



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:38 AM GMT

PDB ID : 2XQY  
Title : CRYSTAL STRUCTURE OF PSEUDORABIES CORE FRAGMENT OF GLYCOPROTEIN H IN COMPLEX WITH FAB D6.3  
Authors : Backovic, M.; Dubois, R.; Cockburn, J.; Sharff, A.; Vaney, M.; Granzow, H.; Klupp, B.; Bricogne, G.; Mettenleiter, T.; Rey, F.  
Deposited on : 2010-09-08  
Resolution : 2.05 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<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

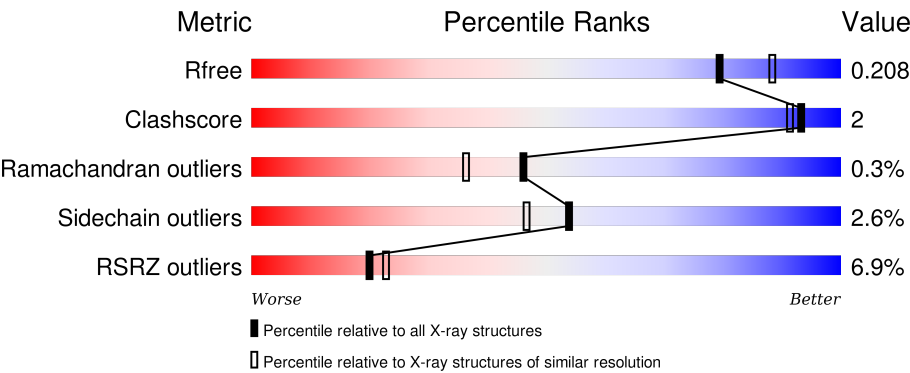
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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 <sub>free</sub>	91344	1192 (2.04-2.04)
Clashscore	102246	1269 (2.04-2.04)
Ramachandran outliers	100387	1258 (2.04-2.04)
Sidechain outliers	100360	1258 (2.04-2.04)
RSRZ outliers	91569	1194 (2.04-2.04)

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  $\geq 3$ , 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	572	<div><div>10%</div><div><div></div><div>74%</div><div>7%</div><div>•</div><div>19%</div></div></div>
1	E	572	<div><div>8%</div><div><div></div><div>74%</div><div>6%</div><div>•</div><div>19%</div></div></div>
2	G	261	<div><div>3%</div><div><div></div><div>80%</div><div>•</div><div>17%</div></div></div>
2	J	261	<div><div>2%</div><div><div></div><div>80%</div><div>•</div><div>17%</div></div></div>
3	K	220	<div><div>2%</div><div><div></div><div>95%</div><div>•</div><div>•</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	L	220	<div><div></div><div>2%</div><div>95%</div><div>5%</div></div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 14739 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 ENVELOPE GLYCOPROTEIN H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	464	Total	C	N	O	S	0	0	0
			3475	2195	624	637	19			
1	E	464	Total	C	N	O	S	0	0	0
			3475	2195	624	637	19			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	105	ARG	-	EXPRESSION TAG	UNP P27416
A	106	SER	-	EXPRESSION TAG	UNP P27416
A	640	PHE	-	EXPRESSION TAG	UNP P27416
A	641	GLU	-	EXPRESSION TAG	UNP P27416
A	642	ASP	-	EXPRESSION TAG	UNP P27416
A	643	ASP	-	EXPRESSION TAG	UNP P27416
A	644	ASP	-	EXPRESSION TAG	UNP P27416
A	645	ASP	-	EXPRESSION TAG	UNP P27416
A	646	LYS	-	EXPRESSION TAG	UNP P27416
A	647	ALA	-	EXPRESSION TAG	UNP P27416
A	648	GLY	-	EXPRESSION TAG	UNP P27416
A	649	TRP	-	EXPRESSION TAG	UNP P27416
A	650	SER	-	EXPRESSION TAG	UNP P27416
A	651	HIS	-	EXPRESSION TAG	UNP P27416
A	652	PRO	-	EXPRESSION TAG	UNP P27416
A	653	GLN	-	EXPRESSION TAG	UNP P27416
A	654	PHE	-	EXPRESSION TAG	UNP P27416
A	655	GLU	-	EXPRESSION TAG	UNP P27416
A	656	LYS	-	EXPRESSION TAG	UNP P27416
A	657	GLY	-	EXPRESSION TAG	UNP P27416
A	658	GLY	-	EXPRESSION TAG	UNP P27416
A	659	GLY	-	EXPRESSION TAG	UNP P27416
A	660	SER	-	EXPRESSION TAG	UNP P27416
A	661	GLY	-	EXPRESSION TAG	UNP P27416
A	662	GLY	-	EXPRESSION TAG	UNP P27416

