Full wwPDB X-ray Structure Validation Report (i)

Feb 19, 2016 – 09:56 PM GMT

PDB ID	:	5BQY
Title	:	Crystal structure of hemagglutinin of A/Chicken/Guangdong/S1311/2010 (H6N6) in complex with avian-like receptor LSTa
Authors	:	Ni, F.; Kondrashkina, E.; Wang, Q.
Deposited on	:	2015-05-29
Resolution	:	2.78 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:

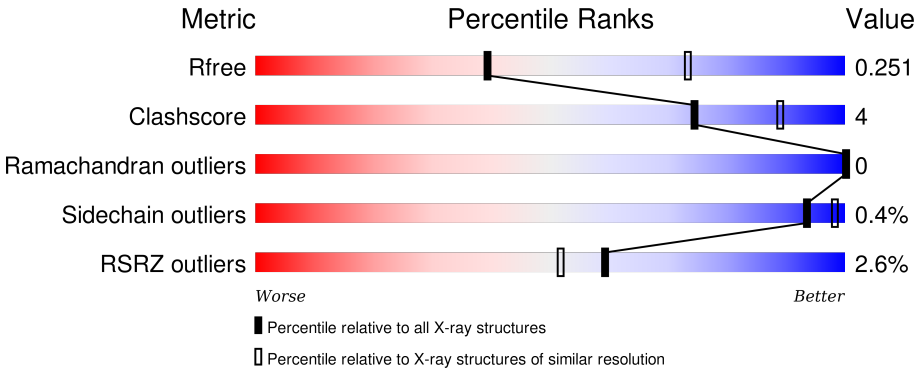
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20026982

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 2.78 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.



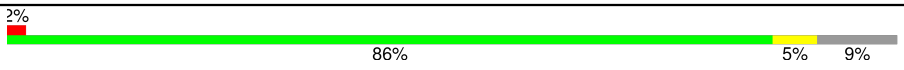
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3004 (2.80-2.76)
Clashscore	102246	3480 (2.80-2.76)
Ramachandran outliers	100387	3423 (2.80-2.76)
Sidechain outliers	100360	3425 (2.80-2.76)
RSRZ outliers	91569	3016 (2.80-2.76)

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	324	3% 89% 10%
1	C	324	% 88% 11%
1	E	324	4% 86% 13%
2	B	191	3% 85% 5% 10%
2	D	191	3% 88% • 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	191	

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	404	-	-	-	X
3	NAG	B	301	-	-	-	X
4	SIA	A	406	-	-	-	X
4	SIA	E	405	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 12368 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called HEMAGGLUTININ HA1 CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	324	Total	C	N	O	S	0	0	0
			2557	1613	451	481	12			
1	C	324	Total	C	N	O	S	0	0	0
			2557	1613	451	481	12			
1	E	324	Total	C	N	O	S	0	0	0
			2557	1613	451	481	12			

- Molecule 2 is a protein called HEMAGGLUTININ HA2 CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	171	Total	C	N	O	S	0	0	0
			1381	857	246	271	7			
2	D	173	Total	C	N	O	S	0	0	0
			1398	869	249	273	7			
2	F	174	Total	C	N	O	S	0	0	0
			1407	874	250	276	7			

There are 18 discrepancies between the modelled and reference sequences:

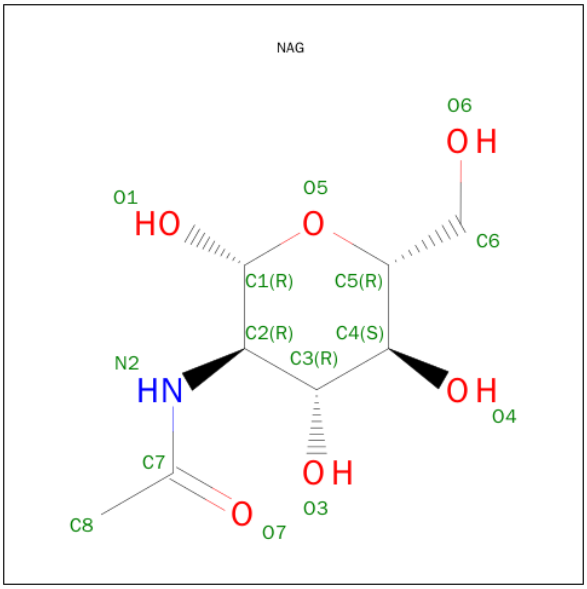
Chain	Residue	Modelled	Actual	Comment	Reference
B	186	GLY	-	expression tag	UNP A0A067YZ73
B	187	ALA	-	expression tag	UNP A0A067YZ73
B	188	LEU	-	expression tag	UNP A0A067YZ73
B	189	VAL	-	expression tag	UNP A0A067YZ73
B	190	PRO	-	expression tag	UNP A0A067YZ73
B	191	ARG	-	expression tag	UNP A0A067YZ73
D	186	GLY	-	expression tag	UNP A0A067YZ73
D	187	ALA	-	expression tag	UNP A0A067YZ73
D	188	LEU	-	expression tag	UNP A0A067YZ73
D	189	VAL	-	expression tag	UNP A0A067YZ73
D	190	PRO	-	expression tag	UNP A0A067YZ73
D	191	ARG	-	expression tag	UNP A0A067YZ73

