



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 15, 2017 – 12:51 am GMT

PDB ID : 4CQW
Title : H5 (tyTy) Del133/Ile155Thr Mutant Haemagglutinin in Complex with Avian Receptor Analogue 3'SLN
Authors : Xiong, X.; Xiao, H.; Martin, S.R.; Coombs, P.J.; Liu, J.; Collins, P.J.; Vachieri, S.G.; Walker, P.A.; Lin, Y.P.; McCauley, J.W.; Gamblin, S.J.; Skehel, J.J.
Deposited on : 2014-02-21
Resolution : 2.30 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

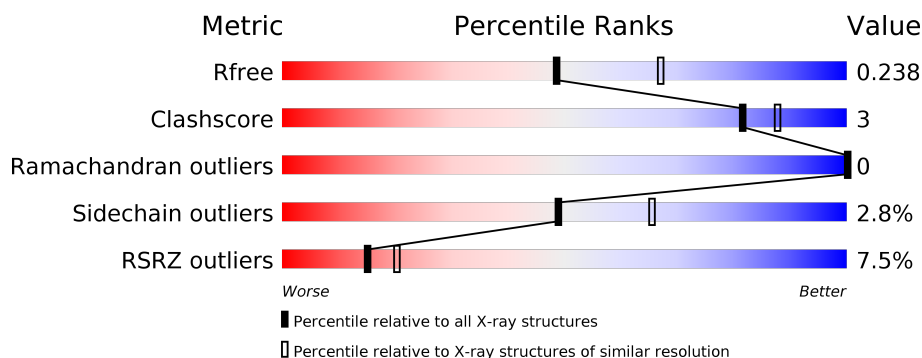
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	4130 (2.30-2.30)
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)
RSRZ outliers	101464	4156 (2.30-2.30)

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	327	<div> <div>89%</div> <div>9%</div> <div>•</div> </div>
1	C	327	<div> <div>91%</div> <div>7%</div> <div>•</div> </div>
1	E	327	<div> <div>90%</div> <div>8%</div> <div>•</div> </div>
2	B	166	<div> <div>82%</div> <div>16%</div> <div>•</div> </div>
2	D	166	<div> <div>25%</div> <div>84%</div> <div>14%</div> <div>••</div> </div>
2	F	166	<div> <div>30%</div> <div>90%</div> <div>7%</div> <div>••</div> </div>

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	1023	X	-	-	-
3	NAG	E	1011	-	-	-	X
5	PO4	A	1325	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12709 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 HAEMAGGLUTININ HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	324	Total	C	N	O	S	0	0	0
			2572	1618	449	491	14			
1	C	322	Total	C	N	O	S	0	0	0
			2552	1607	443	488	14			
1	E	322	Total	C	N	O	S	0	0	0
			2552	1607	443	488	14			

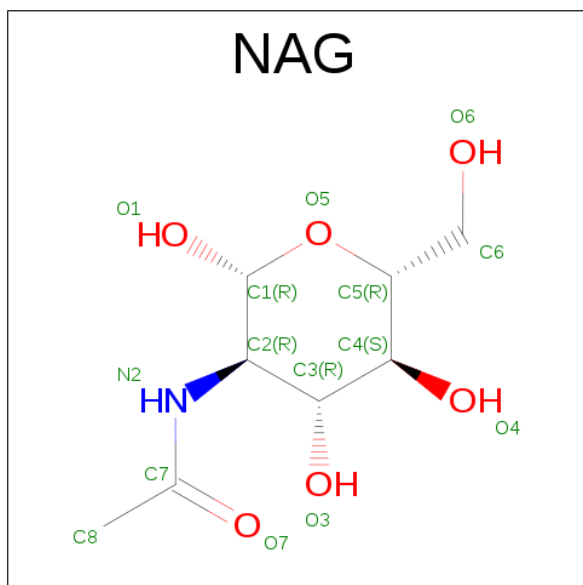
There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ASP	-	EXPRESSION TAG	UNP Q207Z6
A	0	PRO	-	EXPRESSION TAG	UNP Q207Z6
A	.	-	ALA	DELETION	UNP Q207Z6
A	150	THR	ILE	ENGINEERED MUTATION	UNP Q207Z6
A	322	ARG	GLY	CONFLICT	UNP Q207Z6
A	324	THR	ARG	CONFLICT	UNP Q207Z6
C	-1	ASP	-	EXPRESSION TAG	UNP Q207Z6
C	0	PRO	-	EXPRESSION TAG	UNP Q207Z6
C	.	-	ALA	DELETION	UNP Q207Z6
C	150	THR	ILE	ENGINEERED MUTATION	UNP Q207Z6
C	322	ARG	GLY	CONFLICT	UNP Q207Z6
C	324	THR	ARG	CONFLICT	UNP Q207Z6
E	-1	ASP	-	EXPRESSION TAG	UNP Q207Z6
E	0	PRO	-	EXPRESSION TAG	UNP Q207Z6
E	.	-	ALA	DELETION	UNP Q207Z6
E	150	THR	ILE	ENGINEERED MUTATION	UNP Q207Z6
E	322	ARG	GLY	CONFLICT	UNP Q207Z6
E	324	THR	ARG	CONFLICT	UNP Q207Z6