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	663	GLY	-	EXPRESSION TAG	UNP P27416
A	664	SER	-	EXPRESSION TAG	UNP P27416
A	665	GLY	-	EXPRESSION TAG	UNP P27416
A	666	GLY	-	EXPRESSION TAG	UNP P27416
A	667	GLY	-	EXPRESSION TAG	UNP P27416
A	668	SER	-	EXPRESSION TAG	UNP P27416
A	669	TRP	-	EXPRESSION TAG	UNP P27416
A	670	SER	-	EXPRESSION TAG	UNP P27416
A	671	HIS	-	EXPRESSION TAG	UNP P27416
A	672	PRO	-	EXPRESSION TAG	UNP P27416
A	673	GLN	-	EXPRESSION TAG	UNP P27416
A	674	PHE	-	EXPRESSION TAG	UNP P27416
A	675	GLU	-	EXPRESSION TAG	UNP P27416
A	676	LYS	-	EXPRESSION TAG	UNP P27416
E	105	ARG	-	EXPRESSION TAG	UNP P27416
E	106	SER	-	EXPRESSION TAG	UNP P27416
E	105	ARG	-	EXPRESSION TAG	UNP P27416
E	106	SER	-	EXPRESSION TAG	UNP P27416
E	640	PHE	-	EXPRESSION TAG	UNP P27416
E	641	GLU	-	EXPRESSION TAG	UNP P27416
E	642	ASP	-	EXPRESSION TAG	UNP P27416
E	643	ASP	-	EXPRESSION TAG	UNP P27416
E	644	ASP	-	EXPRESSION TAG	UNP P27416
E	645	ASP	-	EXPRESSION TAG	UNP P27416
E	646	LYS	-	EXPRESSION TAG	UNP P27416
E	647	ALA	-	EXPRESSION TAG	UNP P27416
E	648	GLY	-	EXPRESSION TAG	UNP P27416
E	649	TRP	-	EXPRESSION TAG	UNP P27416
E	650	SER	-	EXPRESSION TAG	UNP P27416
E	651	HIS	-	EXPRESSION TAG	UNP P27416
E	652	PRO	-	EXPRESSION TAG	UNP P27416
E	653	GLN	-	EXPRESSION TAG	UNP P27416
E	654	PHE	-	EXPRESSION TAG	UNP P27416
E	655	GLU	-	EXPRESSION TAG	UNP P27416
E	656	LYS	-	EXPRESSION TAG	UNP P27416
E	657	GLY	-	EXPRESSION TAG	UNP P27416
E	658	GLY	-	EXPRESSION TAG	UNP P27416
E	659	GLY	-	EXPRESSION TAG	UNP P27416
E	660	SER	-	EXPRESSION TAG	UNP P27416
E	661	GLY	-	EXPRESSION TAG	UNP P27416
E	662	GLY	-	EXPRESSION TAG	UNP P27416
E	663	GLY	-	EXPRESSION TAG	UNP P27416

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	664	SER	-	EXPRESSION TAG	UNP P27416
E	665	GLY	-	EXPRESSION TAG	UNP P27416
E	666	GLY	-	EXPRESSION TAG	UNP P27416
E	667	GLY	-	EXPRESSION TAG	UNP P27416
E	668	SER	-	EXPRESSION TAG	UNP P27416
E	669	TRP	-	EXPRESSION TAG	UNP P27416
E	670	SER	-	EXPRESSION TAG	UNP P27416
E	671	HIS	-	EXPRESSION TAG	UNP P27416
E	672	PRO	-	EXPRESSION TAG	UNP P27416
E	673	GLN	-	EXPRESSION TAG	UNP P27416
E	674	PHE	-	EXPRESSION TAG	UNP P27416
E	675	GLU	-	EXPRESSION TAG	UNP P27416
E	676	LYS	-	EXPRESSION TAG	UNP P27416

- Molecule 2 is a protein called A13-D6.3 MONOCLONAL ANTIBODY.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	216	Total	C	N	O	S	0	0	0
			1648	1044	270	327	7			
2	J	216	Total	C	N	O	S	0	0	0
			1648	1044	270	327	7			

- Molecule 3 is a protein called A13-D6.3 MONOCLONAL ANTIBODY.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	K	218	Total	C	N	O	S	0	1	0
			1698	1059	288	344	7			
3	L	218	Total	C	N	O	S	0	0	0
			1689	1054	285	343	7			