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	186	GLY	-	expression tag	UNP A0A067YZ73
F	187	ALA	-	expression tag	UNP A0A067YZ73
F	188	LEU	-	expression tag	UNP A0A067YZ73
F	189	VAL	-	expression tag	UNP A0A067YZ73
F	190	PRO	-	expression tag	UNP A0A067YZ73
F	191	ARG	-	expression tag	UNP A0A067YZ73

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



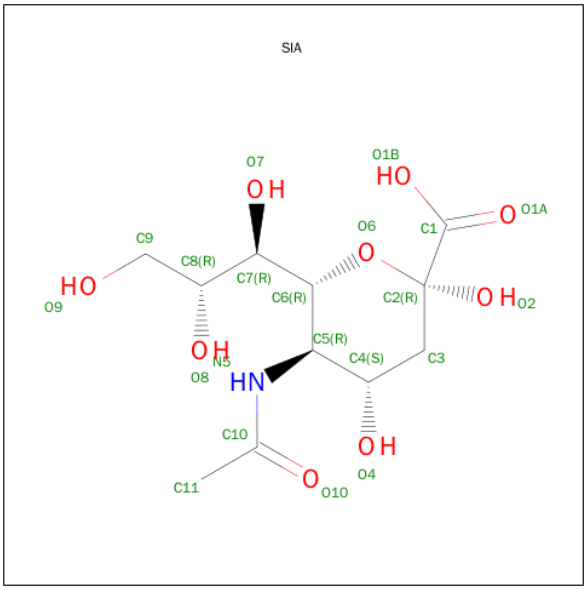
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1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

Continued on next page...

Continued from previous page...

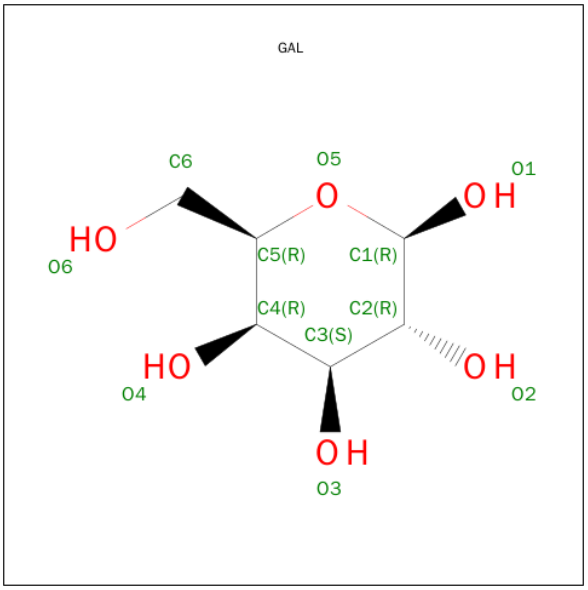
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total C N O 14 8 1 5	0	0
3	C	1	Total C N O 13 8 1 4	0	0
3	C	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	E	1	Total C N O 14 8 1 5	0	0
3	F	1	Total C N O 14 8 1 5	0	0
3	F	1	Total C N O 14 8 1 5	0	0
3	F	1	Total C N O 14 8 1 5	0	0

- Molecule 4 is O-SIALIC ACID (three-letter code: SIA) (formula: $C_{11}H_{19}NO_9$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 20 11 1 8	0	0
4	E	1	Total C N O 20 11 1 8	0	0

- Molecule 5 is BETA-D-GALACTOSE (three-letter code: GAL) (formula: $C_6H_{12}O_6$).



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

- Molecule 6 is water.

