



Full wwPDB X-ray Structure Validation Report ⓘ

Apr 5, 2016 – 08:16 AM EDT

PDB ID : 4ZH1
Title : Complement factor H in complex with the GM1 glycan
Authors : Blaum, B.S.; Stehle, T.
Deposited on : 2015-04-24
Resolution : 2.24 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 : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027107
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027107

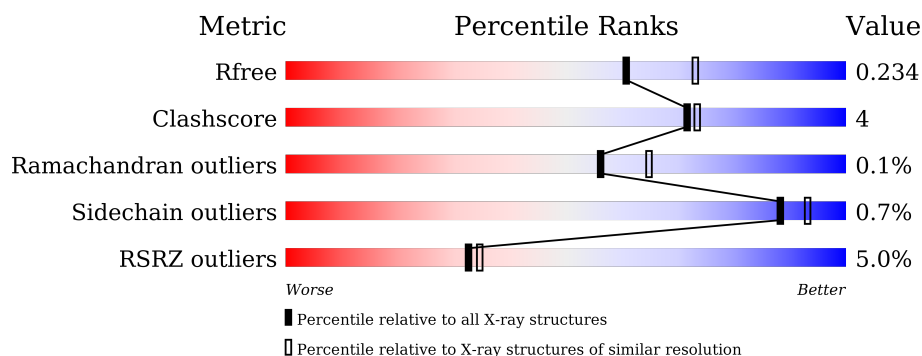
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.24 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	1611 (2.26-2.22)
Clashscore	102246	1764 (2.26-2.22)
Ramachandran outliers	100387	1724 (2.26-2.22)
Sidechain outliers	100360	1724 (2.26-2.22)
RSRZ outliers	91569	1616 (2.26-2.22)

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	317	<div> <div>85%</div> <div>8%</div> <div>8%</div> </div>
1	B	317	<div> <div>80%</div> <div>11%</div> <div>8%</div> </div>
1	C	317	<div> <div>4%</div> <div>85%</div> <div>6%</div> <div>8%</div> </div>
2	D	129	<div> <div>2%</div> <div>85%</div> <div>12%</div> <div>•</div> </div>
2	E	129	<div> <div>3%</div> <div>88%</div> <div>6%</div> <div>5%</div> </div>
2	F	129	<div> <div>28%</div> <div>70%</div> <div>10%</div> <div>•</div> <div>19%</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	GOL	D	1303	-	-	-	X

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 10322 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 Complement C3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	293	Total	C	N	O	S	0	1	0
			2311	1484	391	427	9			
1	B	292	Total	C	N	O	S	0	1	0
			2295	1473	381	432	9			
1	C	291	Total	C	N	O	S	0	3	0
			2280	1471	378	422	9			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	GLY	-	expression tag	UNP P01024
A	-5	PRO	-	expression tag	UNP P01024
A	-4	LEU	-	expression tag	UNP P01024
A	-3	GLY	-	expression tag	UNP P01024
A	-2	SER	-	expression tag	UNP P01024
A	-1	PRO	-	expression tag	UNP P01024
A	0	GLU	-	expression tag	UNP P01024
A	1	PHE	-	expression tag	UNP P01024
A	2	ARG	-	expression tag	UNP P01024
A	17	ALA	CYS	conflict	UNP P01024
B	-6	GLY	-	expression tag	UNP P01024
B	-5	PRO	-	expression tag	UNP P01024
B	-4	LEU	-	expression tag	UNP P01024
B	-3	GLY	-	expression tag	UNP P01024
B	-2	SER	-	expression tag	UNP P01024
B	-1	PRO	-	expression tag	UNP P01024
B	0	GLU	-	expression tag	UNP P01024
B	1	PHE	-	expression tag	UNP P01024
B	2	ARG	-	expression tag	UNP P01024
B	17	ALA	CYS	conflict	UNP P01024
C	-6	GLY	-	expression tag	UNP P01024
C	-5	PRO	-	expression tag	UNP P01024
C	-4	LEU	-	expression tag	UNP P01024

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP P01024
C	-2	SER	-	expression tag	UNP P01024
C	-1	PRO	-	expression tag	UNP P01024
C	0	GLU	-	expression tag	UNP P01024
C	1	PHE	-	expression tag	UNP P01024
C	2	ARG	-	expression tag	UNP P01024
C	17	ALA	CYS	conflict	UNP P01024

- Molecule 2 is a protein called Complement factor H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	126	Total	C	N	O	S	0	2	0
			1011	637	176	189	9			
2	E	122	Total	C	N	O	S	0	2	0
			989	620	176	184	9			
2	F	104	Total	C	N	O	S	0	1	0
			807	512	135	153	7			