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

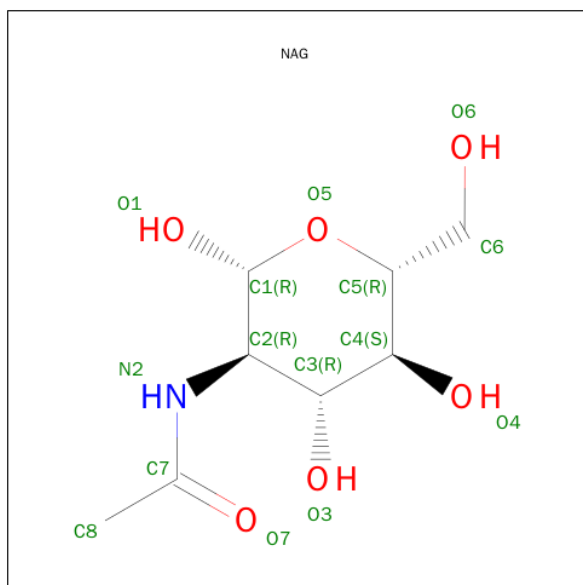
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	2	Total	C	N	O	0	0
			28	16	2	10		
4	A	2	Total	C	N	O	0	0
			28	16	2	10		
4	E	2	Total	C	N	O	0	0
			28	16	2	10		
4	G	2	Total	C	N	O	0	0
			28	16	2	10		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	J	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

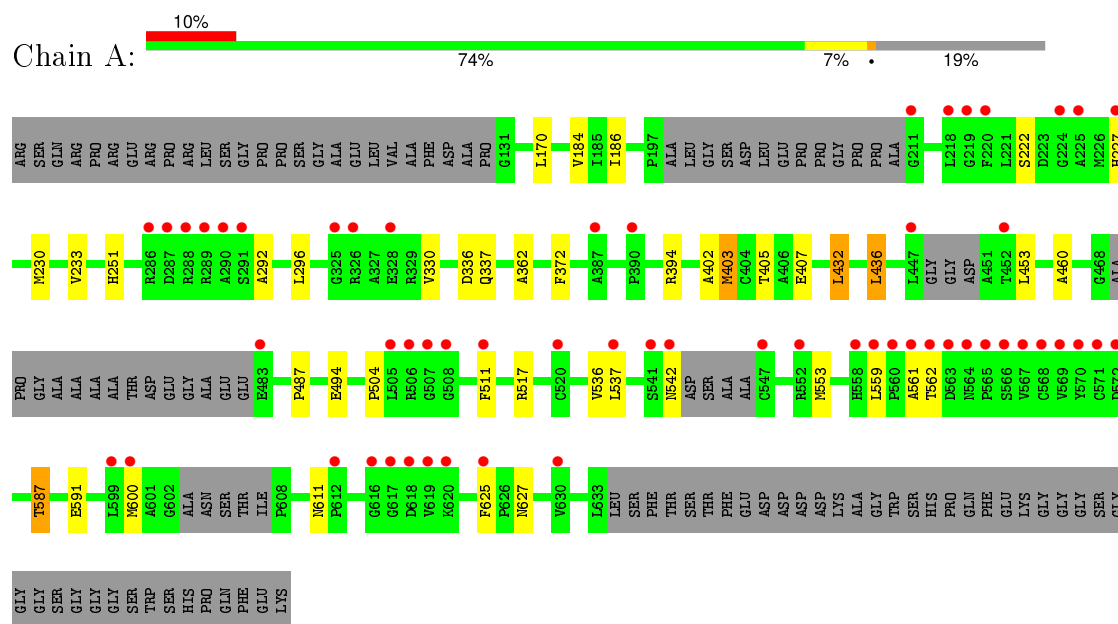
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	135	Total	O	0	0
			135	135		
6	E	159	Total	O	0	0
			159	159		
6	G	141	Total	O	0	0
			141	141		
6	J	184	Total	O	0	0
			184	184		
6	K	171	Total	O	0	0
			171	171		
6	L	162	Total	O	0	0
			162	162		