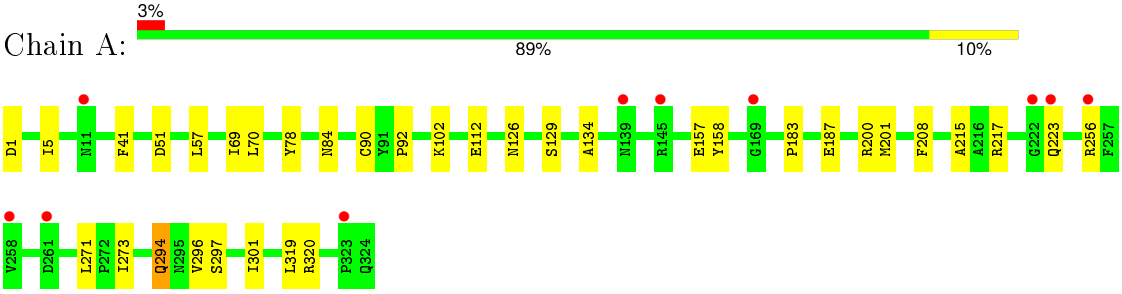
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	30	Total O 30 30	0	0
6	B	23	Total O 23 23	0	0
6	C	46	Total O 46 46	0	0
6	D	13	Total O 13 13	0	0
6	E	40	Total O 40 40	0	0
6	F	32	Total O 32 32	0	0

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 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.

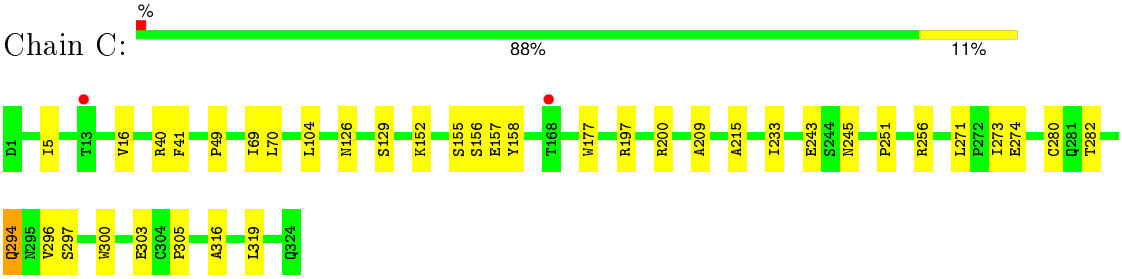
- Molecule 1: HEMAGGLUTININ HA1 CHAIN

Chain A:



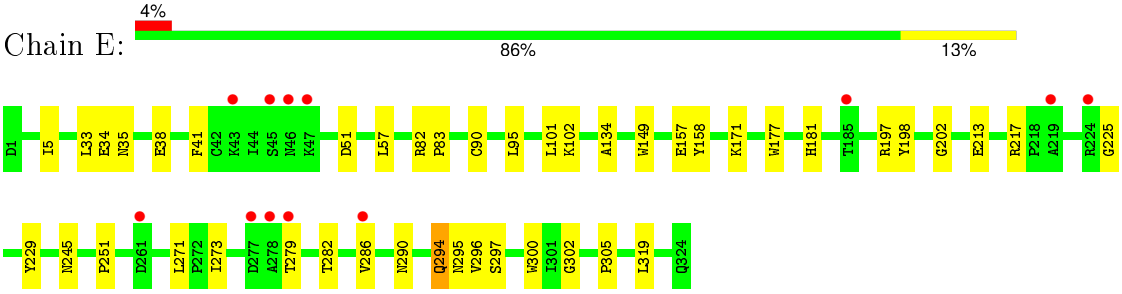
- Molecule 1: HEMAGGLUTININ HA1 CHAIN

Chain C:



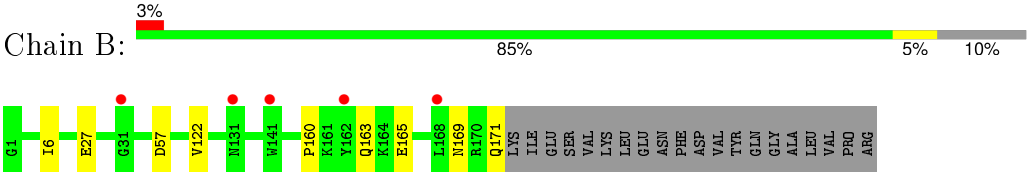
- Molecule 1: HEMAGGLUTININ HA1 CHAIN

Chain E:



