



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 10:07 PM GMT

PDB ID : 4WSU
Title : The crystal structure of hemagglutinin from A/Taiwan/1/2013 in complex with 3'SLN
Authors : Yang, H.; Carney, P.J.; Chang, J.C.; Villanueva, J.M.; Stevens, J.
Deposited on : 2014-10-28
Resolution : 2.70 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 ⓘ 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 (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

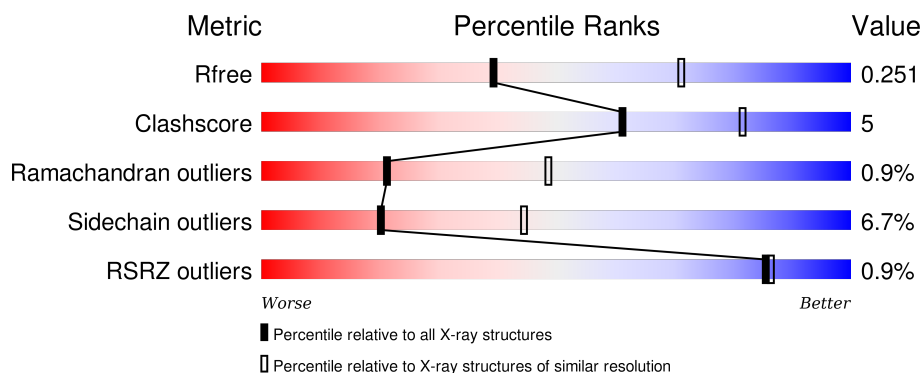
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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 $\leq 5\%$. 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	334	<div> <div>80%</div> <div>14%</div> <div>• •</div> </div>
1	C	334	<div> <div>%</div> <div>82%</div> <div>13%</div> <div>• •</div> </div>
1	E	334	<div> <div>%</div> <div>83%</div> <div>12%</div> <div>• •</div> </div>
2	B	181	<div> <div>%</div> <div>77%</div> <div>13%</div> <div>• 6%</div> </div>
2	D	181	<div> <div>%</div> <div>73%</div> <div>18%</div> <div>• • 6%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	181	 % 78% 13% 6%

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	402	X	-	-	-
4	SIA	C	404	-	-	-	X
5	GAL	A	405	X	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12006 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.

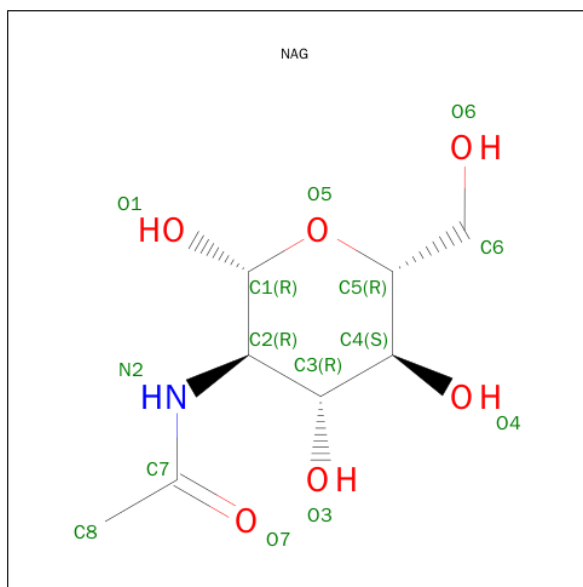
- Molecule 1 is a protein called Hemagglutinin HA1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	325	Total	C	N	O	S	0	0	0
			2566	1626	437	490	13			
1	C	325	Total	C	N	O	S	0	0	0
			2566	1626	437	490	13			
1	E	325	Total	C	N	O	S	0	0	0
			2566	1626	437	490	13			

- Molecule 2 is a protein called Hemagglutinin HA2 chain.

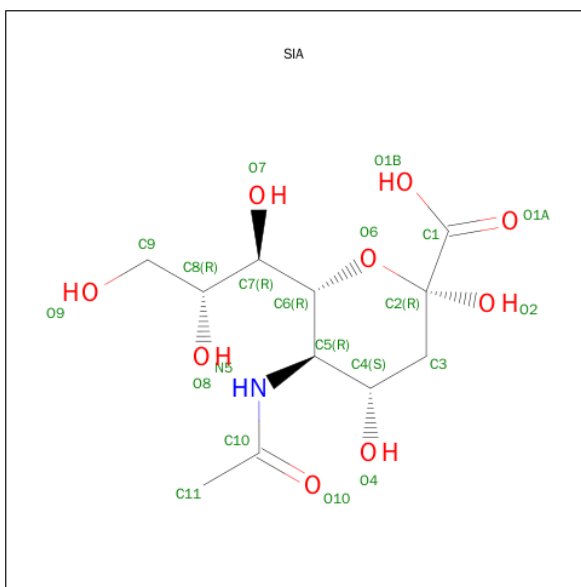
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	170	Total	C	N	O	S	0	0	0
			1372	855	240	270	7			
2	D	170	Total	C	N	O	S	0	0	0
			1372	855	240	270	7			
2	F	170	Total	C	N	O	S	0	0	0
			1372	855	240	270	7			

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C₈H₁₅NO₆).