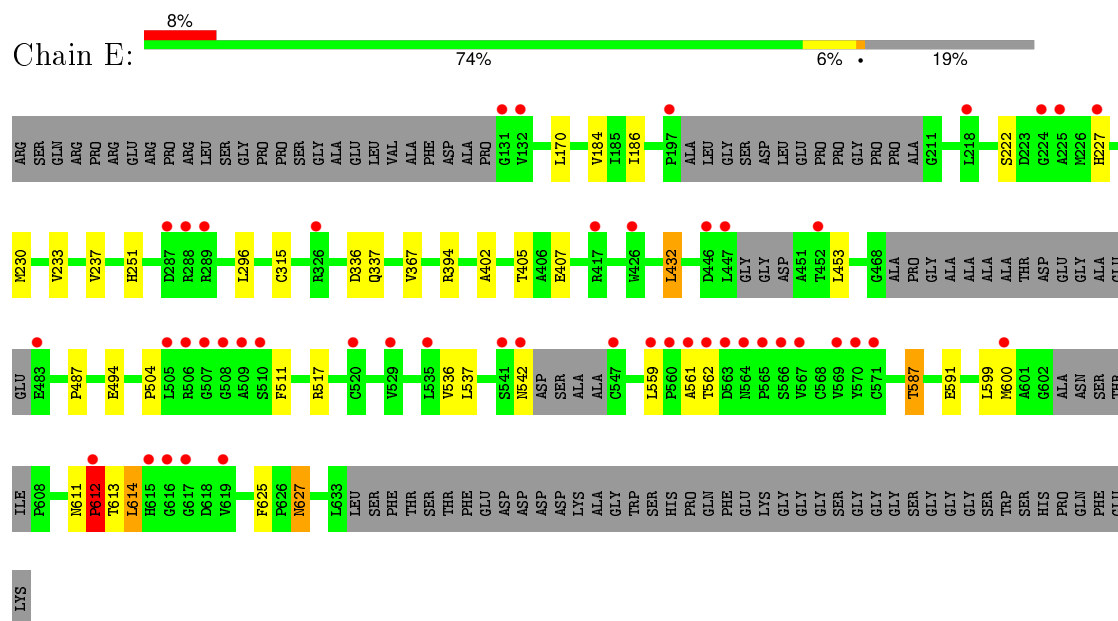
### 3 Residue-property plots

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: ENVELOPE GLYCOPROTEIN H



- Molecule 1: ENVELOPE GLYCOPROTEIN H







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.31Å 100.63Å 161.83Å 90.00° 94.02° 90.00°	Depositor
Resolution (Å)	58.12 – 2.05 55.64 – 2.05	Depositor EDS
% Data completeness (in resolution range)	91.8 (58.12-2.05) 92.1 (55.64-2.05)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.09 (at 2.05Å)	Xtriage
Refinement program	BUSTER 2.11.0	Depositor
R, $R_{free}$	0.188 , 0.205 0.190 , 0.208	Depositor DCC
$R_{free}$ test set	7603 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	49.7	Xtriage
Anisotropy	0.308	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 60.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 151525 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	14739	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.05 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.6582e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 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: 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	0.48	1/3544 (0.0%)	0.63	0/4826
1	E	0.48	0/3544	0.64	1/4826 (0.0%)
2	G	0.49	0/1691	0.59	0/2309
2	J	0.54	0/1691	0.60	0/2309
3	K	0.52	0/1742	0.62	0/2367
3	L	0.53	0/1730	0.62	0/2353
All	All	0.50	1/13942 (0.0%)	0.62	1/18990 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	403	MET	SD-CE	-5.96	1.44	1.77

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	612	PRO	C-N-CA	5.61	135.72	121.70

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	3475	0	3463	18	0
1	E	3475	0	3463	18	0
2	G	1648	0	1601	3	0
2	J	1648	0	1601	4	0
3	K	1698	0	1624	6	0
3	L	1689	0	1611	6	0
4	A	56	0	50	0	0
4	E	28	0	25	0	0
4	G	28	0	25	0	0
4	J	28	0	25	0	0
5	E	14	0	13	1	0
6	A	135	0	0	0	0
6	E	159	0	0	1	0
6	G	141	0	0	0	0
6	J	184	0	0	1	0
6	K	171	0	0	1	0
6	L	162	0	0	0	0
All	All	14739	0	13501	51	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 2.