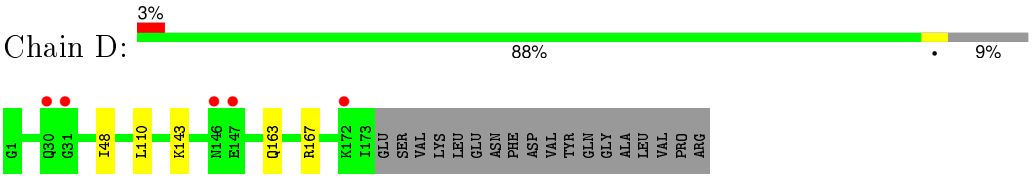
- Molecule 2: HEMAGGLUTININ HA2 CHAIN

Chain B:



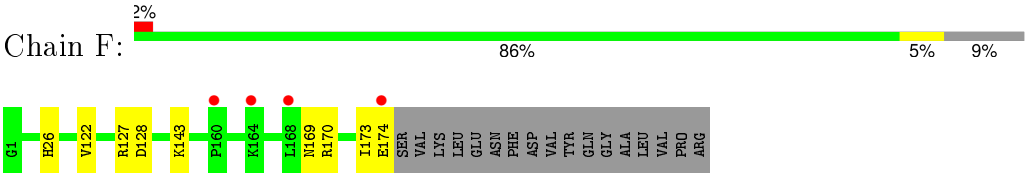
- Molecule 2: HEMAGGLUTININ HA2 CHAIN

Chain D:



• Molecule 2: HEMAGGLUTININ HA2 CHAIN

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	69.35Å 137.77Å 199.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.87 – 2.78 48.87 – 2.78	Depositor EDS
% Data completeness (in resolution range)	99.3 (48.87-2.78) 99.4 (48.87-2.78)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.22 (at 2.77Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.196 , 0.252 0.196 , 0.251	Depositor DCC
R_{free} test set	2461 reflections (5.07%)	DCC
Wilson B-factor (Å^2)	47.1	Xtriage
Anisotropy	0.552	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 43.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 48574 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	12368	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.57% 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: SIA, GAL, 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	0.27	0/2617	0.44	0/3553
1	C	0.30	0/2617	0.48	0/3553
1	E	0.27	0/2617	0.45	1/3553 (0.0%)
2	B	0.27	0/1409	0.42	0/1897
2	D	0.26	0/1426	0.41	0/1919
2	F	0.27	0/1435	0.44	0/1931
All	All	0.27	0/12121	0.44	1/16406 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	319	LEU	CA-CB-CG	5.47	127.88	115.30

