5-C5-C4	2.45	113.78	109.53
4	C	405	BMA	C1-O5-C5	2.79	115.79	112.25
4	C	406	MAN	C1-O5-C5	2.89	115.91	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	405	BMA	1	0

5.6 Ligand geometry (i)

8 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	0.73	1 (7%)	15,19,21	1.18	1 (6%)
3	NAG	A	402	1	14,14,15	0.43	0	15,19,21	0.38	0
3	NAG	A	403	1	14,14,15	0.42	0	15,19,21	1.40	3 (20%)
3	NAG	C	401	1	14,14,15	0.48	0	15,19,21	0.62	1 (6%)
3	NAG	C	402	1	14,14,15	0.62	0	15,19,21	0.77	1 (6%)
3	NAG	C	407	1	14,14,15	0.49	0	15,19,21	0.91	1 (6%)
3	NAG	E	401	1	14,14,15	0.59	0	15,19,21	0.33	0
3	NAG	E	402	1	14,14,15	0.32	0	15,19,21	0.56	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
3	NAG	A	401	1	-	0/6/23/26	0/1/1/1
3	NAG	A	402	1	-	0/6/23/26	0/1/1/1
3	NAG	A	403	1	-	0/6/23/26	0/1/1/1
3	NAG	C	401	1	-	0/6/23/26	0/1/1/1
3	NAG	C	402	1	-	0/6/23/26	0/1/1/1
3	NAG	C	407	1	-	0/6/23/26	0/1/1/1
3	NAG	E	401	1	-	0/6/23/26	0/1/1/1
3	NAG	E	402	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	401	NAG	O5-C1	2.23	1.47	1.43

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	403	NAG	C2-N2-C7	-2.63	119.66	123.04
3	A	403	NAG	C4-C3-C2	-2.39	107.52	111.23
3	C	402	NAG	C2-N2-C7	-2.03	120.43	123.04
3	C	401	NAG	C1-O5-C5	2.02	114.81	112.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	403	NAG	C1-O5-C5	2.80	115.80	112.25
3	C	407	NAG	C1-O5-C5	2.94	115.98	112.25
3	A	401	NAG	C1-O5-C5	4.15	117.52	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	401	NAG	1	0
3	E	401	NAG	1	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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^{th} 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^2)	Q<0.9
1	A	324/331 (97%)	-0.29	0 100 100	23, 35, 51, 61	0
1	C	323/331 (97%)	-0.24	8 (2%) 61 60	24, 35, 56, 75	0
1	E	326/331 (98%)	-0.32	3 (0%) 85 85	22, 34, 52, 63	0
2	B	175/177 (98%)	-0.17	4 (2%) 64 63	23, 33, 54, 82	0
2	D	171/177 (96%)	-0.21	1 (0%) 90 90	23, 33, 52, 67	0
2	F	171/177 (96%)	-0.25	1 (0%) 90 90	20, 29, 44, 57	0
All	All	1490/1524 (97%)	-0.26	17 (1%) 82 82	20, 34, 52, 82	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	134	GLY	5.3
2	B	175	SER	4.6
1	C	80	SER	4.0
2	B	173	ILE	3.9
1	C	277	CYS	3.6
2	B	174	ASP	3.3
1	C	55	ARG	3.2
1	E	78	SER	2.7
1	E	142	ALA	2.6
1	C	274	VAL	2.5
2	F	38	LEU	2.4
1	C	81	THR	2.3
1	C	79	LEU	2.3
1	E	144	ALA	2.2
1	C	276	ASP	2.0
2	D	29	GLU	2.0
1	C	142	ALA	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)

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^{th} 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^2)	Q<0.9
4	MAN	C	406	11/12	0.91	0.33	15.58	50,58,63,74	0
4	BMA	C	405	11/12	0.83	0.30	-	52,59,62,64	0
5	FUL	E	405	10/11	0.88	0.23	-	53,68,71,75	0
5	NAG	E	403	14/15	0.90	0.19	-	44,59,73,75	0
5	NAG	E	404	14/15	0.73	0.40	-	68,80,86,86	0
4	NAG	C	404	14/15	0.87	0.26	-	63,69,85,88	0
4	NAG	C	403	14/15	0.88	0.21	-	56,64,73,74	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^{th} 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^2)	Q<0.9
3	NAG	A	401	14/15	0.83	0.22	2.79	46,62,68,70	0
3	NAG	C	401	14/15	0.90	0.20	1.49	46,55,64,65	0
3	NAG	E	401	14/15	0.91	0.14	1.22	39,48,55,58	0
3	NAG	C	407	14/15	0.78	0.21	-	60,77,86,88	0
3	NAG	A	403	14/15	0.84	0.11	-	56,62,64,67	0
3	NAG	E	402	14/15	0.84	0.29	-	64,71,79,87	0
3	NAG	C	402	14/15	0.80	0.36	-	74,77,80,83	0
3	NAG	A	402	14/15	0.88	0.33	-	59,69,74,75	0

6.5 Other polymers (i)

There are no such residues in this entry.