All (51) 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:612:PRO:HB2	1:E:613:THR:HB	1.45	0.95
2:G:156:SER:H	2:G:196:ASN:HD21	1.30	0.77
1:A:337:GLN:HG2	3:L:56:SER:HB2	1.74	0.70
2:J:156:SER:H	2:J:196:ASN:HD21	1.40	0.69
3:L:150:ILE:HD11	3:L:179:LEU:HD21	1.81	0.61
1:A:251:HIS:CD2	1:A:432:LEU:H	2.22	0.57
1:E:296:LEU:HD23	1:E:336:ASP:HB3	1.86	0.56
1:E:612:PRO:HB2	1:E:613:THR:CB	2.29	0.56
1:A:296:LEU:HD23	1:A:336:ASP:HB3	1.86	0.56
1:E:394:ARG:HG3	1:E:504:PRO:HB2	1.88	0.56
1:E:337:GLN:HG2	3:K:56:SER:HB2	1.86	0.55
1:A:337:GLN:CG	3:L:56:SER:HB2	2.37	0.55
1:A:394:ARG:HG3	1:A:504:PRO:HB2	1.89	0.55
2:J:148:GLU:OE2	6:J:3133:HOH:O	2.18	0.55
3:K:150:ILE:HD11	3:K:179:LEU:HD21	1.90	0.54
1:A:184:VAL:HG11	1:A:233:VAL:HG11	1.89	0.54
1:A:436:LEU:HG	1:A:460:ALA:HB2	1.89	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:251:HIS:CD2	1:E:432:LEU:H	2.25	0.53
1:A:517:ARG:HA	1:A:536:VAL:HG11	1.92	0.52
1:E:517:ARG:HA	1:E:536:VAL:HG11	1.92	0.52
1:E:170:LEU:HD23	1:E:186:ILE:HD12	1.95	0.49
1:E:511:PHE:CD1	1:E:537:LEU:HD21	2.47	0.49
1:A:511:PHE:CD1	1:A:537:LEU:HD21	2.47	0.48
1:A:405:THR:HG22	1:A:407:GLU:H	1.77	0.48
1:E:402:ALA:HA	1:E:587:THR:HG22	1.96	0.48
1:E:405:THR:HG22	1:E:407:GLU:H	1.78	0.47
3:L:141:PRO:O	3:L:198:HIS:HE1	1.98	0.47
3:K:113:PRO:HG3	3:K:144:ILE:HD11	1.95	0.47
2:J:156:SER:H	2:J:196:ASN:ND2	2.09	0.47
1:A:170:LEU:HD23	1:A:186:ILE:HD12	1.97	0.46
1:A:402:ALA:HA	1:A:587:THR:HG22	1.98	0.46
1:A:625:PHE:C	1:A:627:ASN:H	2.18	0.46
2:J:38:LYS:HB2	2:J:48:ILE:HD11	1.98	0.46
1:E:184:VAL:HG11	1:E:233:VAL:HG11	1.98	0.45
1:E:315:CYS:HB3	6:E:3064:HOH:O	2.17	0.44
2:G:38:LYS:HB2	2:G:48:ILE:HD11	1.99	0.44
3:L:198:HIS:HD2	3:L:200:THR:OG1	2.00	0.44
3:K:198:HIS:HD2	3:K:200:THR:OG1	2.01	0.43
1:E:337:GLN:CG	3:K:56:SER:HB2	2.49	0.42
1:A:494:GLU:HB3	1:A:542:ASN:HD21	1.84	0.42
2:G:115:LYS:HE2	6:K:3081:HOH:O	2.18	0.42
1:A:292:ALA:O	1:A:330:VAL:HG11	2.19	0.42
1:E:494:GLU:HB3	1:E:542:ASN:HD21	1.85	0.42
3:K:36:TYR:HE2	3:K:89:GLN:HE21	1.68	0.42
3:L:40:PRO:HG2	3:L:165:ASP:CG	2.40	0.41
1:A:372:PHE:CZ	1:A:403:MET:HE2	2.56	0.41
1:E:625:PHE:C	1:E:627:ASN:H	2.24	0.41
1:E:599:LEU:HD12	1:E:611:ASN:HB2	2.02	0.41
1:E:625:PHE:CD1	5:E:2101:NAG:H61	2.56	0.40
1:A:511:PHE:HD1	1:A:537:LEU:HD21	1.87	0.40
1:A:362:ALA:HB1	1:A:403:MET:HE1	2.03	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	452/572 (79%)	434 (96%)	17 (4%)	1 (0%)	52	43
1	E	452/572 (79%)	433 (96%)	15 (3%)	4 (1%)	21	10
2	G	212/261 (81%)	206 (97%)	6 (3%)	0	100	100
2	J	212/261 (81%)	207 (98%)	5 (2%)	0	100	100
3	K	217/220 (99%)	213 (98%)	4 (2%)	0	100	100
3	L	216/220 (98%)	209 (97%)	7 (3%)	0	100	100
All	All	1761/2106 (84%)	1702 (97%)	54 (3%)	5 (0%)	46	36

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	627	ASN
1	A	561	ALA
1	E	561	ALA
1	E	614	LEU
1	E	612	PRO