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	2557	0	2511	24	0
1	C	2557	0	2511	24	0
1	E	2557	0	2512	31	0
2	B	1381	0	1292	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1398	0	1317	3	0
2	F	1407	0	1321	9	0
3	A	70	0	64	0	0
3	B	28	0	25	1	0
3	C	55	0	49	1	0
3	E	70	0	63	1	0
3	F	42	0	37	1	0
4	A	20	0	17	1	0
4	E	20	0	17	2	0
5	A	11	0	9	0	0
5	E	11	0	9	0	0
6	A	30	0	0	3	0
6	B	23	0	0	3	0
6	C	46	0	0	3	0
6	D	13	0	0	0	0
6	E	40	0	0	5	0
6	F	32	0	0	2	0
All	All	12368	0	11754	96	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 (96) 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:5:ILE:O	6:C:501:HOH:O	1.84	0.94
2:F:127:ARG:O	6:F:401:HOH:O	1.92	0.88
1:A:84:ASN:ND2	6:A:503:HOH:O	2.15	0.80
1:E:34:GLU:OE2	6:E:501:HOH:O	2.00	0.79
1:C:49:PRO:O	6:C:502:HOH:O	2.05	0.75
1:C:156:SER:OG	1:C:157:GLU:N	2.20	0.74
1:C:294:GLN:HG2	1:C:296:VAL:H	1.55	0.71
1:A:294:GLN:HG2	1:A:296:VAL:H	1.57	0.70
2:B:57:ASP:OD1	6:B:401:HOH:O	2.09	0.70
1:E:35:ASN:N	6:E:501:HOH:O	2.26	0.69
2:B:6:ILE:O	6:B:402:HOH:O	2.10	0.69
1:A:256:ARG:NH1	6:A:501:HOH:O	1.88	0.68
1:E:57:LEU:HD22	1:E:102:LYS:HG2	1.75	0.68
1:C:126:ASN:ND2	1:C:129:SER:OG	2.28	0.67
1:E:290:ASN:ND2	6:E:502:HOH:O	2.25	0.65
1:E:294:GLN:HG2	1:E:296:VAL:H	1.64	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:200:ARG:HD3	1:C:215:ALA:HB2	1.82	0.61
2:B:27:GLU:OE1	6:B:403:HOH:O	2.16	0.61
1:E:294:GLN:NE2	1:E:297:SER:O	2.34	0.60
1:E:5:ILE:HD11	2:F:122:VAL:HG21	1.84	0.60
2:B:165:GLU:O	2:B:169:ASN:ND2	2.35	0.58
1:E:157:GLU:HG2	1:E:158:TYR:H	1.69	0.57
1:C:294:GLN:NE2	1:C:297:SER:O	2.38	0.57
1:E:217:ARG:HH21	1:E:225:GLY:HA2	1.71	0.56
1:A:126:ASN:ND2	1:A:129:SER:OG	2.39	0.56
1:A:157:GLU:HG2	1:A:158:TYR:H	1.71	0.55
1:E:157:GLU:HG2	1:E:158:TYR:N	2.23	0.54
1:E:181:HIS:O	1:E:217:ARG:NH2	2.40	0.54
1:C:40:ARG:NH1	1:C:274:GLU:OE1	2.39	0.54
1:C:294:GLN:HB3	1:C:305:PRO:HG2	1.91	0.52
1:A:1:ASP:N	6:A:502:HOH:O	1.95	0.52
1:C:200:ARG:NH1	1:C:243:GLU:OE1	2.38	0.52
1:A:51:ASP:HB2	1:A:273:ILE:HD12	1.92	0.52
1:A:41:PHE:CE2	1:A:271:LEU:HB2	2.45	0.51
1:A:112:GLU:HB2	1:A:256:ARG:HB2	1.93	0.51
1:C:41:PHE:CE2	1:C:271:LEU:HB2	2.47	0.50
1:A:157:GLU:HG2	1:A:158:TYR:N	2.26	0.50
2:F:26:HIS:HD2	6:F:405:HOH:O	1.95	0.49
1:A:187:GLU:OE2	4:A:406:SIA:O8	2.21	0.49
1:E:294:GLN:HB3	1:E:305:PRO:HG2	1.94	0.49
2:D:48:ILE:HD11	2:D:110:LEU:HD23	1.94	0.49
2:F:128:ASP:O	2:F:170:ARG:NH1	2.46	0.48
3:B:301:NAG:H3	3:B:301:NAG:H83	1.95	0.48
1:C:41:PHE:HB2	1:C:273:ILE:HG12	1.95	0.48
1:A:92:PRO:HG3	1:A:223:GLN:HB2	1.96	0.48
3:E:402:NAG:H61	3:E:403:NAG:O5	2.14	0.47
1:C:209:ALA:HB3	1:E:213:GLU:HB3	1.97	0.47
1:A:69:ILE:HG23	1:A:70:LEU:HG	1.96	0.47
1:E:149:TRP:HH2	4:E:405:SIA:H91	1.80	0.47
1:C:256:ARG:NH2	6:C:508:HOH:O	2.46	0.47