There are 12 discrepancies between the modelled and reference sequences:

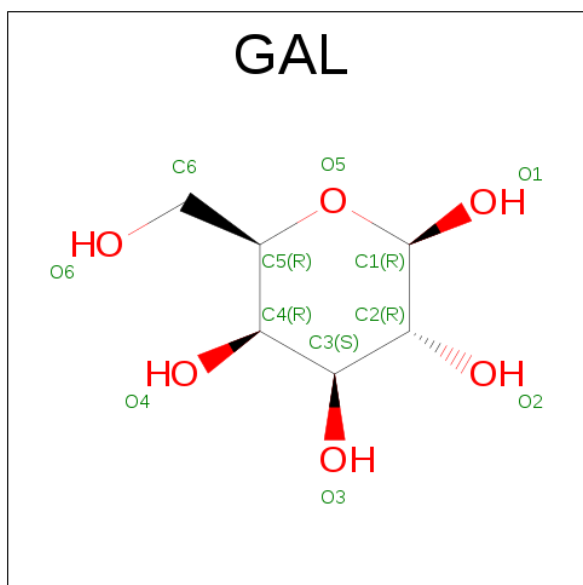
Chain	Residue	Modelled	Actual	Comment	Reference
D	1103	GLU	-	expression tag	UNP P08603
D	1104	ALA	-	expression tag	UNP P08603
D	1105	GLU	-	expression tag	UNP P08603
D	1106	PHE	-	expression tag	UNP P08603
E	1103	GLU	-	expression tag	UNP P08603
E	1104	ALA	-	expression tag	UNP P08603
E	1105	GLU	-	expression tag	UNP P08603
E	1106	PHE	-	expression tag	UNP P08603
F	1103	GLU	-	expression tag	UNP P08603
F	1104	ALA	-	expression tag	UNP P08603
F	1105	GLU	-	expression tag	UNP P08603
F	1106	PHE	-	expression tag	UNP P08603

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



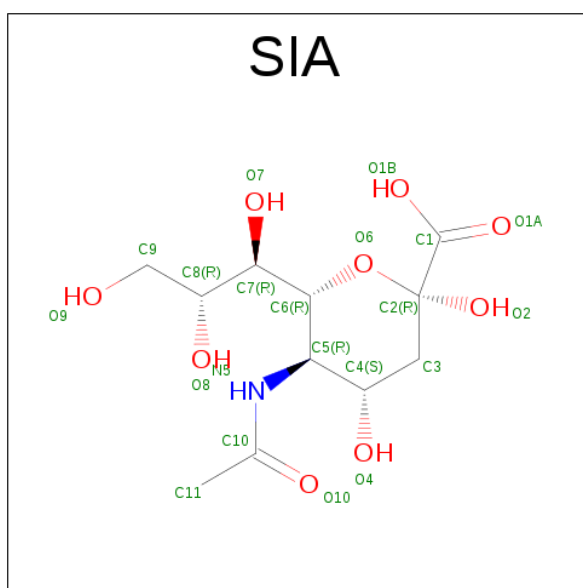
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is BETA-D-GALACTOSE (three-letter code: GAL) (formula: C₆H₁₂O₆).



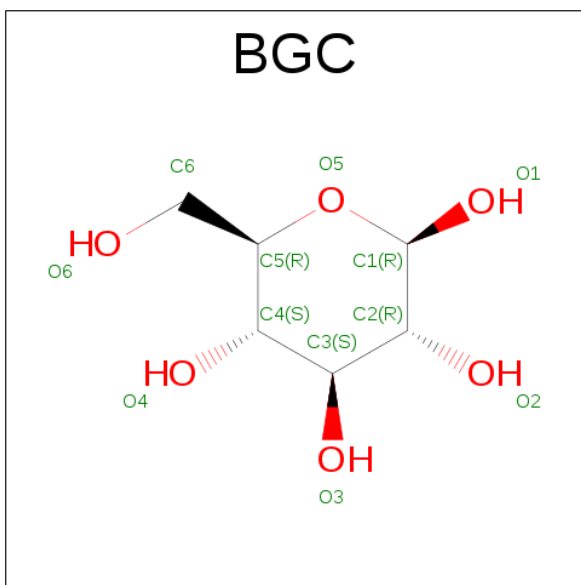
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			12	6	6		
4	E	1	Total	C	O	0	0
			11	6	5		
4	E	1	Total	C	O	0	0
			11	6	5		
4	F	1	Total	C	O	0	0
			11	6	5		

- Molecule 5 is O-SIALIC ACID (three-letter code: SIA) (formula: $C_{11}H_{19}NO_9$).