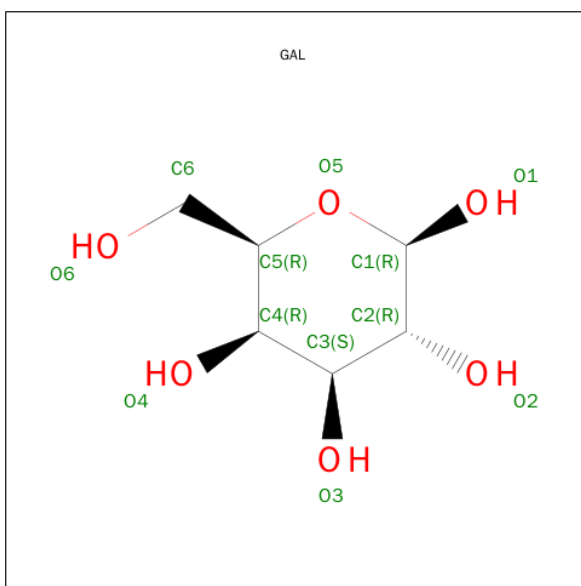
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		
3	E	1	Total	C	N	O	0	0
			14	8	1	5		
3	E	1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

- Molecule 4 is SUGAR (O-SIALIC ACID) (three-letter code: SIA) (formula: C₁₁H₁₉NO₉).



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

- Molecule 5 is SUGAR (3-MER) (three-letter code: GAL, NAG) (formula: $C_6H_{12}O_6$, $C_8H_{15}NO_6$).

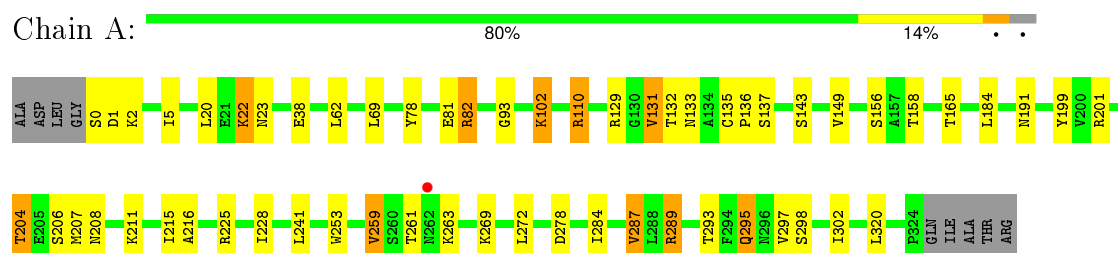


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	2	Total	C	N	O	0	0
			26	14	1	11		

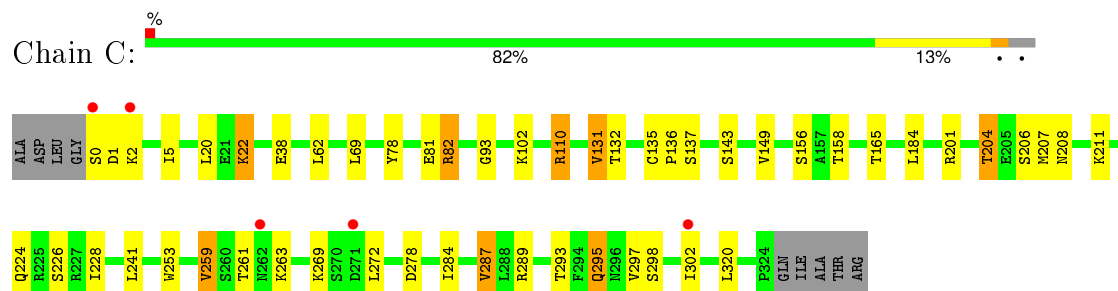
3 Residue-property plots

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 ($\text{RSRZ} > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