### 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	347/423 (82%)	333 (96%)	14 (4%)	38	29
1	E	347/423 (82%)	333 (96%)	14 (4%)	38	29
2	G	189/222 (85%)	187 (99%)	2 (1%)	80	79

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	J	189/222 (85%)	185 (98%)	4 (2%)	61	56
3	K	192/193 (100%)	190 (99%)	2 (1%)	82	81
3	L	191/193 (99%)	189 (99%)	2 (1%)	82	81
All	All	1455/1676 (87%)	1417 (97%)	38 (3%)	54	47

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	222	SER
1	A	227	HIS
1	A	230	MET
1	A	432	LEU
1	A	436	LEU
1	A	453	LEU
1	A	487	PRO
1	A	553	MET
1	A	559	LEU
1	A	562	THR
1	A	587	THR
1	A	591	GLU
1	A	600	MET
1	A	611	ASN
1	E	222	SER
1	E	227	HIS
1	E	230	MET
1	E	237	VAL
1	E	367	VAL
1	E	432	LEU
1	E	453	LEU
1	E	487	PRO
1	E	559	LEU
1	E	562	THR
1	E	587	THR
1	E	591	GLU
1	E	600	MET
1	E	614	LEU
2	G	82(B)	SER
2	G	196	ASN
2	J	43	ARG
2	J	82(B)	SER
2	J	173	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	J	196	ASN
3	K	89	GLN
3	K	202	THR
3	L	89	GLN
3	L	214	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	251	HIS
1	A	542	ASN
1	A	557	GLN
1	A	615	HIS
1	E	251	HIS
1	E	542	ASN
1	E	557	GLN
1	E	611	ASN
1	E	615	HIS
2	G	196	ASN
2	J	196	ASN
3	K	89	GLN
3	K	198	HIS
3	L	89	GLN
3	L	198	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 ⓘ

10 carbohydrates are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link



column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	A	1101	1,4	14,14,15	1.35	2 (14%)	15,19,21	1.50	1 (6%)
4	NAG	A	1102	4	14,14,15	1.38	2 (14%)	15,19,21	1.09	2 (13%)
4	NAG	A	2101	1,4	14,14,15	1.52	1 (7%)	15,19,21	1.69	3 (20%)
4	NAG	A	2102	4	14,14,15	1.46	1 (7%)	15,19,21	1.05	1 (6%)
4	NAG	E	1101	1,4	14,14,15	1.28	3 (21%)	15,19,21	1.45	1 (6%)
4	NAG	E	1102	4	14,14,15	1.39	2 (14%)	15,19,21	0.96	1 (6%)
4	NAG	G	1101	2,4	14,14,15	1.93	2 (14%)	15,19,21	2.32	5 (33%)
4	NAG	G	1102	4	14,14,15	2.00	2 (14%)	15,19,21	1.82	3 (20%)
4	NAG	J	1101	2,4	14,14,15	1.90	3 (21%)	15,19,21	2.42	5 (33%)
4	NAG	J	1102	4	14,14,15	2.05	3 (21%)	15,19,21	1.80	3 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	1101	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1102	4	-	0/6/23/26	0/1/1/1
4	NAG	A	2101	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	2102	4	-	0/6/23/26	0/1/1/1
4	NAG	E	1101	1,4	-	0/6/23/26	0/1/1/1
4	NAG	E	1102	4	-	0/6/23/26	0/1/1/1
4	NAG	G	1101	2,4	-	0/6/23/26	0/1/1/1
4	NAG	G	1102	4	-	0/6/23/26	0/1/1/1
4	NAG	J	1101	2,4	-	0/6/23/26	0/1/1/1
4	NAG	J	1102	4	-	0/6/23/26	0/1/1/1

All (21) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1102	NAG	C3-C2	2.05	1.57	1.52
4	E	1101	NAG	C3-C2	2.07	1.57	1.52
4	J	1101	NAG	O4-C4	2.13	1.48	1.43
4	E	1101	NAG	C1-C2	2.17	1.55	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	1101	NAG	O5-C1	2.21	1.47	1.43
4	E	1102	NAG	C3-C2	2.21	1.57	1.52
4	A	1101	NAG	C3-C2	2.26	1.57	1.52
4	G	1102	NAG	C3-C2	2.33	1.57	1.52
4	J	1102	NAG	C3-C2	2.35	1.57	1.52
4	J	1102	NAG	C2-N2	2.59	1.50	1.46
4	A	1101	NAG	C1-C2	2.59	1.56	1.52
4	J	1101	NAG	C4-C5	2.76	1.58	1.53
4	G	1101	NAG	C4-C5	2.94	1.59	1.53
4	A	1102	NAG	C1-C2	2.99	1.56	1.52
4	E	1102	NAG	C1-C2	3.26	1.57	1.52
4	A	2101	NAG	C1-C2	3.93	1.57	1.52
4	A	2102	NAG	C1-C2	3.99	1.58	1.52
4	J	1101	NAG	C1-C2	4.29	1.58	1.52
4	G	1101	NAG	C1-C2	4.65	1.58	1.52
4	J	1102	NAG	C1-C2	5.64	1.60	1.52
4	G	1102	NAG	C1-C2	5.68	1.60	1.52