- Molecule 2 is a protein called HAEMAGGLUTININ HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	163	Total	C	N	O	S	0	0	0
			1320	822	229	261	8			
2	D	163	Total	C	N	O	S	0	0	0
			1320	822	229	261	8			
2	F	163	Total	C	N	O	S	0	0	0
			1320	822	229	261	8			

- 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		
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		

Continued on next page...

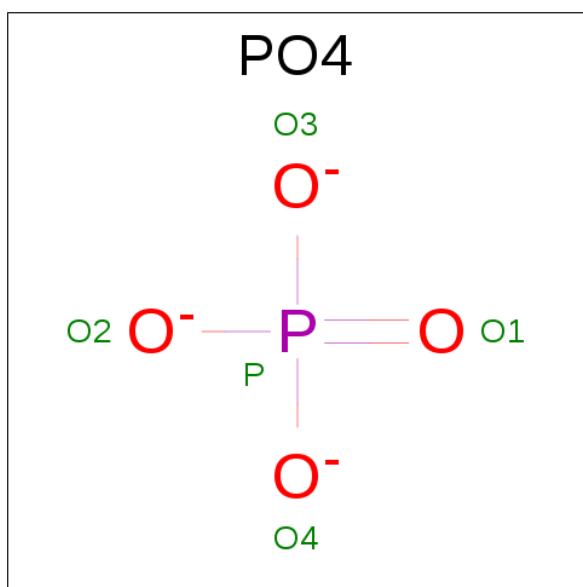
Continued from previous page...

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

- Molecule 4 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	3	Total	C	N	O	0	0
			46	25	2	19		
4	C	3	Total	C	N	O	0	0
			46	25	2	19		
4	E	3	Total	C	N	O	0	0
			46	25	2	19		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	P	0	0
			5	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	197	Total	O	0	0
			197	197		
6	B	64	Total	O	0	0
			64	64		

Continued on next page...

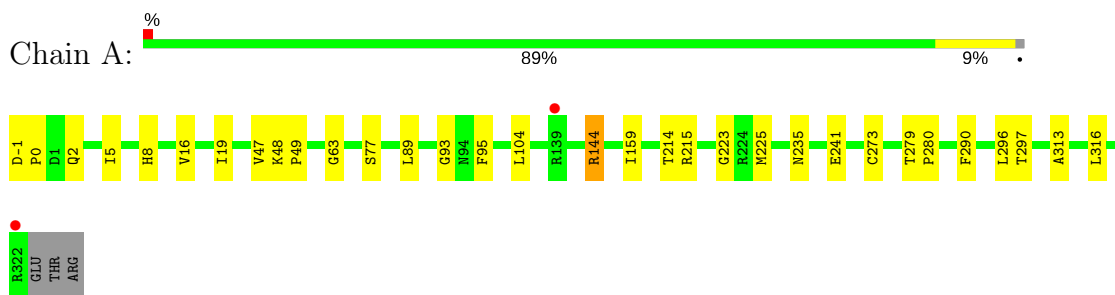
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	C	262	Total 262	O 262	0	0
6	D	30	Total 30	O 30	0	0
6	E	225	Total 225	O 225	0	0
6	F	26	Total 26	O 26	0	0

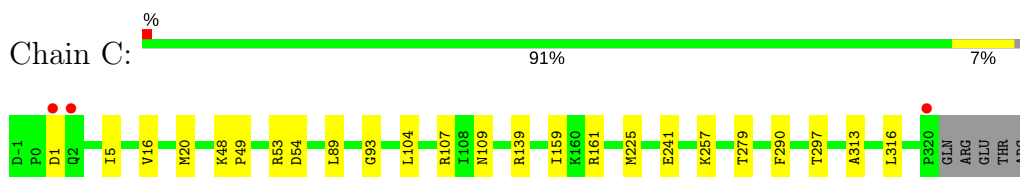
3 Residue-property plots [i](#)