3:F:302:NAG:H3	3:F:303:NAG:H61	1.97	0.47
1:C:197:ARG:HD2	1:C:245:ASN:O	2.14	0.46
2:B:160:PRO:HA	2:B:163:GLN:HB2	1.96	0.46
6:E:509:HOH:O	2:F:143:LYS:HE3	2.14	0.46
1:E:282:THR:HG22	1:E:300:TRP:HB3	1.96	0.46
1:C:282:THR:HG22	1:C:300:TRP:HB3	1.97	0.46
3:C:401:NAG:H3	3:C:401:NAG:O7	2.15	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:69:ILE:HG23	1:C:70:LEU:HG	1.97	0.45
1:E:51:ASP:HB2	1:E:273:ILE:HD12	1.99	0.45
1:A:5:ILE:HD11	2:B:122:VAL:HG21	1.97	0.45
1:A:294:GLN:NE2	1:A:297:SER:O	2.50	0.45
2:F:143:LYS:HB3	2:F:143:LYS:HE2	1.71	0.45
1:A:319:LEU:HD23	1:A:320:ARG:O	2.18	0.45
1:E:171:LYS:NZ	6:E:505:HOH:O	2.37	0.44
1:E:41:PHE:CE2	1:E:271:LEU:HB2	2.52	0.44
1:E:90:CYS:HB2	1:E:134:ALA:O	2.18	0.44
1:C:49:PRO:HG2	1:C:273:ILE:HD13	2.00	0.44
1:A:217:ARG:CD	1:E:202:GLY:HA3	2.47	0.44
1:E:38:GLU:HB2	1:E:286:VAL:HB	1.99	0.44
1:A:183:PRO:HA	1:A:215:ALA:O	2.18	0.44
1:E:95:LEU:HD21	1:E:101:LEU:HD23	1.99	0.44
1:E:149:TRP:CH2	4:E:405:SIA:H91	2.53	0.44
1:A:57:LEU:HD22	1:A:102:LYS:HG2	1.99	0.44
2:F:173:ILE:HG22	2:F:174:GLU:N	2.33	0.43
1:C:280:CYS:HB2	1:C:303:GLU:O	2.19	0.43
1:C:16:VAL:HG21	1:C:316:ALA:HB2	2.01	0.43
2:F:173:ILE:HG22	2:F:174:GLU:H	1.83	0.43
1:A:78:TYR:HB3	1:A:301:ILE:HD13	2.01	0.43
1:C:177:TRP:HB3	1:C:251:PRO:HG3	2.01	0.42
1:E:197:ARG:HD2	1:E:245:ASN:O	2.18	0.42
2:D:143:LYS:HD3	2:D:143:LYS:HA	1.91	0.42
1:E:198:TYR:N	1:E:245:ASN:OD1	2.45	0.42
2:D:163:GLN:O	2:D:167:ARG:HG3	2.18	0.42
2:B:171:GLN:N	2:B:171:GLN:OE1	2.41	0.42
1:E:279:THR:O	1:E:302:GLY:HA3	2.20	0.41
1:E:95:LEU:HA	1:E:229:TYR:HB2	2.02	0.41
1:A:90:CYS:HB2	1:A:134:ALA:O	2.20	0.41
1:C:294:GLN:HE21	1:C:297:SER:H	1.67	0.41
1:C:152:LYS:HB3	1:C:158:TYR:HB2	2.03	0.41
1:E:82:ARG:HA	1:E:83:PRO:HD3	1.95	0.41
1:A:217:ARG:HD3	1:E:202:GLY:HA3	2.02	0.41
1:E:33:LEU:HD21	1:E:295:ASN:HB2	2.02	0.41
1:E:177:TRP:HB3	1:E:251:PRO:HD3	2.04	0.40
2:F:169:ASN:O	2:F:173:ILE:HD13	2.21	0.40
1:C:104:LEU:HD21	1:C:233:ILE:HD11	2.03	0.40
1:A:201:MET:HB2	1:A:208:PHE:HB3	2.04	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	322/324 (99%)	312 (97%)	10 (3%)	0	100	100
1	C	322/324 (99%)	313 (97%)	9 (3%)	0	100	100
1	E	322/324 (99%)	312 (97%)	10 (3%)	0	100	100
2	B	169/191 (88%)	165 (98%)	4 (2%)	0	100	100
2	D	171/191 (90%)	168 (98%)	3 (2%)	0	100	100
2	F	172/191 (90%)	168 (98%)	4 (2%)	0	100	100
All	All	1478/1545 (96%)	1438 (97%)	40 (3%)	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	281/281 (100%)	280 (100%)	1 (0%)	93	98
1	C	281/281 (100%)	278 (99%)	3 (1%)	80	95
1	E	281/281 (100%)	280 (100%)	1 (0%)	93	98
2	B	146/164 (89%)	146 (100%)	0	100	100
2	D	148/164 (90%)	148 (100%)	0	100	100
2	F	149/164 (91%)	149 (100%)	0	100	100
All	All	1286/1335 (96%)	1281 (100%)	5 (0%)	93	98