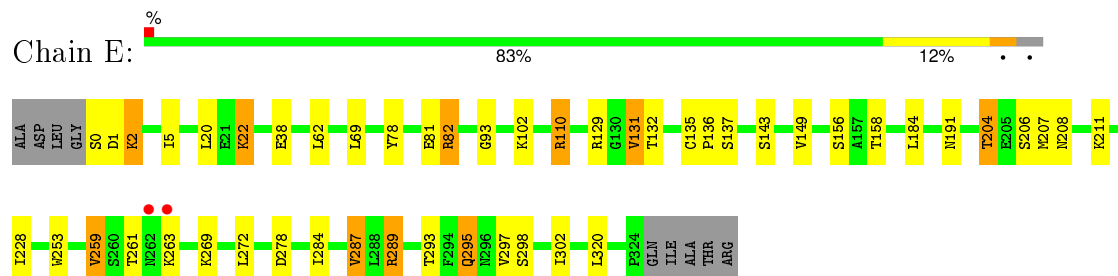
• Molecule 1: Hemagglutinin HA1 chain



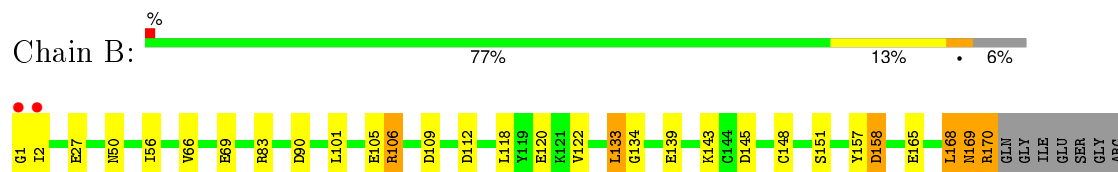
• Molecule 1: Hemagglutinin HA1 chain



• Molecule 1: Hemagglutinin HA1 chain

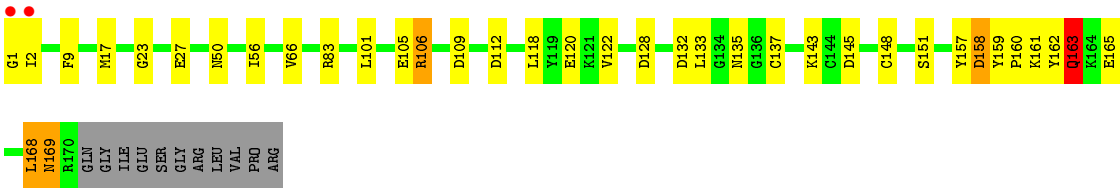


• Molecule 2: Hemagglutinin HA2 chain

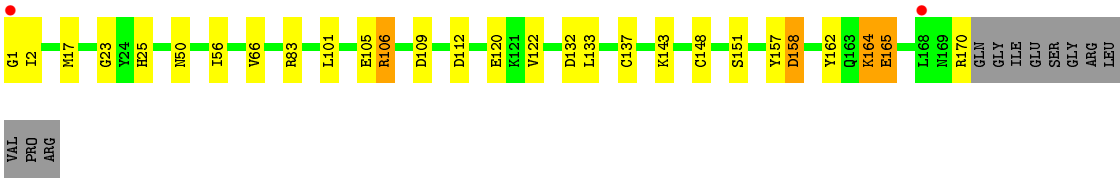
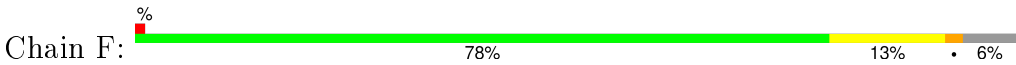


LEU
VAL
PRO
ARG

• 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, α , β , γ	186.98Å 100.08Å 135.24Å 90.00° 126.04° 90.00°	Depositor
Resolution (Å)	49.94 – 2.70 46.37 – 2.70	Depositor EDS
% Data completeness (in resolution range)	98.3 (49.94-2.70) 98.3 (46.37-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.22 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.212 , 0.244 0.221 , 0.251	Depositor DCC
R_{free} test set	2779 reflections (5.37%)	DCC
Wilson B-factor (Å ²)	48.3	Xtriage
Anisotropy	0.064	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 25.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 54578 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	12006	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 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	0.65	1/2627 (0.0%)	0.80	1/3574 (0.0%)
1	C	0.67	1/2627 (0.0%)	0.82	1/3574 (0.0%)
1	E	0.68	2/2627 (0.1%)	0.81	1/3574 (0.0%)
2	B	0.71	0/1400	0.88	5/1885 (0.3%)
2	D	0.72	0/1400	0.87	5/1885 (0.3%)
2	F	0.73	0/1400	0.86	4/1885 (0.2%)
All	All	0.68	4/12081 (0.0%)	0.83	17/16377 (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	A	0	1
1	C	0	1
1	E	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1	ASP	C-N	9.24	1.55	1.34
1	C	1	ASP	C-N	7.29	1.50	1.34
1	E	2	LYS	C-N	-6.43	1.19	1.34
1	E	1	ASP	C-N	6.32	1.48	1.34

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	102	LYS	CD-CE-NZ	8.04	130.20	111.70
1	E	102	LYS	CD-CE-NZ	7.91	129.89	111.70
1	C	102	LYS	CD-CE-NZ	7.61	129.20	111.70
2	F	83	ARG	NE-CZ-NH2	-7.24	116.68	120.30
2	B	83	ARG	NE-CZ-NH2	-6.62	116.99	120.30
2	D	163	GLN	N-CA-CB	6.42	122.16	110.60
2	B	170	ARG	NE-CZ-NH1	6.35	123.47	120.30
2	F	106	ARG	NE-CZ-NH2	-6.25	117.18	120.30
2	D	83	ARG	NE-CZ-NH2	-6.17	117.21	120.30
2	D	106	ARG	NE-CZ-NH2	-6.15	117.23	120.30
2	B	106	ARG	NE-CZ-NH2	-5.61	117.50	120.30
2	F	106	ARG	NE-CZ-NH1	5.52	123.06	120.30
2	D	106	ARG	NE-CZ-NH1	5.29	122.94	120.30
2	B	90	ASP	CB-CG-OD1	5.29	123.06	118.30
2	F	83	ARG	NE-CZ-NH1	5.06	122.83	120.30
2	B	145	ASP	CB-CG-OD1	5.02	122.81	118.30
2	D	145	ASP	CB-CG-OD1	5.01	122.81	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	156	SER	Peptide
1	C	156	SER	Peptide
1	E	156	SER	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	2566	0	2513	28	0
1	C	2566	0	2514	27	0
1	E	2566	0	2512	20	0
2	B	1372	0	1287	23	0
2	D	1372	0	1287	28	0
2	F	1372	0	1287	21	0
3	A	42	0	39	1	0
3	C	42	0	39	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	42	0	39	0	0
4	A	20	0	17	5	0
4	C	20	0	17	3	0
5	A	26	0	23	8	0
All	All	12006	0	11574	129	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 5.