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

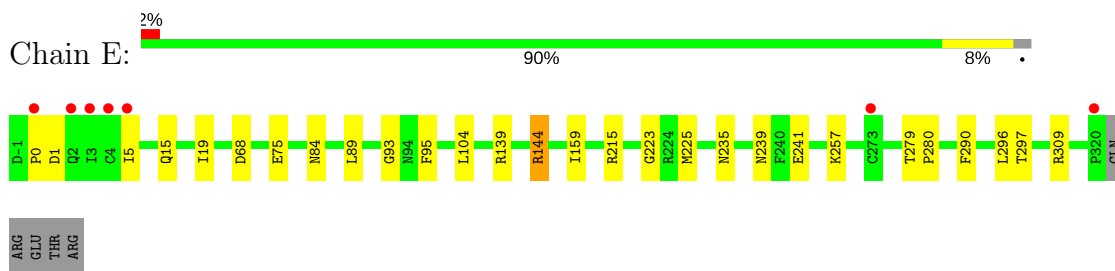
- Molecule 1: HAEMAGGLUTININ HA1



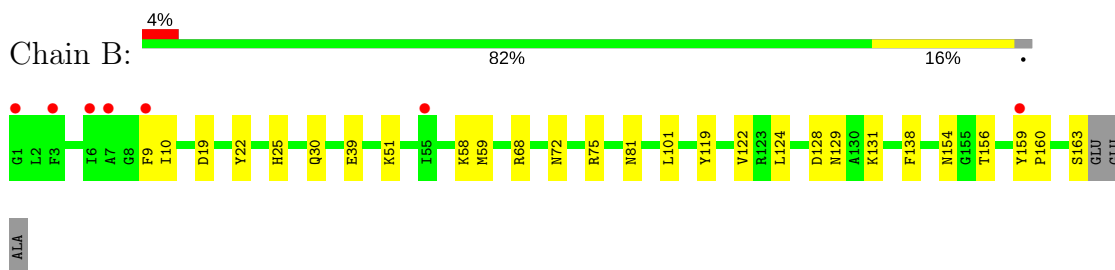
- Molecule 1: HAEMAGGLUTININ HA1



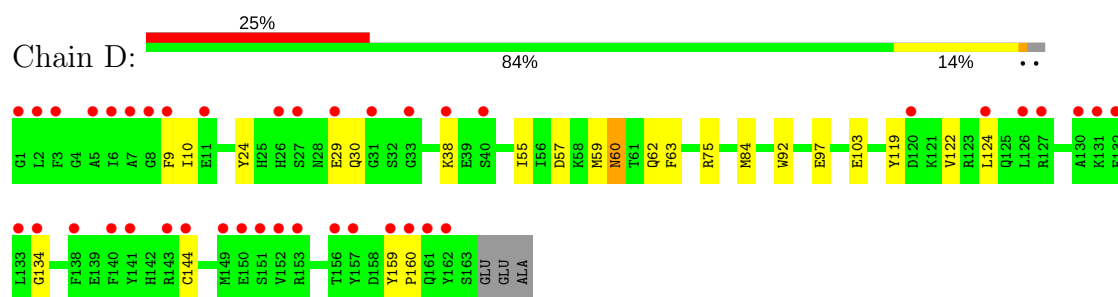
- Molecule 1: HAEMAGGLUTININ HA1



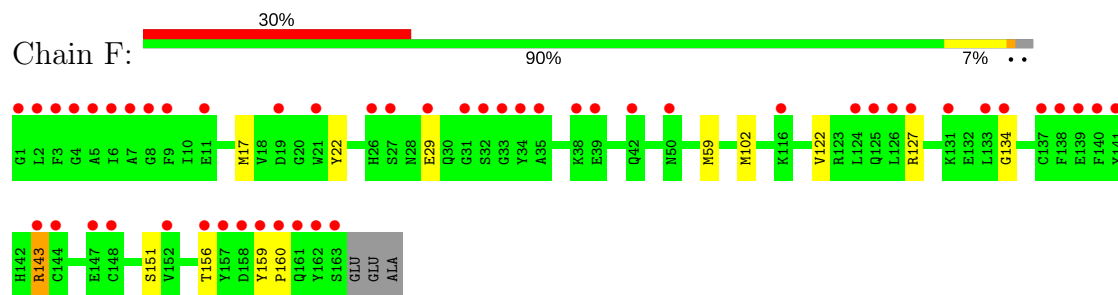
- Molecule 2: HAEMAGGLUTININ HA2



- Molecule 2: HAEMAGGLUTININ HA2



• Molecule 2: HAEMAGGLUTININ HA2



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	87.64Å 117.55Å 101.16Å 90.00° 92.50° 90.00°	Depositor
Resolution (Å)	101.06 – 2.30 87.55 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.3 (101.06-2.30) 98.3 (87.55-2.30)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.05 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0046	Depositor
R, R_{free}	0.196 , 0.238 0.196 , 0.238	Depositor DCC
R_{free} test set	4479 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	35.4	Xtriage
Anisotropy	0.592	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 42.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.024 for h,-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	12709	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.52% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, GAL, NAG, SIA