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	294	GLN
1	C	155	SER
1	C	294	GLN
1	C	319	LEU
1	E	294	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	87	ASN
1	C	126	ASN
1	C	294	GLN
2	D	163	GLN
1	E	324	GLN
2	F	26	HIS

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

23 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	0.42	0	15,19,21	0.28	0
3	NAG	A	402	1	14,14,15	0.33	0	15,19,21	0.34	0
3	NAG	A	403	1	14,14,15	0.45	0	15,19,21	0.55	0
3	NAG	A	404	1,3	14,14,15	0.43	0	15,19,21	0.45	0
3	NAG	A	405	3	14,14,15	0.50	0	15,19,21	1.06	2 (13%)
4	SIA	A	406	5	17,20,21	0.93	1 (5%)	18,28,31	1.63	4 (22%)
5	GAL	A	407	4	11,11,12	0.80	0	15,15,17	1.15	1 (6%)
3	NAG	B	301	3,2	14,14,15	0.31	0	15,19,21	1.38	1 (6%)
3	NAG	B	302	3	14,14,15	0.33	0	15,19,21	0.45	0
3	NAG	C	401	1	14,14,15	0.49	0	15,19,21	1.27	1 (6%)
3	NAG	C	402	1	14,14,15	0.80	1 (7%)	15,19,21	1.10	1 (6%)
3	NAG	C	403	1	13,13,15	0.32	0	13,18,21	0.24	0
3	NAG	C	404	1	14,14,15	0.45	0	15,19,21	0.75	1 (6%)
3	NAG	E	401	1	14,14,15	0.79	1 (7%)	15,19,21	0.74	1 (6%)
3	NAG	E	402	1,3	14,14,15	0.66	0	15,19,21	0.74	1 (6%)
3	NAG	E	403	3	14,14,15	1.11	1 (7%)	15,19,21	0.88	0
3	NAG	E	404	1	14,14,15	1.27	1 (7%)	15,19,21	0.85	1 (6%)
4	SIA	E	405	5	17,20,21	1.00	1 (5%)	18,28,31	1.59	3 (16%)
5	GAL	E	406	3,4	11,11,12	0.55	0	15,15,17	1.14	1 (6%)
3	NAG	E	407	5	14,14,15	0.21	0	15,19,21	0.42	0
3	NAG	F	301	3,2	14,14,15	0.82	0	15,19,21	0.85	1 (6%)
3	NAG	F	302	3	14,14,15	0.63	0	15,19,21	1.27	1 (6%)
3	NAG	F	303	3	14,14,15	0.97	2 (14%)	15,19,21	0.98	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	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	A	404	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	405	3	-	0/6/23/26	0/1/1/1
4	SIA	A	406	5	-	0/14/34/38	0/1/1/1
5	GAL	A	407	4	-	0/2/19/22	0/1/1/1
3	NAG	B	301	3,2	-	0/6/23/26	0/1/1/1
3	NAG	B	302	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	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	403	1	-	0/4/21/26	0/1/1/1
3	NAG	C	404	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	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	403	3	-	0/6/23/26	0/1/1/1
3	NAG	E	404	1	-	0/6/23/26	0/1/1/1
4	SIA	E	405	5	-	0/14/34/38	0/1/1/1
5	GAL	E	406	3,4	-	0/2/19/22	0/1/1/1
3	NAG	E	407	5	-	0/6/23/26	0/1/1/1
3	NAG	F	301	3,2	-	0/6/23/26	0/1/1/1
3	NAG	F	302	3	-	0/6/23/26	0/1/1/1
3	NAG	F	303	3	-	0/6/23/26	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	404	NAG	O5-C1	-4.26	1.36	1.43
4	A	406	SIA	C10-N5	2.15	1.42	1.34
3	F	303	NAG	O5-C1	2.16	1.47	1.43
4	E	405	SIA	C10-N5	2.18	1.42	1.34
3	E	401	NAG	C1-C2	2.32	1.55	1.52
3	C	402	NAG	O5-C1	2.70	1.48	1.43
3	F	303	NAG	C1-C2	2.72	1.56	1.52
3	E	403	NAG	O5-C1	3.59	1.49	1.43