All (129) 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)
4:A:404:SIA:O6	5:A:405:GAL:O4	1.96	0.84
4:A:404:SIA:O6	5:A:405:GAL:C4	2.33	0.77
1:C:82:ARG:NH2	1:C:272:LEU:O	2.20	0.75
2:D:105:GLU:CD	2:F:106:ARG:HH22	1.91	0.74
5:A:405:GAL:O5	5:A:406:NAG:H4	1.88	0.73
1:C:2:LYS:HE2	2:D:27:GLU:OE1	1.88	0.73
1:E:5:ILE:HD11	2:F:122:VAL:HG21	1.72	0.72
1:A:82:ARG:NH2	1:A:272:LEU:O	2.22	0.71
2:D:148:CYS:O	2:D:151:SER:HB3	1.91	0.71
2:B:148:CYS:O	2:B:151:SER:HB3	1.91	0.71
1:A:78:TYR:CD1	1:A:302:ILE:HD11	2.26	0.71
1:C:2:LYS:HG2	2:D:27:GLU:HB3	1.73	0.70
1:E:82:ARG:NH2	1:E:272:LEU:O	2.23	0.70
2:F:148:CYS:O	2:F:151:SER:HB3	1.91	0.70
1:E:78:TYR:CD1	1:E:302:ILE:HD11	2.27	0.70
1:C:78:TYR:CD1	1:C:302:ILE:HD11	2.29	0.68
1:A:5:ILE:HD11	2:B:122:VAL:HG21	1.76	0.68
1:C:5:ILE:HD11	2:D:122:VAL:HG21	1.77	0.67
1:C:293:THR:HG21	2:D:56:ILE:HG12	1.75	0.67
2:B:106:ARG:HH22	2:F:105:GLU:CD	1.97	0.67
2:B:1:GLY:HA3	2:B:112:ASP:OD2	1.95	0.66
2:F:132:ASP:O	2:F:137:CYS:O	2.13	0.66
2:F:164:LYS:HG3	2:F:165:GLU:H	1.60	0.66
5:A:405:GAL:O5	5:A:406:NAG:C4	2.42	0.66
2:D:132:ASP:O	2:D:137:CYS:O	2.15	0.65
1:A:278:ASP:O	1:A:289:ARG:NH2	2.28	0.65
1:C:204:THR:HG22	1:C:206:SER:H	1.61	0.64
1:A:293:THR:HG21	2:B:56:ILE:HG12	1.80	0.64
2:B:105:GLU:CD	2:D:106:ARG:HH22	2.01	0.64