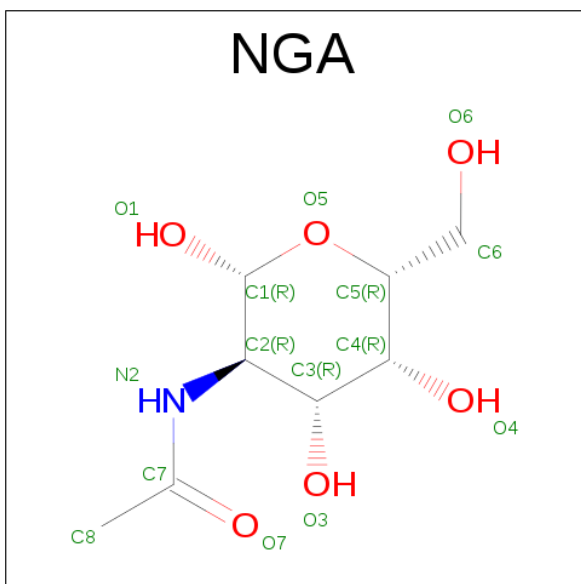
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	D	1	Total	C	N	O	0	0
			20	11	1	8		
5	E	1	Total	C	N	O	0	0
			20	11	1	8		
5	F	1	Total	C	N	O	0	0
			20	11	1	8		

- Molecule 6 is BETA-D-GLUCOSE (three-letter code: BGC) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	E	1	Total	C	O	0	0
			12	6	6		
6	F	1	Total	C	O	0	0
			12	6	6		

- Molecule 7 is N-ACETYL-D-GALACTOSAMINE (three-letter code: NGA) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	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
7	F	1	Total	C	N	O	0	0
			14	8	1	5		

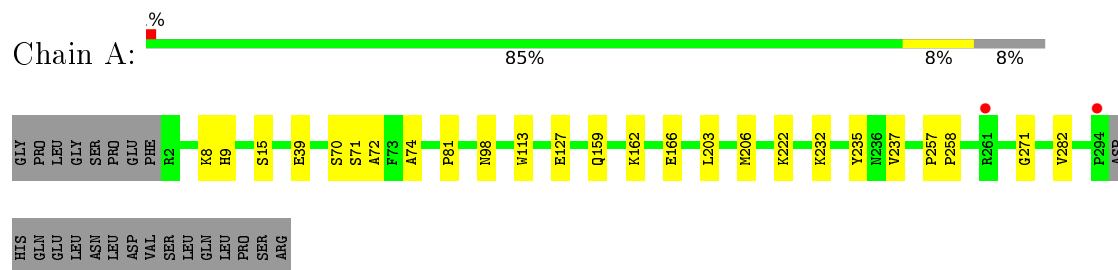
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	143	Total	O	0	0
			143	143		
8	B	101	Total	O	0	0
			101	101		
8	C	73	Total	O	0	0
			73	73		
8	D	75	Total	O	0	0
			75	75		
8	E	50	Total	O	0	0
			50	50		
8	F	6	Total	O	0	0
			6	6		