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	0.32	0/2634	0.53	0/3580
1	C	0.34	0/2614	0.53	0/3554
1	E	0.33	0/2614	0.54	0/3554
2	B	0.34	0/1347	0.49	0/1813
2	D	0.32	0/1347	0.48	0/1813
2	F	0.32	0/1347	0.47	0/1813
All	All	0.33	0/11903	0.51	0/16127

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2572	0	2504	18	0
1	C	2552	0	2483	12	0
1	E	2552	0	2483	17	0
2	B	1320	0	1224	17	0
2	D	1320	0	1224	14	0
2	F	1320	0	1224	8	0
3	A	42	0	39	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	42	0	39	1	0
3	E	42	0	39	4	0
4	A	46	0	40	0	0
4	C	46	0	40	0	0
4	E	46	0	40	0	0
5	A	5	0	0	0	0
6	A	197	0	0	0	0
6	B	64	0	0	2	0
6	C	262	0	0	0	0
6	D	30	0	0	0	0
6	E	225	0	0	0	0
6	F	26	0	0	0	0
All	All	12709	0	11379	70	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 3.

All (70) 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:E:19:ILE:HD11	2:F:102:MET:HA	1.57	0.85
3:E:1023:NAG:H83	3:E:1023:NAG:H3	1.66	0.77
1:E:290:PHE:HZ	2:F:59:MET:HG3	1.57	0.69
1:C:290:PHE:HZ	2:D:59:MET:HG3	1.60	0.66
1:E:0:PRO:HA	2:F:143:ARG:HH22	1.60	0.66
1:E:1:ASP:H	2:F:143:ARG:HH12	1.44	0.66
1:A:19:ILE:HD12	2:B:101:LEU:HB3	1.78	0.65
1:C:279:THR:HG22	1:C:297:THR:HG22	1.77	0.65
2:B:131:LYS:HE2	2:F:127:ARG:HH21	1.63	0.63
1:A:290:PHE:HZ	2:B:59:MET:HG3	1.65	0.62
1:C:5:ILE:HD11	2:D:122:VAL:HG21	1.83	0.61
2:B:19:ASP:HB2	6:B:2005:HOH:O	2.01	0.61
3:E:1023:NAG:H3	3:E:1023:NAG:C8	2.31	0.61
2:D:57:ASP:O	2:D:60:ASN:HB2	2.03	0.59
1:E:279:THR:HG22	1:E:297:THR:HG22	1.84	0.58
1:C:48:LYS:HG3	1:C:49:PRO:HD2	1.86	0.57
1:A:279:THR:HG22	1:A:297:THR:HG22	1.86	0.57
2:D:159:TYR:N	2:D:160:PRO:HD2	2.22	0.55
1:C:159:ILE:O	1:C:241:GLU:HA	2.07	0.54
2:B:58:LYS:HE3	2:D:97:GLU:HB3	1.90	0.54
2:B:156:THR:HG22	2:B:156:THR:O	2.09	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:25:HIS:HE1	6:B:2015:HOH:O	1.91	0.52
1:E:19:ILE:HD12	1:E:19:ILE:H	1.74	0.52
1:E:280:PRO:HD3	1:E:296:LEU:O	2.10	0.51
1:E:68:ASP:OD2	1:E:144:ARG:NH1	2.44	0.50
1:E:5:ILE:HD11	2:F:122:VAL:HG21	1.92	0.50
1:A:16:VAL:HG21	1:A:313:ALA:HB2	1.93	0.49
1:C:16:VAL:HG21	1:C:313:ALA:HB2	1.94	0.48
2:D:62:GLN:HG2	2:D:92:TRP:CE3	2.48	0.48
1:E:235:ASN:HD22	3:E:1164:NAG:H5	1.78	0.48
1:A:159:ILE:O	1:A:241:GLU:HA	2.14	0.47
1:A:-1:ASP:N	1:A:0:PRO:HD2	2.30	0.47
1:E:215:ARG:HH21	1:E:223:GLY:HA2	1.80	0.47