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:131:VAL:HG22	1:C:143:SER:HA	1.81	0.63
1:C:278:ASP:O	1:C:289:ARG:NH2	2.30	0.63
1:A:204:THR:HG22	1:A:206:SER:H	1.64	0.63
1:E:278:ASP:O	1:E:289:ARG:NH2	2.29	0.62
1:C:22:LYS:HD3	2:F:50:ASN:HD21	1.64	0.62
1:E:131:VAL:HG22	1:E:143:SER:HA	1.82	0.62
2:D:160:PRO:O	2:D:162:TYR:O	2.17	0.62
2:F:1:GLY:HA3	2:F:112:ASP:OD2	1.99	0.62
4:A:404:SIA:O6	5:A:405:GAL:H4	2.00	0.62
2:F:2:ILE:HG12	2:F:109:ASP:OD1	2.00	0.62
2:D:1:GLY:HA3	2:D:112:ASP:OD2	2.00	0.62
1:E:204:THR:HG22	1:E:206:SER:H	1.63	0.62
4:C:404:SIA:O6	4:C:404:SIA:O8	2.00	0.61
1:A:131:VAL:HG22	1:A:143:SER:HA	1.83	0.60
2:B:2:ILE:HG12	2:B:109:ASP:OD1	2.00	0.60
2:B:133:LEU:HD11	2:B:139:GLU:HG3	1.82	0.60
4:A:404:SIA:C2	5:A:405:GAL:O4	2.50	0.59
1:E:293:THR:HG21	2:F:56:ILE:HG12	1.85	0.58
1:E:22:LYS:HD3	2:B:50:ASN:HD21	1.69	0.58
2:D:165:GLU:O	2:D:168:LEU:O	2.22	0.57
1:C:2:LYS:HE2	2:D:27:GLU:HB2	1.86	0.57
2:B:168:LEU:O	2:B:169:ASN:HB2	2.05	0.57
2:D:132:ASP:O	2:D:133:LEU:HB2	2.05	0.56
2:B:169:ASN:O	2:B:170:ARG:HB2	2.05	0.56
5:A:405:GAL:HO6	5:A:406:NAG:HO4	1.49	0.56
1:A:295:GLN:HE22	1:A:298:SER:H	1.54	0.56
1:A:38:GLU:HB2	1:A:287:VAL:HG13	1.88	0.55
2:D:157:TYR:O	2:D:158:ASP:CB	2.54	0.55
1:C:224:GLN:HE22	4:C:404:SIA:H6	1.70	0.54
2:F:157:TYR:O	2:F:158:ASP:CB	2.55	0.54
1:A:215:ILE:O	1:C:201:ARG:NE	2.40	0.54
1:E:295:GLN:HE22	1:E:298:SER:H	1.56	0.54
1:C:295:GLN:HE22	1:C:298:SER:H	1.55	0.54
2:B:157:TYR:O	2:B:158:ASP:CB	2.54	0.54
1:C:38:GLU:HB2	1:C:287:VAL:HG13	1.89	0.54
2:F:151:SER:OG	2:F:157:TYR:HA	2.08	0.53
1:E:38:GLU:HB2	1:E:287:VAL:HG13	1.90	0.53
2:B:151:SER:OG	2:B:157:TYR:HA	2.07	0.53
2:D:161:LYS:C	2:D:162:TYR:O	2.43	0.53
2:D:151:SER:OG	2:D:157:TYR:HA	2.09	0.53
2:B:165:GLU:O	2:B:168:LEU:O	2.26	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:22:LYS:CD	2:F:50:ASN:HD21	2.21	0.52
2:B:169:ASN:O	2:B:170:ARG:CB	2.58	0.52
2:B:168:LEU:O	2:B:169:ASN:CB	2.59	0.50
1:C:131:VAL:HG22	1:C:143:SER:CA	2.40	0.50
2:D:2:ILE:HG12	2:D:109:ASP:OD1	2.12	0.50
1:A:216:ALA:HB2	1:C:201:ARG:HD2	1.94	0.49
1:A:23:ASN:OD1	3:A:402:NAG:O5	2.23	0.49
1:A:131:VAL:HG22	1:A:143:SER:CA	2.43	0.49
2:D:168:LEU:O	2:D:169:ASN:HB2	2.13	0.48
1:A:225:ARG:CZ	5:A:406:NAG:H61	2.44	0.48
1:C:295:GLN:HE21	1:C:297:VAL:H	1.60	0.48
1:A:133:ASN:OD1	4:A:404:SIA:O1A	2.30	0.48
1:A:204:THR:HB	1:A:207:MET:HG3	1.96	0.47
1:C:204:THR:HB	1:C:207:MET:HG3	1.96	0.47
1:E:295:GLN:HE21	1:E:297:VAL:H	1.61	0.47
1:A:149:VAL:HG23	1:A:253:TRP:HB2	1.97	0.47
1:E:131:VAL:HG22	1:E:143:SER:CA	2.44	0.47
2:D:2:ILE:HG23	2:D:109:ASP:OD1	2.14	0.47
1:C:110:ARG:HB3	1:C:259:VAL:CG1	2.44	0.47
2:D:105:GLU:OE2	2:F:106:ARG:NH2	2.47	0.47
2:F:164:LYS:CG	2:F:165:GLU:H	2.27	0.47
2:D:168:LEU:O	2:D:169:ASN:CB	2.63	0.46
1:E:110:ARG:HB3	1:E:259:VAL:CG1	2.44	0.46
2:B:157:TYR:O	2:B:158:ASP:HB3	2.15	0.46
2:F:132:ASP:O	2:F:133:LEU:HB2	2.13	0.46
1:A:110:ARG:HB3	1:A:259:VAL:CG1	2.46	0.46
1:A:295:GLN:HE21	1:A:297:VAL:H	1.62	0.46
1:E:149:VAL:HG23	1:E:253:TRP:HB2	1.97	0.46
1:E:204:THR:HB	1:E:207:MET:HG3	1.98	0.45
1:E:22:LYS:CD	2:B:50:ASN:HD21	2.28	0.45
1:E:93:GLY:HA3	1:E:228:ILE:O	2.16	0.45
2:F:157:TYR:O	2:F:158:ASP:HB3	2.16	0.45
1:C:149:VAL:HG23	1:C:253:TRP:HB2	1.97	0.45
2:B:106:ARG:NH2	2:F:105:GLU:OE2	2.50	0.45
2:D:128:ASP:OD1	2:D:159:TYR:OH	2.23	0.45
2:D:157:TYR:O	2:D:158:ASP:HB3	2.17	0.44
1:A:81:GLU:O	1:A:269:LYS:HA	2.17	0.44
1:A:93:GLY:HA3	1:A:228:ILE:O	2.18	0.44
1:E:135:CYS:N	1:E:136:PRO:HD3	2.33	0.43
1:C:81:GLU:O	1:C:269:LYS:HA	2.19	0.43
1:C:93:GLY:HA3	1:C:228:ILE:O	2.17	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:2:ILE:HG23	2:F:109:ASP:OD1	2.19	0.43
1:E:81:GLU:O	1:E:269:LYS:HA	2.19	0.43
1:A:2:LYS:HG2	2:B:27:GLU:HB3	2.00	0.42
2:F:17:MET:HE1	2:F:23:GLY:C	2.41	0.42
1:E:2:LYS:HE3	2:F:25:HIS:CD2	2.55	0.42
1:A:135:CYS:N	1:A:136:PRO:HD3	2.35	0.41
1:C:165:THR:HA	1:C:241:LEU:O	2.20	0.41
1:A:22:LYS:HD3	2:D:50:ASN:HD21	1.84	0.41
2:D:9:PHE:O	2:D:135:ASN:HA	2.20	0.41
2:D:17:MET:HE2	2:D:23:GLY:HA3	2.02	0.41
1:C:135:CYS:N	1:C:136:PRO:HD3	2.35	0.41
1:A:102:LYS:HD2	2:B:69:GLU:OE1	2.21	0.41
1:C:226:SER:OG	4:C:404:SIA:H92	2.21	0.41
1:A:165:THR:HA	1:A:241:LEU:O	2.22	0.40
2:B:105:GLU:OE2	2:D:106:ARG:NH2	2.54	0.40
2:D:162:TYR:O	2:D:163:GLN:CB	2.70	0.40
1:A:293:THR:CG2	2:B:56:ILE:HG12	2.50	0.40
1:A:199:TYR:CD1	1:A:201:ARG:HG3	2.56	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	323/334 (97%)	310 (96%)	12 (4%)	1 (0%)	46	75
1	C	323/334 (97%)	309 (96%)	13 (4%)	1 (0%)	46	75
1	E	323/334 (97%)	310 (96%)	12 (4%)	1 (0%)	46	75
2	B	168/181 (93%)	159 (95%)	5 (3%)	4 (2%)	7	19
2	D	168/181 (93%)	159 (95%)	6 (4%)	3 (2%)	11	27
2	F	168/181 (93%)	160 (95%)	5 (3%)	3 (2%)	11	27