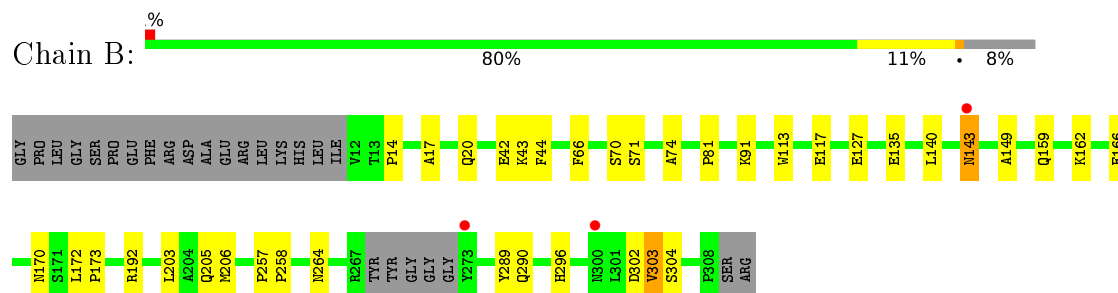
3 Residue-property plots

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

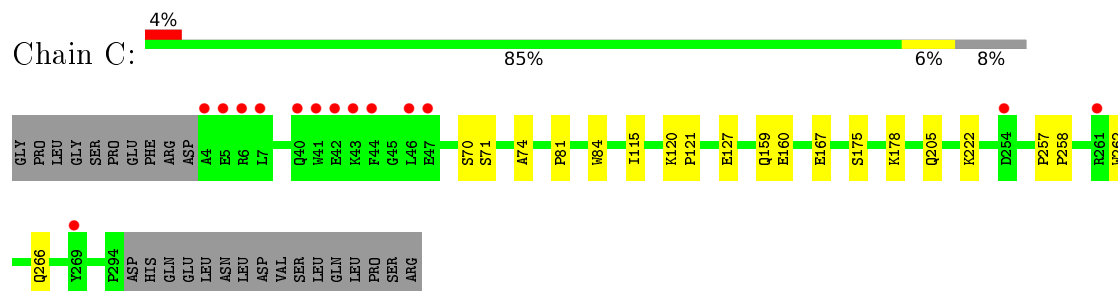
• Molecule 1: Complement C3



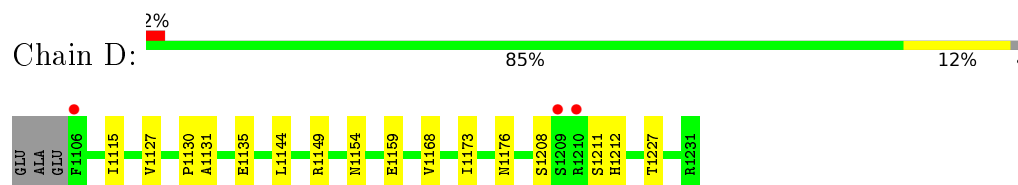
• Molecule 1: Complement C3



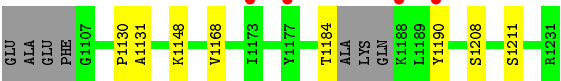
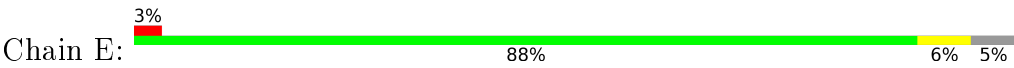
• Molecule 1: Complement C3



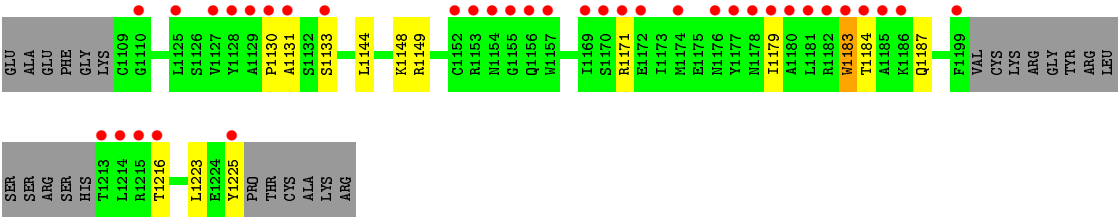
• Molecule 2: Complement factor H



● Molecule 2: Complement factor H



● Molecule 2: Complement factor H



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	75.77Å 82.86Å 86.57Å 112.97° 111.43° 99.63°	Depositor
Resolution (Å)	46.18 – 2.24 46.18 – 2.24	Depositor EDS
% Data completeness (in resolution range)	98.0 (46.18-2.24) 80.4 (46.18-2.24)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.11 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.181 , 0.228 0.188 , 0.234	Depositor DCC
R_{free} test set	3965 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	42.7	Xtriage
Anisotropy	0.486	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 50.8	EDS
Estimated twinning fraction	0.009 for -h,-k,h+k+l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 79286 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	10322	wwPDB-VP
Average B, all atoms (Å ²)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.19% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GAL, GOL, SIA, BGC, NGA

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	0.61	0/2360	0.67	0/3198
1	B	0.59	0/2342	0.68	0/3179
1	C	0.51	0/2329	0.63	0/3164
2	D	0.65	0/1038	0.68	0/1408
2	E	0.59	0/1015	0.71	2/1376 (0.1%)
2	F	0.50	0/829	0.67	0/1132
All	All	0.58	0/9913	0.67	2/13457 (0.0%)

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	B	0	1
2	F	0	1
All	All	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	1148	LYS	CB-CA-C	-6.51	97.37	110.40
2	E	1148	LYS	CA-CB-CG	5.08	124.57	113.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	143	ASN	Peptide
2	F	1183	TRP	Peptide