3:A:1023:NAG:H83	3:A:1023:NAG:H3	1.97	0.46
1:A:235:ASN:HD22	3:A:1164:NAG:H5	1.79	0.46
1:C:109:ASN:HB2	1:C:257:LYS:HD3	1.96	0.46
1:C:290:PHE:HE1	2:D:62:GLN:OE1	1.97	0.46
2:B:159:TYR:HB3	2:B:160:PRO:HD3	1.97	0.46
2:F:159:TYR:HB3	2:F:160:PRO:HD3	1.98	0.46
2:D:55:ILE:HD11	2:D:103:GLU:HG3	1.98	0.45
1:A:93:GLY:HA3	1:A:225:MET:O	2.17	0.45
1:A:2:GLN:HB2	2:B:138:PHE:O	2.17	0.45
1:C:5:ILE:HG13	2:D:119:TYR:HA	1.99	0.45
2:D:9:PHE:CD1	2:D:10:ILE:HG13	2.52	0.44
1:A:280:PRO:HD3	1:A:296:LEU:O	2.18	0.44
1:C:93:GLY:HA3	1:C:225:MET:O	2.18	0.44
1:A:215:ARG:HH21	1:A:223:GLY:HA2	1.83	0.43
2:B:124:LEU:HD22	2:D:134:GLY:HA2	2.00	0.43
1:A:47:VAL:HB	1:A:77:SER:HB3	2.01	0.43
2:B:68:ARG:HD2	2:B:81:ASN:OD1	2.19	0.43
1:E:235:ASN:ND2	3:E:1164:NAG:H5	2.33	0.43
1:E:159:ILE:O	1:E:241:GLU:HA	2.19	0.43
1:A:5:ILE:HG13	2:B:119:TYR:HA	2.00	0.42
2:B:72:ASN:OD1	2:B:75:ARG:NH2	2.53	0.42
2:B:129:ASN:OD1	2:B:163:SER:HB3	2.19	0.42
1:A:63:GLY:O	1:A:144:ARG:HD2	2.20	0.42
2:D:63:PHE:CZ	2:D:84:MET:CE	3.02	0.42
1:E:93:GLY:HA3	1:E:225:MET:O	2.20	0.42
1:E:15:GLN:O	1:E:309:ARG:NH2	2.53	0.42
1:E:75:GLU:OE2	1:E:257:LYS:HE2	2.19	0.42
1:A:48:LYS:HG2	1:A:49:PRO:HD2	2.02	0.41
2:D:124:LEU:HD22	2:F:134:GLY:HA2	2.01	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:53:ARG:HB3	1:C:54:ASP:H	1.63	0.41
2:D:62:GLN:HG2	2:D:92:TRP:CD2	2.56	0.41
1:A:214:THR:O	1:E:239:ASN:HB2	2.21	0.41
2:B:9:PHE:CD1	2:B:10:ILE:HG13	2.55	0.41
3:C:1023:NAG:H3	3:C:1023:NAG:C8	2.51	0.41
2:B:51:LYS:HE3	1:C:20:MET:HE2	2.02	0.41
1:A:5:ILE:HD11	2:B:122:VAL:HG21	2.03	0.41
1:A:8:HIS:O	1:A:316:LEU:HD11	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	322/327 (98%)	312 (97%)	10 (3%)	0	100	100
1	C	320/327 (98%)	308 (96%)	12 (4%)	0	100	100
1	E	320/327 (98%)	311 (97%)	9 (3%)	0	100	100
2	B	161/166 (97%)	156 (97%)	5 (3%)	0	100	100
2	D	161/166 (97%)	155 (96%)	6 (4%)	0	100	100
2	F	161/166 (97%)	158 (98%)	3 (2%)	0	100	100
All	All	1445/1479 (98%)	1400 (97%)	45 (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	291/294 (99%)	286 (98%)	5 (2%)	66	81
1	C	289/294 (98%)	282 (98%)	7 (2%)	54	72
1	E	289/294 (98%)	283 (98%)	6 (2%)	59	76
2	B	139/141 (99%)	134 (96%)	5 (4%)	40	55
2	D	139/141 (99%)	132 (95%)	7 (5%)	28	39
2	F	139/141 (99%)	133 (96%)	6 (4%)	33	45
All	All	1286/1305 (98%)	1250 (97%)	36 (3%)	49	65