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1473/1545 (95%)	1407 (96%)	53 (4%)	13 (1%)	21	49

All (13) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	163	GLN
2	F	164	LYS
1	A	263	LYS
1	C	263	LYS
1	E	263	LYS
2	B	133	LEU
2	B	158	ASP
2	B	169	ASN
2	D	158	ASP
2	D	169	ASN
2	F	158	ASP
2	F	162	TYR
2	B	134	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	286/292 (98%)	262 (92%)	24 (8%)	14	30
1	C	286/292 (98%)	265 (93%)	21 (7%)	17	39
1	E	286/292 (98%)	262 (92%)	24 (8%)	14	30
2	B	145/154 (94%)	139 (96%)	6 (4%)	37	69
2	D	145/154 (94%)	139 (96%)	6 (4%)	37	69
2	F	145/154 (94%)	139 (96%)	6 (4%)	37	69
All	All	1293/1338 (97%)	1206 (93%)	87 (7%)	20	44

All (87) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	0	SER
1	A	20	LEU
1	A	22	LYS
1	A	62	LEU
1	A	69	LEU
1	A	82	ARG
1	A	110	ARG
1	A	129	ARG
1	A	131	VAL
1	A	132	THR
1	A	137	SER
1	A	158	THR
1	A	184	LEU
1	A	191	ASN
1	A	204	THR
1	A	208	ASN
1	A	211	LYS
1	A	259	VAL
1	A	261	THR
1	A	284	ILE
1	A	287	VAL
1	A	289	ARG
1	A	295	GLN
1	A	320	LEU
1	C	0	SER
1	C	20	LEU
1	C	22	LYS
1	C	62	LEU
1	C	69	LEU
1	C	82	ARG
1	C	110	ARG
1	C	131	VAL
1	C	132	THR
1	C	137	SER
1	C	158	THR
1	C	184	LEU
1	C	204	THR
1	C	208	ASN
1	C	211	LYS
1	C	259	VAL
1	C	261	THR
1	C	284	ILE
1	C	287	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	295	GLN
1	C	320	LEU
1	E	0	SER
1	E	20	LEU
1	E	22	LYS
1	E	62	LEU
1	E	69	LEU
1	E	82	ARG
1	E	110	ARG
1	E	129	ARG
1	E	131	VAL
1	E	132	THR
1	E	137	SER
1	E	158	THR
1	E	184	LEU
1	E	191	ASN
1	E	204	THR
1	E	208	ASN
1	E	211	LYS
1	E	259	VAL
1	E	261	THR
1	E	284	ILE
1	E	287	VAL
1	E	289	ARG
1	E	295	GLN
1	E	320	LEU
2	B	66	VAL
2	B	101	LEU
2	B	118	LEU
2	B	120	GLU
2	B	143	LYS
2	B	168	LEU
2	D	66	VAL
2	D	101	LEU
2	D	118	LEU
2	D	120	GLU
2	D	143	LYS
2	D	168	LEU
2	F	66	VAL
2	F	101	LEU
2	F	120	GLU
2	F	143	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	F	165	GLU
2	F	170	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	36	GLN
1	A	208	ASN
1	A	224	GLN
1	A	295	GLN
1	C	36	GLN
1	C	208	ASN
1	C	224	GLN
1	C	295	GLN
1	E	36	GLN
1	E	295	GLN
2	B	50	ASN
2	B	114	ASN
2	D	50	ASN
2	D	114	ASN
2	F	50	ASN
2	F	114	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

11 non-standard protein/DNA/RNA residues 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	0.27	0	15,19,21	0.53	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	NAG	A	402	1	14,14,15	0.28	0	15,19,21	0.53	0
3	NAG	A	403	1	14,14,15	0.59	0	15,19,21	1.93	1 (6%)
5	GAL	A	405	5	11,11,12	0.26	0	14,15,17	0.61	0
5	NAG	A	406	5	15,15,15	0.93	0	17,21,21	2.13	2 (11%)
3	NAG	C	401	1	14,14,15	0.45	0	15,19,21	2.64	2 (13%)
3	NAG	C	402	1	14,14,15	0.27	0	15,19,21	0.53	0
3	NAG	C	403	1	14,14,15	0.70	0	15,19,21	2.69	6 (40%)
3	NAG	E	400	1	14,14,15	0.66	0	15,19,21	1.64	5 (33%)
3	NAG	E	401	1	14,14,15	0.27	0	15,19,21	0.53	0
3	NAG	E	402	1	14,14,15	0.79	0	15,19,21	2.83	5 (33%)