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	2311	0	2306	21	0
1	B	2295	0	2280	23	0
1	C	2280	0	2256	15	0
2	D	1011	0	967	15	0
2	E	989	0	941	6	0
2	F	807	0	744	7	0
3	A	6	0	8	1	0
3	B	6	0	8	1	0
3	C	6	0	8	3	0
3	D	6	0	8	0	0
4	D	12	0	11	0	0
4	E	22	0	18	0	0
4	F	11	0	8	0	0
5	D	20	0	17	0	0
5	E	20	0	17	0	0
5	F	20	0	17	1	0
6	E	12	0	11	0	0
6	F	12	0	11	0	0
7	E	14	0	12	0	0
7	F	14	0	13	0	0
8	A	143	0	0	5	0
8	B	101	0	0	3	0
8	C	73	0	0	3	0
8	D	75	0	0	5	0
8	E	50	0	0	0	0
8	F	6	0	0	0	0
All	All	10322	0	9661	84	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

All (84) 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)
2:E:1168:VAL:HG22	2:E:1190[A]:TYR:CE1	1.93	1.03
1:C:121:PRO:O	1:C:178:LYS:HE3	1.69	0.92
2:E:1168:VAL:HG22	2:E:1190[A]:TYR:CZ	2.06	0.90
1:B:264:ASN:OD1	8:B:501:HOH:O	1.95	0.84
1:A:166:GLU:HG2	8:A:590:HOH:O	1.78	0.82
1:B:159:GLN:NE2	1:B:205:GLN:OE1	2.13	0.82
1:A:98[A]:ASN:ND2	8:A:501:HOH:O	2.15	0.80
2:D:1154:ASN:N	8:D:1402:HOH:O	2.25	0.69
1:C:262:TRP:CZ2	1:C:266:GLN:OE1	2.46	0.69
1:C:121:PRO:O	1:C:178:LYS:CE	2.41	0.68
1:C:159:GLN:OE1	1:C:205[A]:GLN:OE1	2.11	0.66
2:D:1149:ARG:NH1	8:D:1403:HOH:O	2.29	0.66
1:A:9:HIS:O	2:D:1173:ILE:HD11	1.96	0.65
1:A:166:GLU:CG	8:A:590:HOH:O	2.39	0.63
1:C:115[B]:ILE:HD12	1:C:175:SER:OG	2.00	0.61
2:F:1133[B]:SER:OG	2:F:1149:ARG:CG	2.49	0.59
2:F:1216:THR:OG1	2:F:1223:LEU:HD11	2.03	0.58
1:B:70:SER:O	1:B:71:SER:HB2	2.03	0.58
1:C:70:SER:O	1:C:71:SER:HB2	2.02	0.58
2:F:1144:LEU:HD21	2:F:1148:LYS:HB2	1.86	0.58
1:A:70:SER:O	1:A:71:SER:HB2	2.05	0.56
1:C:167:GLU:OE1	8:C:501:HOH:O	2.18	0.55
2:D:1212:HIS:HB2	8:D:1401:HOH:O	2.06	0.55
2:F:1171:ARG:CG	2:F:1187:GLN:HB3	2.35	0.55
1:B:127:GLU:HG2	3:B:401:GOL:H2	1.89	0.54
1:C:120:LYS:NZ	8:C:507:HOH:O	2.41	0.54
1:A:39:GLU:C	8:A:532:HOH:O	2.45	0.53
1:A:237:VAL:CG1	1:A:282:VAL:HG11	2.39	0.53
1:C:160:GLU:HG3	8:C:537:HOH:O	2.08	0.52
2:D:1208:SER:O	2:D:1211[B]:SER:OG	2.07	0.52
1:B:135:GLU:HG3	8:B:581:HOH:O	2.11	0.51
1:A:237:VAL:CG1	1:A:282:VAL:CG1	2.90	0.50
2:D:1130:PRO:O	2:D:1131:ALA:HB3	2.12	0.50
2:E:1168:VAL:HG22	2:E:1190[A]:TYR:OH	2.10	0.50
1:A:271:GLY:O	2:D:1176:ASN:HB2	2.11	0.49
1:B:14:PRO:HG2	1:B:20:GLN:OE1	2.12	0.49
1:A:232:LYS:HD3	1:A:235:TYR:CZ	2.48	0.49
1:C:74:ALA:HB2	1:C:81:PRO:HA	1.93	0.49
2:E:1130:PRO:O	2:E:1131:ALA:HB3	2.12	0.49
1:B:257:PRO:N	1:B:258:PRO:HD2	2.28	0.49
1:A:237:VAL:HG12	1:A:282:VAL:CG1	2.43	0.49
2:D:1208:SER:OG	2:D:1211[A]:SER:HB2	2.13	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:15:SER:OG	2:D:1227:THR:HG23	2.13	0.48
1:B:74:ALA:HB2	1:B:81:PRO:HA	1.95	0.48