All (25) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	J	1101	NAG	O4-C4-C3	-4.56	100.07	110.34
4	G	1101	NAG	O4-C4-C3	-4.53	100.13	110.34
4	J	1102	NAG	C3-C2-N2	-3.15	103.01	110.56
4	G	1102	NAG	C3-C2-N2	-2.78	103.91	110.56
4	A	1102	NAG	C6-C5-C4	2.03	118.03	113.02
4	A	2101	NAG	C2-N2-C7	2.09	125.72	123.04
4	J	1101	NAG	O4-C4-C5	2.10	114.82	109.24
4	E	1102	NAG	C6-C5-C4	2.11	118.21	113.02
4	A	1102	NAG	O4-C4-C5	2.24	115.17	109.24
4	G	1102	NAG	C2-N2-C7	2.47	126.21	123.04
4	J	1102	NAG	C2-N2-C7	2.50	126.25	123.04
4	A	2101	NAG	O4-C4-C3	2.65	116.31	110.34
4	G	1101	NAG	O4-C4-C5	2.72	116.44	109.24
4	A	2102	NAG	C1-O5-C5	2.84	115.85	112.25
4	J	1101	NAG	C4-C3-C2	2.90	115.73	111.23
4	G	1101	NAG	C4-C3-C2	3.36	116.46	111.23
4	G	1101	NAG	C2-N2-C7	3.54	127.59	123.04
4	E	1101	NAG	C1-O5-C5	4.12	117.47	112.25
4	A	2101	NAG	C1-O5-C5	4.30	117.70	112.25
4	A	1101	NAG	C1-O5-C5	4.30	117.71	112.25
4	J	1101	NAG	C2-N2-C7	4.46	128.77	123.04

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	1101	NAG	C1-O5-C5	4.48	117.94	112.25
4	J	1102	NAG	C4-C3-C2	4.71	118.55	111.23
4	J	1101	NAG	C1-O5-C5	4.81	118.35	112.25
4	G	1102	NAG	C4-C3-C2	4.91	118.86	111.23

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

1 ligand is 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
5	NAG	E	2101	1	14,14,15	1.44	2 (14%)	15,19,21	0.91	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
5	NAG	E	2101	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)
5	E	2101	NAG	O5-C5	2.18	1.48	1.43
5	E	2101	NAG	C1-C2	2.41	1.55	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
5	E	2101	NAG	C3-C2-N2	2.04	115.44	110.56

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	2101	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<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	464/572 (81%)	0.70	57 (12%) 5 5	47, 73, 135, 168	0
1	E	464/572 (81%)	0.58	47 (10%) 9 10	45, 70, 124, 160	0
2	G	216/261 (82%)	0.49	7 (3%) 51 58	43, 61, 93, 130	1 (0%)
2	J	216/261 (82%)	0.50	5 (2%) 64 70	38, 53, 89, 141	1 (0%)
3	K	218/220 (99%)	0.28	4 (1%) 71 76	41, 58, 84, 123	0
3	L	218/220 (99%)	0.42	4 (1%) 71 76	40, 54, 87, 120	0
All	All	1796/2106 (85%)	0.53	124 (6%) 20 23	38, 63, 118, 168	2 (0%)