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	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	403	1	-	0/6/23/26	0/1/1/1
5	GAL	A	405	5	1/1/4/5	0/2/19/22	0/1/1/1
5	NAG	A	406	5	-	0/6/26/26	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	403	1	-	0/6/23/26	0/1/1/1
3	NAG	E	400	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

There are no bond length outliers.

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	402	NAG	O3-C3-C2	-3.95	101.28	109.11
3	C	403	NAG	C6-C5-C4	-2.91	105.83	113.02
3	E	400	NAG	C3-C4-C5	-2.61	105.65	110.20
3	E	402	NAG	O7-C7-C8	-2.11	118.19	122.06
3	E	400	NAG	O4-C4-C5	2.02	114.60	109.24
3	E	400	NAG	O3-C3-C2	2.44	113.94	109.11
3	E	400	NAG	C1-O5-C5	2.48	115.39	112.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	403	NAG	O6-C6-C5	2.51	119.63	111.33
3	E	400	NAG	O5-C5-C6	2.71	113.22	107.35
3	C	403	NAG	O5-C5-C6	2.94	113.72	107.35
3	C	401	NAG	C3-C4-C5	2.95	115.34	110.20
3	E	402	NAG	C4-C3-C2	3.30	116.36	111.23
3	C	403	NAG	C4-C3-C2	3.34	116.42	111.23
3	E	402	NAG	C3-C4-C5	3.44	116.20	110.20
3	C	403	NAG	C3-C4-C5	3.55	116.39	110.20
5	A	406	NAG	C3-C4-C5	5.31	119.45	110.20
5	A	406	NAG	O3-C3-C2	5.41	120.68	109.66
3	A	403	NAG	C1-O5-C5	6.72	120.78	112.25
3	C	403	NAG	C1-O5-C5	7.69	122.00	112.25
3	E	402	NAG	C1-O5-C5	8.09	122.52	112.25
3	C	401	NAG	C1-O5-C5	9.29	124.04	112.25

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	405	GAL	C1
3	A	402	NAG	C1

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	402	NAG	1	0
5	A	405	GAL	7	0
5	A	406	NAG	4	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

13 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	0.27	0	15,19,21	0.53	0
3	NAG	A	402	1	14,14,15	0.28	0	15,19,21	0.53	0
3	NAG	A	403	1	14,14,15	0.59	0	15,19,21	1.93	1 (6%)
4	SIA	A	404	-	16,20,21	0.97	1 (6%)	18,28,31	1.85	3 (16%)
5	GAL	A	405	5	11,11,12	0.26	0	14,15,17	0.61	0
5	NAG	A	406	5	15,15,15	0.93	0	17,21,21	2.13	2 (11%)
3	NAG	C	401	1	14,14,15	0.45	0	15,19,21	2.64	2 (13%)
3	NAG	C	402	1	14,14,15	0.27	0	15,19,21	0.53	0
3	NAG	C	403	1	14,14,15	0.70	0	15,19,21	2.69	6 (40%)
4	SIA	C	404	-	16,20,21	0.94	1 (6%)	18,28,31	1.89	7 (38%)
3	NAG	E	400	1	14,14,15	0.66	0	15,19,21	1.64	5 (33%)
3	NAG	E	401	1	14,14,15	0.27	0	15,19,21	0.53	0
3	NAG	E	402	1	14,14,15	0.79	0	15,19,21	2.83	5 (33%)

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	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	403	1	-	0/6/23/26	0/1/1/1
4	SIA	A	404	-	-	0/14/34/38	0/1/1/1
5	GAL	A	405	5	1/1/4/5	0/2/19/22	0/1/1/1
5	NAG	A	406	5	-	0/6/26/26	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	403	1	-	0/6/23/26	0/1/1/1
4	SIA	C	404	-	-	0/14/34/38	0/1/1/1
3	NAG	E	400	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 (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	404	SIA	C3-C2	2.32	1.56	1.52
4	C	404	SIA	C7-C6	2.60	1.56	1.52