2:E:1208:SER:HB2	2:E:1211:SER:HB2	1.96	0.47
2:F:1130:PRO:O	2:F:1131:ALA:HB3	2.14	0.47
2:D:1135:GLU:OE2	2:D:1149:ARG:NH2	2.48	0.47
1:B:302:ASP:OD1	1:B:304:SER:OG	2.26	0.47
1:B:143:ASN:OD1	1:B:192:ARG:NH1	2.46	0.47
1:B:203:LEU:HD23	1:B:206:MET:CE	2.46	0.46
1:C:257:PRO:N	1:C:258:PRO:HD2	2.30	0.46
1:B:289:TYR:CE2	1:B:296:HIS:HB2	2.50	0.46
1:A:203:LEU:HD23	1:A:206:MET:CE	2.46	0.46
1:A:39:GLU:HA	8:A:532:HOH:O	2.15	0.45
2:D:1208:SER:HG	2:D:1211[A]:SER:HB2	1.81	0.45
1:C:262:TRP:CH2	1:C:266:GLN:OE1	2.69	0.45
1:C:127:GLU:HG2	3:C:401:GOL:C3	2.47	0.45
2:D:1127:VAL:HG12	8:D:1442:HOH:O	2.15	0.45
2:E:1184:THR:O	2:E:1184:THR:HG22	2.16	0.45
1:A:74:ALA:HB2	1:A:81:PRO:HA	1.99	0.44
1:B:91:LYS:CE	8:B:549:HOH:O	2.65	0.44
1:B:143:ASN:OD1	1:B:192:ARG:HD3	2.16	0.44
1:A:162:LYS:O	1:A:166:GLU:HB2	2.18	0.44
1:A:159:GLN:OE1	1:A:206:MET:HG2	2.18	0.44
2:F:1179:ILE:HD11	2:F:1225:TYR:HD2	1.83	0.43
1:B:140:LEU:CD2	1:B:149:ALA:HB1	2.49	0.43
1:A:257:PRO:N	1:A:258:PRO:HD2	2.33	0.43
1:B:14:PRO:HG2	1:B:20:GLN:CD	2.40	0.42
1:B:203:LEU:HD23	1:B:206:MET:HE1	2.00	0.42
1:B:44:PHE:CE2	1:B:303:VAL:HG12	2.54	0.42
1:B:17:ALA:HB3	1:B:66:PHE:CZ	2.54	0.42
1:A:127:GLU:HG2	3:A:401:GOL:H2	2.01	0.42
1:A:8:LYS:HE3	2:D:1168:VAL:HG21	2.02	0.41
2:D:1135:GLU:OE2	2:D:1149:ARG:NH1	2.53	0.41
2:D:1115:ILE:HG22	8:D:1425:HOH:O	2.20	0.41
2:F:1183:TRP:O	2:F:1184:THR:C	2.59	0.41
1:B:162:LYS:O	1:B:166:GLU:HB2	2.21	0.41
1:B:290:GLN:HG2	1:B:296:HIS:CE1	2.56	0.41
1:B:113:TRP:CD1	1:B:117:GLU:HG3	2.56	0.41
1:B:172:LEU:HB3	1:B:173:PRO:HD3	2.03	0.40
5:F:1303:SIA:O8	5:F:1303:SIA:O1A	2.37	0.40
1:C:127:GLU:HG2	3:C:401:GOL:H31	2.04	0.40
1:A:72:ALA:HB2	1:A:113:TRP:CD2	2.56	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:84:TRP:HB2	3:C:401:GOL:H12	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	292/317 (92%)	285 (98%)	7 (2%)	0	100	100
1	B	289/317 (91%)	279 (96%)	9 (3%)	1 (0%)	46	51
1	C	292/317 (92%)	285 (98%)	7 (2%)	0	100	100
2	D	126/129 (98%)	124 (98%)	2 (2%)	0	100	100
2	E	120/129 (93%)	119 (99%)	1 (1%)	0	100	100
2	F	101/129 (78%)	97 (96%)	4 (4%)	0	100	100
All	All	1220/1338 (91%)	1189 (98%)	30 (2%)	1 (0%)	56	65

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	303	VAL

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	240/264 (91%)	239 (100%)	1 (0%)	93	96
1	B	243/264 (92%)	240 (99%)	3 (1%)	78	87
1	C	234/264 (89%)	233 (100%)	1 (0%)	93	96
2	D	111/115 (96%)	109 (98%)	2 (2%)	66	77
2	E	109/115 (95%)	109 (100%)	0	100	100
2	F	86/115 (75%)	86 (100%)	0	100	100
All	All	1023/1137 (90%)	1016 (99%)	7 (1%)	88	93