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	E	405	SIA	O6-C2-C3	-4.45	101.05	109.77
4	A	406	SIA	C7-C6-C5	-3.55	109.17	114.06
4	A	406	SIA	O6-C2-C3	-3.38	103.15	109.77
4	E	405	SIA	C8-C7-C6	-2.31	108.22	113.08
3	F	301	NAG	O4-C4-C3	-2.30	105.17	110.36
4	E	405	SIA	C6-C5-N5	-2.19	107.32	111.06
4	A	406	SIA	C6-C5-N5	-2.09	107.48	111.06
3	E	402	NAG	C1-O5-C5	2.02	115.11	112.14
3	A	405	NAG	O5-C5-C4	2.04	113.51	110.13
3	E	404	NAG	C4-C3-C2	2.05	114.52	111.34
3	E	401	NAG	C1-O5-C5	2.17	115.33	112.14
3	A	405	NAG	C3-C4-C5	2.34	114.40	110.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	407	GAL	O5-C5-C4	2.40	114.10	110.13
3	C	404	NAG	C1-O5-C5	2.53	115.86	112.14
5	E	406	GAL	C1-O5-C5	2.65	116.03	112.14
4	A	406	SIA	O6-C6-C5	2.66	112.83	108.48
3	F	303	NAG	C1-O5-C5	3.35	117.06	112.14
3	F	302	NAG	C1-O5-C5	3.70	117.59	112.14
3	C	402	NAG	C1-O5-C5	3.81	117.74	112.14
3	C	401	NAG	C1-O5-C5	3.94	117.94	112.14
3	B	301	NAG	C2-N2-C7	4.61	129.10	123.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	406	SIA	1	0
3	B	301	NAG	1	0
3	C	401	NAG	1	0
3	E	402	NAG	1	0
3	E	403	NAG	1	0
4	E	405	SIA	2	0
3	F	302	NAG	1	0
3	F	303	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^{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	324/324 (100%)	0.06	10 (3%) 52 45	28, 44, 66, 91	0
1	C	324/324 (100%)	0.06	2 (0%) 90 87	29, 42, 64, 94	0
1	E	324/324 (100%)	0.15	12 (3%) 45 37	29, 45, 65, 90	0
2	B	171/191 (89%)	0.14	5 (2%) 55 47	27, 42, 74, 90	0
2	D	173/191 (90%)	-0.01	5 (2%) 55 47	28, 42, 68, 88	0
2	F	174/191 (91%)	0.01	4 (2%) 64 56	28, 41, 65, 95	0
All	All	1490/1545 (96%)	0.07	38 (2%) 59 52	27, 43, 68, 95	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	46	ASN	4.8
2	F	160	PRO	3.6
1	E	279	THR	3.5
1	A	11	ASN	3.5
2	B	131	ASN	3.4
2	D	147	GLU	3.3
1	E	261	ASP	3.3
1	A	222	GLY	3.0
2	F	168	LEU	3.0
1	E	277	ASP	2.9
2	F	164	LYS	2.9
1	C	168	THR	2.9
1	E	278	ALA	2.6
1	A	139	ASN	2.5
2	B	31	GLY	2.5
1	E	286	VAL	2.5
1	E	45	SER	2.5
2	D	146	ASN	2.5
1	A	258	VAL	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	162	TYR	2.4
1	A	169	GLY	2.4
1	E	224	ARG	2.4
1	A	261	ASP	2.3
2	B	168	LEU	2.3
2	D	172	LYS	2.3
2	D	31	GLY	2.3
1	E	185	THR	2.2
1	A	145	ARG	2.2
1	A	256	ARG	2.2
2	F	174	GLU	2.2
2	D	30	GLN	2.1
1	E	219	ALA	2.1
1	A	223	GLN	2.0
2	B	141	TRP	2.0
1	C	13	THR	2.0
1	A	323	PRO	2.0
1	E	43	LYS	2.0
1	E	47	LYS	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^{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	SIA	A	406	20/21	0.72	0.54	6.28	61,87,103,112	0
3	NAG	B	301	14/15	0.77	0.33	4.86	53,78,102,111	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	NAG	A	404	14/15	0.80	0.24	2.96	53,85,102,103	0
4	SIA	E	405	20/21	0.91	0.26	2.26	50,56,68,77	0
3	NAG	E	401	14/15	0.70	0.47	-	81,105,114,117	0
3	NAG	E	404	14/15	0.76	0.35	-	60,88,100,103	0
3	NAG	C	402	14/15	0.85	0.24	-	65,82,103,107	0
3	NAG	C	403	13/15	0.63	0.69	-	45,111,131,138	0
3	NAG	F	303	14/15	0.74	0.49	-	83,110,124,126	0
3	NAG	E	403	14/15	0.51	0.55	-	88,116,140,148	0
3	NAG	A	405	14/15	0.74	0.41	-	87,111,124,125	0
3	NAG	E	402	14/15	0.60	0.48	-	77,104,119,120	0
3	NAG	C	401	14/15	0.81	0.32	-	86,101,110,110	0
5	GAL	E	406	11/12	0.85	0.26	-	62,68,83,86	0
3	NAG	A	401	14/15	0.66	0.40	-	72,93,117,121	0
3	NAG	A	402	14/15	0.76	0.42	-	87,106,114,117	0
3	NAG	C	404	14/15	0.60	0.41	-	85,107,120,121	0
3	NAG	B	302	14/15	0.83	0.30	-	55,76,87,94	0
3	NAG	F	302	14/15	0.90	0.42	-	66,80,97,99	0
3	NAG	A	403	14/15	0.74	0.34	-	69,90,104,119	0
5	GAL	A	407	11/12	0.79	0.47	-	81,95,107,109	0
3	NAG	E	407	14/15	0.85	0.41	-	82,100,105,106	0
3	NAG	F	301	14/15	0.87	0.28	-	58,72,87,115	0

6.5 Other polymers (i)

There are no such residues in this entry.