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	404	SIA	C6-C5-N5	-4.81	102.69	111.07
3	E	402	NAG	O3-C3-C2	-3.95	101.28	109.11
4	A	404	SIA	O6-C2-C3	-3.78	102.59	109.86
4	C	404	SIA	O4-C4-C3	-3.18	102.27	110.06
3	C	403	NAG	C6-C5-C4	-2.91	105.83	113.02
3	E	400	NAG	C3-C4-C5	-2.61	105.65	110.20
4	C	404	SIA	O10-C10-C11	-2.33	117.79	122.06
3	E	402	NAG	O7-C7-C8	-2.11	118.19	122.06
3	E	400	NAG	O4-C4-C5	2.02	114.60	109.24
4	C	404	SIA	C4-C5-N5	2.07	114.90	110.41
4	C	404	SIA	O7-C7-C6	2.31	114.78	109.43
3	E	400	NAG	O3-C3-C2	2.44	113.94	109.11
3	E	400	NAG	C1-O5-C5	2.48	115.39	112.25
3	C	403	NAG	O6-C6-C5	2.51	119.63	111.33
4	A	404	SIA	C11-C10-N5	2.54	120.96	116.11
4	C	404	SIA	C3-C4-C5	2.70	114.49	111.47
3	E	400	NAG	O5-C5-C6	2.71	113.22	107.35
4	C	404	SIA	C11-C10-N5	2.88	121.63	116.11
3	C	403	NAG	O5-C5-C6	2.94	113.72	107.35
3	C	401	NAG	C3-C4-C5	2.95	115.34	110.20
4	C	404	SIA	O6-C2-C3	3.28	116.17	109.86
3	E	402	NAG	C4-C3-C2	3.30	116.36	111.23
3	C	403	NAG	C4-C3-C2	3.34	116.42	111.23
3	E	402	NAG	C3-C4-C5	3.44	116.20	110.20
3	C	403	NAG	C3-C4-C5	3.55	116.39	110.20
5	A	406	NAG	C3-C4-C5	5.31	119.45	110.20
5	A	406	NAG	O3-C3-C2	5.41	120.68	109.66
3	A	403	NAG	C1-O5-C5	6.72	120.78	112.25
3	C	403	NAG	C1-O5-C5	7.69	122.00	112.25
3	E	402	NAG	C1-O5-C5	8.09	122.52	112.25
3	C	401	NAG	C1-O5-C5	9.29	124.04	112.25

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	A	405	GAL	C1
3	A	402	NAG	C1

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	402	NAG	1	0
4	A	404	SIA	5	0
5	A	405	GAL	7	0
5	A	406	NAG	4	0
4	C	404	SIA	3	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	E	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	2:LYS	C	3:ILE	N	1.19

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, 95th 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	325/334 (97%)	-0.22	1 (0%) 94 95	29, 50, 73, 118	0
1	C	325/334 (97%)	-0.38	5 (1%) 76 76	29, 48, 72, 119	0
1	E	325/334 (97%)	-0.35	2 (0%) 90 91	30, 47, 72, 135	0
2	B	170/181 (93%)	-0.22	2 (1%) 81 81	29, 42, 65, 86	0
2	D	170/181 (93%)	-0.20	2 (1%) 81 81	23, 41, 61, 79	0
2	F	170/181 (93%)	-0.13	2 (1%) 81 81	30, 44, 73, 96	0
All	All	1485/1545 (96%)	-0.27	14 (0%) 85 86	23, 46, 71, 135	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1	GLY	3.4
2	F	1	GLY	3.4
1	C	262	ASN	3.2
2	D	1	GLY	2.6
1	E	263	LYS	2.6
1	C	302	ILE	2.5
1	E	262	ASN	2.5
1	C	0	SER	2.3
1	C	2	LYS	2.3
2	D	2	ILE	2.2
1	A	262	ASN	2.2
1	C	271	ASP	2.1
2	B	2	ILE	2.1
2	F	168	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [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, 95th 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	E	402	14/15	0.91	0.17	1.02	51,57,65,67	0
5	NAG	A	406	15/15	0.82	0.23	0.99	78,102,120,121	0
3	NAG	A	403	14/15	0.89	0.19	-	62,67,73,74	0
3	NAG	C	402	14/15	0.80	0.17	-	52,80,93,98	0
5	GAL	A	405	11/12	0.81	0.20	-	66,78,93,98	0
3	NAG	A	401	14/15	0.73	0.28	-	83,110,121,126	0
3	NAG	C	401	14/15	0.84	0.21	-	78,92,98,104	0
3	NAG	E	401	14/15	0.65	0.25	-	55,72,81,84	0
3	NAG	A	402	14/15	0.86	0.19	-	61,88,94,94	0
3	NAG	E	400	14/15	0.76	0.22	-	67,90,101,105	0
3	NAG	C	403	14/15	0.85	0.17	-	60,73,78,78	0

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, 95th 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	C	404	20/21	0.71	0.41	12.18	64,100,120,122	0
4	SIA	A	404	20/21	0.94	0.21	1.09	44,54,62,65	0
3	NAG	E	402	14/15	0.91	0.17	1.02	51,57,65,67	0
5	NAG	A	406	15/15	0.82	0.23	0.99	78,102,120,121	0
5	GAL	A	405	11/12	0.81	0.20	-	66,78,93,98	0
3	NAG	A	401	14/15	0.73	0.28	-	83,110,121,126	0
3	NAG	C	401	14/15	0.84	0.21	-	78,92,98,104	0
3	NAG	E	401	14/15	0.65	0.25	-	55,72,81,84	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	NAG	A	402	14/15	0.86	0.19	-	61,88,94,94	0
3	NAG	A	403	14/15	0.89	0.19	-	62,67,73,74	0
3	NAG	C	402	14/15	0.80	0.17	-	52,80,93,98	0
3	NAG	E	400	14/15	0.76	0.22	-	67,90,101,105	0
3	NAG	C	403	14/15	0.85	0.17	-	60,73,78,78	0

6.5 Other polymers [i](#)

There are no such residues in this entry.