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	222	LYS
1	B	42	GLU
1	B	43	LYS
1	B	170	ASN
1	C	222	LYS
2	D	1144	LEU
2	D	1159	GLU

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

Mol	Chain	Res	Type
1	B	142	ASN
1	B	159	GLN
1	B	266	GLN
1	C	40	GLN
1	C	142	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue 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	$\# Z > 2$
4	GAL	E	1305	7	11,11,12	0.55	0	15,15,17	1.16	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
4	GAL	E	1305	7	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
4	E	1305	GAL	C1-O5-C5	2.63	116.00	112.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

15 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	GOL	A	401	-	5,5,5	0.36	0	5,5,5	0.44	0
3	GOL	B	401	-	5,5,5	0.43	0	5,5,5	0.63	0
3	GOL	C	401	-	5,5,5	0.38	0	5,5,5	1.30	1 (20%)
4	GAL	D	1301	5	12,12,12	0.88	0	17,17,17	1.35	2 (11%)
5	SIA	D	1302	4	17,20,21	0.51	0	18,28,31	1.27	2 (11%)
3	GOL	D	1303	-	5,5,5	0.53	0	5,5,5	0.87	0
6	BGC	E	1301	4	12,12,12	0.46	0	17,17,17	0.71	0
4	GAL	E	1302	5,7,6	11,11,12	0.56	0	15,15,17	1.71	3 (20%)
5	SIA	E	1303	4	17,20,21	0.57	0	18,28,31	1.20	2 (11%)
7	NGA	E	1304	4	14,14,15	0.55	0	15,19,21	1.48	1 (6%)
4	GAL	E	1305	7	11,11,12	0.55	0	15,15,17	1.16	1 (6%)
6	BGC	F	1301	4	12,12,12	0.82	0	17,17,17	0.94	0
4	GAL	F	1302	5,7,6	11,11,12	0.41	0	15,15,17	1.55	3 (20%)
5	SIA	F	1303	4	17,20,21	0.53	0	18,28,31	1.25	1 (5%)
7	NGA	F	1304	4	14,14,15	0.46	0	15,19,21	0.93	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	A	401	-	-	0/4/4/4	0/0/0/0
3	GOL	B	401	-	-	0/4/4/4	0/0/0/0
3	GOL	C	401	-	-	0/4/4/4	0/0/0/0
4	GAL	D	1301	5	-	0/2/22/22	0/1/1/1
5	SIA	D	1302	4	-	0/14/34/38	0/1/1/1
3	GOL	D	1303	-	-	0/4/4/4	0/0/0/0
6	BGC	E	1301	4	-	0/2/22/22	0/1/1/1
4	GAL	E	1302	5,7,6	-	0/2/19/22	0/1/1/1
5	SIA	E	1303	4	-	0/14/34/38	0/1/1/1
7	NGA	E	1304	4	-	0/6/23/26	0/1/1/1
4	GAL	E	1305	7	-	0/2/19/22	0/1/1/1
6	BGC	F	1301	4	-	0/2/22/22	0/1/1/1
4	GAL	F	1302	5,7,6	-	0/2/19/22	0/1/1/1
5	SIA	F	1303	4	-	0/14/34/38	0/1/1/1
7	NGA	F	1304	4	-	0/6/23/26	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(°)
5	F	1303	SIA	C3-C4-C5	-4.38	106.59	111.47
4	E	1302	GAL	O3-C3-C2	-3.57	103.46	110.01
4	E	1302	GAL	O5-C5-C4	-3.37	104.55	110.13
3	C	401	GOL	C3-C2-C1	-2.26	101.52	111.06
4	F	1302	GAL	C2-C3-C4	-2.17	107.27	111.05
5	E	1303	SIA	O7-C7-C6	-2.14	104.55	109.48
4	E	1302	GAL	O4-C4-C5	-2.13	103.62	109.23
5	D	1302	SIA	O9-C9-C8	-2.08	106.46	111.07
4	D	1301	GAL	C4-C3-C2	2.01	114.49	110.79
4	F	1302	GAL	O5-C5-C6	2.15	111.95	107.34
5	D	1302	SIA	C3-C4-C5	2.56	114.33	111.47
4	E	1305	GAL	C1-O5-C5	2.63	116.00	112.14
4	D	1301	GAL	O1-C1-C2	2.77	116.71	109.05
5	E	1303	SIA	O8-C8-C9	2.84	115.97	109.23
4	F	1302	GAL	C1-O5-C5	3.77	117.69	112.14
7	E	1304	NGA	O5-C5-C6	3.84	115.56	107.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	401	GOL	1	0