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89	LEU
1	A	95	PHE
1	A	104	LEU
1	A	144	ARG
1	A	273	CYS
2	B	22	TYR
2	B	30	GLN
2	B	39	GLU
2	B	128	ASP
2	B	154	ASN
1	C	1	ASP
1	C	89	LEU
1	C	104	LEU
1	C	107	ARG
1	C	139	ARG
1	C	161	ARG
1	C	316	LEU
2	D	24	TYR
2	D	29	GLU
2	D	30	GLN
2	D	38	LYS
2	D	60	ASN
2	D	75	ARG
2	D	144	CYS
1	E	84	ASN
1	E	89	LEU
1	E	95	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	104	LEU
1	E	139	ARG
1	E	144	ARG
2	F	17	MET
2	F	22	TYR
2	F	29	GLU
2	F	143	ARG
2	F	151	SER
2	F	156	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	103	HIS
1	A	235	ASN
2	B	154	ASN
1	C	110	HIS
2	D	60	ASN
1	E	235	ASN
2	F	30	GLN

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 ⓘ

9 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	SIA	A	1321	4	17,20,21	0.39	0	19,28,31	0.94	1 (5%)
4	GAL	A	1322	4	11,11,12	0.33	0	13,15,17	1.07	1 (7%)
4	NAG	A	1323	4	15,15,15	0.46	0	21,21,21	1.29	2 (9%)
4	SIA	C	1321	4	17,20,21	0.40	0	19,28,31	0.93	1 (5%)
4	GAL	C	1322	4	11,11,12	0.30	0	13,15,17	0.83	0
4	NAG	C	1323	4	15,15,15	0.43	0	21,21,21	2.01	4 (19%)
4	SIA	E	1321	4	17,20,21	0.49	0	19,28,31	1.01	2 (10%)
4	GAL	E	1322	4	11,11,12	0.35	0	13,15,17	1.25	3 (23%)
4	NAG	E	1323	4	15,15,15	0.45	0	21,21,21	1.23	2 (9%)

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	SIA	A	1321	4	-	0/14/34/38	0/1/1/1
4	GAL	A	1322	4	-	0/2/19/22	0/1/1/1
4	NAG	A	1323	4	-	0/6/26/26	0/1/1/1
4	SIA	C	1321	4	-	0/14/34/38	0/1/1/1
4	GAL	C	1322	4	-	0/2/19/22	0/1/1/1
4	NAG	C	1323	4	-	0/6/26/26	0/1/1/1
4	SIA	E	1321	4	-	0/14/34/38	0/1/1/1
4	GAL	E	1322	4	-	0/2/19/22	0/1/1/1
4	NAG	E	1323	4	-	0/6/26/26	0/1/1/1

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	1323	NAG	C1-C2-N2	-6.05	103.72	110.73
4	A	1323	NAG	C1-C2-N2	-4.79	105.18	110.73
4	C	1323	NAG	C3-C2-N2	-4.32	102.33	110.61
4	E	1323	NAG	C1-C2-N2	-3.57	106.59	110.73
4	E	1321	SIA	C4-C5-N5	-3.18	103.85	110.40
4	C	1321	SIA	C4-C5-N5	-2.69	104.85	110.40
4	A	1321	SIA	C4-C5-N5	-2.53	105.19	110.40
4	A	1323	NAG	C3-C2-N2	-2.41	105.99	110.61
4	E	1322	GAL	O3-C3-C2	-2.29	105.86	110.02
4	E	1321	SIA	C6-C5-N5	-2.04	107.42	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	1323	NAG	C3-C2-N2	-2.01	106.76	110.61
4	E	1322	GAL	C1-O5-C5	2.06	115.00	112.17
4	C	1323	NAG	C1-C2-C3	2.08	113.38	110.54
4	E	1322	GAL	C1-C2-C3	2.46	112.77	109.65
4	A	1322	GAL	C1-C2-C3	2.80	113.19	109.65
4	C	1323	NAG	O5-C1-C2	2.97	112.50	109.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.6 Ligand geometry [i](#)

10 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	1011	1	14,14,15	0.44	0	15,19,21	1.34	1 (6%)
3	NAG	A	1023	1	14,14,15	0.72	1 (7%)	15,19,21	2.62	6 (40%)
3	NAG	A	1164	1	14,14,15	0.45	0	15,19,21	1.03	1 (6%)
5	PO4	A	1325	-	4,4,4	0.76	0	6,6,6	0.48	0
3	NAG	C	1011	1	14,14,15	0.47	0	15,19,21	1.00	1 (6%)
3	NAG	C	1023	1	14,14,15	0.66	0	15,19,21	2.30	6 (40%)
3	NAG	C	1164	1	14,14,15	0.48	0	15,19,21	1.19	2 (13%)
3	NAG	E	1011	1	14,14,15	0.49	0	15,19,21	1.13	1 (6%)
3	NAG	E	1023	1	14,14,15	0.66	0	15,19,21	1.98	4 (26%)
3	NAG	E	1164	1	14,14,15	0.41	0	15,19,21	1.00	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	1011	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1023	1	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	A	1164	1	-	0/6/23/26	0/1/1/1
5	PO4	A	1325	-	-	0/0/0/0	0/0/0/0
3	NAG	C	1011	1	-	0/6/23/26	0/1/1/1
3	NAG	C	1023	1	-	0/6/23/26	0/1/1/1
3	NAG	C	1164	1	-	0/6/23/26	0/1/1/1
3	NAG	E	1011	1	-	0/6/23/26	0/1/1/1
3	NAG	E	1023	1	-	0/6/23/26	0/1/1/1
3	NAG	E	1164	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	1023	NAG	C1-C2	2.13	1.55	1.52