All (124) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	561	ALA	10.4
1	A	561	ALA	9.5
1	E	612	PRO	8.9
1	A	562	THR	8.7
1	A	560	PRO	7.9
1	E	562	THR	7.9
1	E	507	GLY	7.8
1	E	541	SER	7.0
3	L	214	CYS	7.0
3	K	214	CYS	6.5
1	A	567	VAL	6.3
1	E	560	PRO	6.3
1	A	564	ASN	6.2
1	A	559	LEU	5.6
1	E	509	ALA	5.5
1	E	563	ASP	5.3
1	A	568	CYS	5.2
1	A	565	PRO	5.1
1	A	220	PHE	5.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	452	THR	5.1
1	A	570	TYR	5.0
1	E	619	VAL	4.6
1	E	564	ASN	4.6
1	E	567	VAL	4.6
1	E	566	SER	4.4
1	E	542	ASN	4.4
2	J	193	ILE	4.4
1	A	569	VAL	4.4
3	L	213	GLU	4.3
1	E	569	VAL	4.2
1	A	452	THR	4.1
1	A	506	ARG	3.9
1	A	541	SER	3.9
1	E	506	ARG	3.9
1	A	563	ASP	3.6
1	A	571	CYS	3.6
1	A	290	ALA	3.6
2	J	210	ILE	3.6
1	A	227	HIS	3.5
1	A	600	MET	3.5
1	A	618	ASP	3.4
1	A	542	ASN	3.4
1	A	218	LEU	3.3
1	A	287	ASP	3.3
1	E	570	TYR	3.3
1	A	599	LEU	3.3
1	A	616	GLY	3.3
1	E	508	GLY	3.3
1	E	289	ARG	3.3
1	E	616	GLY	3.2
1	E	218	LEU	3.2
1	E	559	LEU	3.1
1	A	390	PRO	3.1
1	E	131	GLY	3.1
1	A	566	SER	3.1
1	A	289	ARG	3.1
1	E	565	PRO	3.0
1	E	505	LEU	3.0
2	G	193	ILE	3.0
2	G	132	THR	2.9
1	A	326	ARG	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	552	ARG	2.9
1	A	507	GLY	2.9
1	A	225	ALA	2.9
1	A	630	VAL	2.9
1	E	520	CYS	2.8
1	E	483	GLU	2.7
1	E	326	ARG	2.7
1	A	612	PRO	2.7
1	A	286	ARG	2.7
1	A	288	ARG	2.7
1	E	617	GLY	2.7
2	G	173	ASP	2.7
1	E	535	LEU	2.6
1	E	510	SER	2.6
1	E	447	LEU	2.6
1	A	520	CYS	2.6
1	A	508	GLY	2.6
1	E	227	HIS	2.5
1	E	615	HIS	2.5
1	A	619	VAL	2.5
1	A	387	ALA	2.4
3	K	212	ASN	2.4
1	A	625	PHE	2.4
3	K	125	LEU	2.4
1	A	325	GLY	2.4
1	A	537	LEU	2.3
2	G	9	ALA	2.3
1	A	224	GLY	2.3
1	A	291	SER	2.3
1	A	572	ASP	2.3
1	A	547	CYS	2.3
1	E	132	VAL	2.3
2	G	89	VAL	2.3
1	A	505	LEU	2.3
3	L	212	ASN	2.2
2	J	186	SER	2.2
2	G	210	ILE	2.2
1	A	617	GLY	2.2
2	J	82(B)	SER	2.2
1	A	328	GLU	2.2
1	E	547	CYS	2.1
1	E	426	TRP	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	197	PRO	2.1
3	K	213	GLU	2.1
3	L	30	TYR	2.1
1	E	600	MET	2.1
2	J	18	VAL	2.1
1	A	558	HIS	2.1
1	E	287	ASP	2.1
1	E	446	ASP	2.1
1	E	529	VAL	2.1
1	A	483	GLU	2.0
1	A	211	GLY	2.0
1	A	511	PHE	2.0
2	G	194	THR	2.0
1	A	620	LYS	2.0
1	E	288	ARG	2.0
1	E	571	CYS	2.0
1	A	219	GLY	2.0
1	E	224	GLY	2.0
1	A	447	LEU	2.0
1	E	225	ALA	2.0
1	E	417	ARG	2.0

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

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

## 6.3 Carbohydrates [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	E	1102	14/15	0.91	0.15	1.03	81,86,90,92	0
4	NAG	A	1102	14/15	0.94	0.12	0.14	77,82,85,88	0
4	NAG	E	1101	14/15	0.97	0.10	-0.40	72,77,81,81	0
4	NAG	A	1101	14/15	0.95	0.09	-0.96	66,71,75,75	0
4	NAG	J	1102	14/15	0.69	0.26	-	102,108,112,112	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	A	2102	14/15	0.78	0.45	-	179,183,189,192	0
4	NAG	G	1102	14/15	0.77	0.20	-	112,117,123,123	0
4	NAG	A	2101	14/15	0.76	0.40	-	169,172,179,183	0
4	NAG	J	1101	14/15	0.78	0.20	-	86,91,95,97	0
4	NAG	G	1101	14/15	0.86	0.18	-	98,104,108,108	0

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NAG	E	2101	14/15	0.65	0.26	-	128,133,140,141	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.