3	B	401	GOL	1	0
3	C	401	GOL	3	0
5	F	1303	SIA	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, 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	293/317 (92%)	-0.29	2 (0%) 89 89	37, 54, 85, 124	0
1	B	292/317 (92%)	-0.27	3 (1%) 84 85	39, 58, 94, 130	0
1	C	291/317 (91%)	0.10	14 (4%) 34 36	45, 71, 117, 135	0
2	D	126/129 (97%)	-0.31	3 (2%) 62 64	43, 56, 92, 131	0
2	E	122/129 (94%)	-0.17	4 (3%) 50 53	46, 68, 96, 121	0
2	F	104/129 (80%)	1.51	36 (34%) 0 0	54, 108, 152, 170	0
All	All	1228/1338 (91%)	-0.03	62 (5%) 32 34	37, 63, 117, 170	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	41	TRP	6.2
2	F	1179	ILE	5.9
2	F	1180	ALA	5.9
2	F	1213	THR	5.6
2	F	1130	PRO	5.6
2	F	1214	LEU	5.5
2	F	1128	TYR	5.4
2	F	1153	ARG	5.3
2	F	1177	TYR	5.3
2	F	1178	ASN	5.1
2	F	1199	PHE	5.0
2	F	1154	ASN	4.9
2	F	1176	ASN	4.7
2	F	1157	TRP	4.6
2	F	1184	THR	4.6
2	F	1129	ALA	4.4
2	F	1127	VAL	4.3
1	C	44	PHE	4.2
1	C	43	LYS	4.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	1181	LEU	3.9
2	F	1131	ALA	3.8
2	F	1110	GLY	3.8
1	C	4	ALA	3.7
2	E	1173	ILE	3.6
1	C	269	TYR	3.6
2	F	1225	TYR	3.6
1	C	46	LEU	3.5
2	F	1174	MET	3.4
1	C	5	GLU	3.3
1	C	42	GLU	3.3
1	B	300	ASN	3.2
1	C	7	LEU	3.2
2	D	1210	ARG	3.1
2	F	1170	SER	3.1
2	F	1125	LEU	3.1
2	F	1156	GLN	3.1
2	F	1155	GLY	3.0
2	E	1188	LYS	3.0
2	F	1171	ARG	2.9
1	B	273	TYR	2.9
1	C	6	ARG	2.8
1	C	40	GLN	2.7
2	F	1215	ARG	2.6
1	C	254	ASP	2.6
2	F	1169	ILE	2.6
2	D	1209	SER	2.5
2	F	1172	GLU	2.4
1	A	294	PRO	2.4
2	F	1183	TRP	2.4
2	F	1152	CYS	2.4
1	C	261	ARG	2.3
2	D	1106	PHE	2.3
2	F	1186	LYS	2.3
2	F	1185	ALA	2.3
2	F	1182	ARG	2.3
2	E	1177	TYR	2.2
1	C	47	GLU	2.2
2	F	1216	THR	2.1
1	B	143	ASN	2.1
1	A	261	ARG	2.1
2	E	1190[A]	TYR	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	1133[A]	SER	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
4	GAL	E	1305	11/12	0.93	0.10	-	89,95,103,116	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
3	GOL	D	1303	6/6	0.91	0.25	4.50	78,86,91,97	0
5	SIA	D	1302	20/21	0.91	0.25	1.99	52,88,104,109	0
3	GOL	A	401	6/6	0.98	0.17	1.30	52,58,61,71	0
3	GOL	C	401	6/6	0.95	0.15	0.99	53,72,73,84	0
5	SIA	E	1303	20/21	0.97	0.10	-0.61	52,59,82,91	0
5	SIA	F	1303	20/21	0.92	0.15	-0.82	93,105,123,125	0
6	BGC	E	1301	12/12	0.95	0.08	-1.43	66,72,80,87	0
3	GOL	B	401	6/6	0.98	0.11	-1.68	40,53,64,64	0
4	GAL	E	1302	11/12	0.97	0.06	-2.29	58,68,78,96	0
7	NGA	E	1304	14/15	0.94	0.08	-	70,83,90,98	0
6	BGC	F	1301	12/12	0.64	0.32	-	121,134,140,143	0
4	GAL	F	1302	11/12	0.92	0.12	-	93,107,121,126	0
4	GAL	D	1301	12/12	0.76	0.21	-	78,98,101,105	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	GAL	E	1305	11/12	0.93	0.10	-	89,95,103,116	0
7	NGA	F	1304	14/15	0.90	0.16	-	97,109,112,120	0

6.5 Other polymers [i](#)

There are no such residues in this entry.