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1023	NAG	O5-C1-C2	-4.42	105.33	111.47
3	E	1023	NAG	O7-C7-C8	-2.45	117.59	122.06
3	C	1023	NAG	O7-C7-C8	-2.34	117.80	122.06
3	A	1023	NAG	O7-C7-C8	-2.32	117.84	122.06
3	C	1023	NAG	O5-C1-C2	-2.14	108.49	111.47
3	C	1164	NAG	C3-C4-C5	-2.10	106.51	110.22
3	C	1023	NAG	C1-C2-N2	2.01	113.93	110.49
3	E	1023	NAG	C4-C3-C2	2.63	114.87	111.02
3	C	1011	NAG	C1-O5-C5	2.72	115.92	112.17
3	E	1164	NAG	C1-O5-C5	2.79	116.00	112.17
3	A	1023	NAG	C1-C2-N2	3.04	115.69	110.49
3	A	1164	NAG	C1-O5-C5	3.15	116.51	112.17
3	A	1023	NAG	C8-C7-N2	3.18	121.85	116.11
3	E	1023	NAG	C8-C7-N2	3.23	121.94	116.11
3	C	1023	NAG	C8-C7-N2	3.31	122.08	116.11
3	C	1023	NAG	C1-O5-C5	3.33	116.76	112.17
3	C	1164	NAG	C1-O5-C5	3.39	116.84	112.17
3	E	1011	NAG	C1-O5-C5	3.68	117.24	112.17
3	A	1011	NAG	C1-O5-C5	3.78	117.37	112.17
3	A	1023	NAG	C1-O5-C5	3.99	117.67	112.17
3	E	1023	NAG	C2-N2-C7	5.41	130.83	122.94

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1023	NAG	C2-N2-C7	6.32	132.17	122.94
3	C	1023	NAG	C2-N2-C7	6.43	132.33	122.94

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	1023	NAG	C1

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1023	NAG	1	0
3	A	1164	NAG	1	0
3	C	1023	NAG	1	0
3	E	1023	NAG	2	0
3	E	1164	NAG	2	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, 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	324/327 (99%)	-0.10	2 (0%)	89 92	27, 42, 57, 84	0
1	C	322/327 (98%)	0.02	3 (0%)	84 87	24, 35, 60, 114	0
1	E	322/327 (98%)	0.06	7 (2%)	62 69	25, 37, 71, 122	0
2	B	163/166 (98%)	0.27	7 (4%)	36 43	23, 51, 80, 99	0
2	D	163/166 (98%)	1.31	41 (25%)	1 1	25, 74, 147, 156	0
2	F	163/166 (98%)	1.59	50 (30%)	0 0	23, 80, 145, 171	0
All	All	1457/1479 (98%)	0.35	110 (7%)	15 20	23, 41, 126, 171	0

All (110) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	33	GLY	9.3
2	F	160	PRO	9.0
2	F	29	GLU	7.0
2	F	5	ALA	6.9
2	D	159	TYR	6.8
2	D	31	GLY	6.5
2	F	141	TYR	6.4
2	F	156	THR	6.2
2	F	163	SER	6.1
2	D	5	ALA	5.7
2	F	26	HIS	5.5
2	F	140	PHE	5.4
2	F	157	TYR	5.3
2	D	29	GLU	5.3
2	D	7	ALA	5.0
2	F	159	TYR	4.9
2	F	32	SER	4.8
2	D	160	PRO	4.7
2	D	1	GLY	4.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	140	PHE	4.5
2	F	143	ARG	4.5
2	D	153	ARG	4.4
2	D	9	PHE	4.4
1	E	320	PRO	4.3
2	D	143	ARG	4.3
1	A	322	ARG	4.3
2	F	162	TYR	4.3
2	B	7	ALA	4.1
2	D	8	GLY	4.1
2	F	152	VAL	4.0
2	F	3	PHE	4.0
2	F	138	PHE	4.0
2	D	157	TYR	3.9
2	D	133	LEU	3.9
2	F	9	PHE	3.9
2	F	144	CYS	3.9
2	F	148	CYS	3.9
2	F	7	ALA	3.9
2	D	156	THR	3.7
2	D	124	LEU	3.7
2	D	161	GLN	3.7
2	F	124	LEU	3.7
2	F	161	GLN	3.6
1	C	1	ASP	3.6
2	F	127	ARG	3.5
2	D	3	PHE	3.5
2	D	132	GLU	3.5
2	D	11	GLU	3.5
2	F	158	ASP	3.4
2	B	1	GLY	3.4
2	F	39	GLU	3.4
2	B	9	PHE	3.3
1	C	320	PRO	3.3
2	D	150	GLU	3.3
2	D	134	GLY	3.3
2	F	139	GLU	3.2
2	F	38	LYS	3.2
2	D	130	ALA	3.2
1	E	3	ILE	3.2
2	D	144	CYS	3.2
2	B	3	PHE	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	38	LYS	3.2
2	D	162	TYR	3.2
2	D	6	ILE	3.1
2	D	138	PHE	3.1
2	D	149	MET	3.1
1	E	2	GLN	3.1
1	E	5	ILE	3.0
2	F	31	GLY	2.9
2	D	151	SER	2.8
2	F	147	GLU	2.8
2	D	127	ARG	2.8
2	D	2	LEU	2.7
2	D	27	SER	2.7
2	F	4	GLY	2.7
2	D	33	GLY	2.7
2	F	125	GLN	2.7
2	F	27	SER	2.6
2	F	1	GLY	2.6
2	F	19	ASP	2.6
2	F	126	LEU	2.6
2	F	116	LYS	2.6
2	F	11	GLU	2.5
2	D	141	TYR	2.5
2	D	26	HIS	2.5
1	E	4	CYS	2.4
2	F	8	GLY	2.4
2	F	34	TYR	2.4
1	A	139	ARG	2.4
2	F	6	ILE	2.4
2	D	126	LEU	2.4
2	F	21	TRP	2.4
2	F	133	LEU	2.4
2	D	152	VAL	2.3
2	F	134	GLY	2.3
1	E	0	PRO	2.3
2	F	35	ALA	2.3
2	D	131	LYS	2.3
1	C	2	GLN	2.3
2	F	137	CYS	2.2
2	D	40	SER	2.2
2	F	2	LEU	2.2
2	F	50	ASN	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	6	ILE	2.1
2	B	159	TYR	2.1
2	D	120	ASP	2.1
2	B	55	ILE	2.1
1	E	273	CYS	2.1
2	F	131	LYS	2.0
2	F	42	GLN	2.0

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

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

6.3 Carbohydrates [i](#)

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	E	1321	20/21	0.98	0.14	-0.52	29,31,36,36	0
4	SIA	A	1321	20/21	0.95	0.12	-0.62	34,39,43,44	0
4	SIA	C	1321	20/21	0.96	0.12	-0.86	27,31,33,35	0
4	GAL	E	1322	11/12	0.93	0.10	-2.99	38,43,45,47	0
4	GAL	C	1322	11/12	0.96	0.10	-4.16	35,37,41,42	0
4	NAG	C	1323	15/15	0.92	0.14	-	48,58,67,67	0
4	NAG	E	1323	15/15	0.87	0.22	-	52,60,66,67	0
4	NAG	A	1323	15/15	0.91	0.22	-	59,68,76,77	0
4	GAL	A	1322	11/12	0.94	0.11	-	44,47,51,53	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, 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(\AA^2)	Q<0.9
5	PO4	A	1325	5/5	0.85	0.30	8.10	79,82,86,87	0
3	NAG	E	1011	14/15	0.90	0.27	4.05	55,64,68,70	0
3	NAG	C	1164	14/15	0.95	0.13	-0.93	33,39,45,51	0
3	NAG	A	1164	14/15	0.94	0.12	-0.96	35,40,46,47	0
3	NAG	A	1011	14/15	0.97	0.12	-2.75	33,40,45,46	0
3	NAG	E	1164	14/15	0.92	0.10	-2.79	33,38,41,44	0
3	NAG	A	1023	14/15	0.86	0.19	-	50,61,65,68	0
3	NAG	C	1023	14/15	0.87	0.27	-	65,78,87,90	0
3	NAG	E	1023	14/15	0.88	0.24	-	60,70,76,76	0
3	NAG	C	1011	14/15	0.88	0.25	-	47,54,58,59	0

6.5 Other polymers [i](#)

There are no such residues in this entry.