Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 03:56 PM BST

PDB ID	:	6APJ
Title	:	Crystal Structure of human ST6GALNAC2
Authors	:	Forouhar, F.; Moremen, K.W.; Northeast Structural Genomics Consortium (NESG); Tong, L.
Deposited on	:	2017-08-17
Resolution	:	3.10 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see references ①) were used in the production of this report:

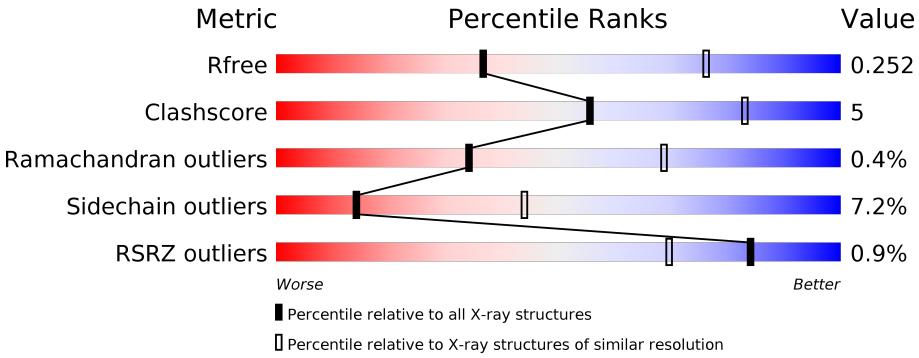
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 3.10 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}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	374	63% 13% • 22%
1	B	374	63% 11% • 24%
1	C	374	% 65% 11% • 23%
1	D	374	2% 63% 11% • 23%
1	E	374	% 65% 11% • 23%
1	F	374	% 66% 11% • 21%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13939 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.

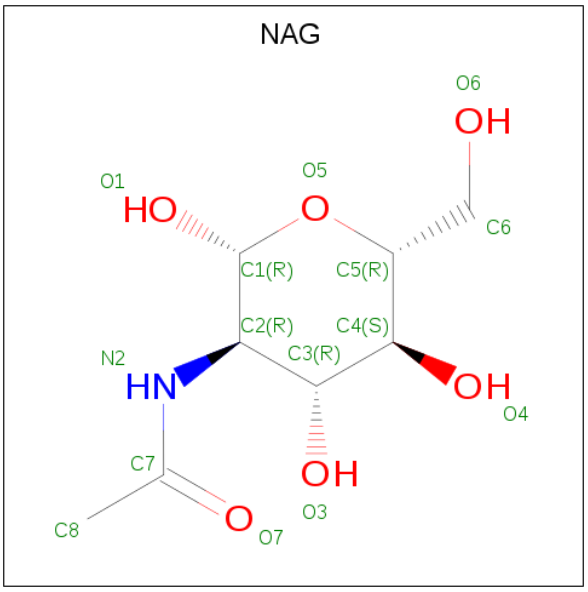
- Molecule 1 is a protein called Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	291	Total	C	N	O	S	Se	0	0	0
			2325	1507	404	406	4	4			
1	B	285	Total	C	N	O	S	Se	0	0	0
			2273	1472	395	398	4	4			
1	C	289	Total	C	N	O	S	Se	0	0	0
			2309	1497	400	404	4	4			
1	D	288	Total	C	N	O	S	Se	0	0	0
			2297	1488	399	402	4	4			
1	E	288	Total	C	N	O	S	Se	0	0	0
			2307	1497	401	400	5	4			
1	F	296	Total	C	N	O	S	Se	0	0	0
			2372	1540	412	411	5	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	initiating methionine	UNP Q9UJ37
B	1	MSE	-	initiating methionine	UNP Q9UJ37
C	1	MSE	-	initiating methionine	UNP Q9UJ37
D	1	MSE	-	initiating methionine	UNP Q9UJ37
E	1	MSE	-	initiating methionine	UNP Q9UJ37
F	1	MSE	-	initiating methionine	UNP Q9UJ37

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



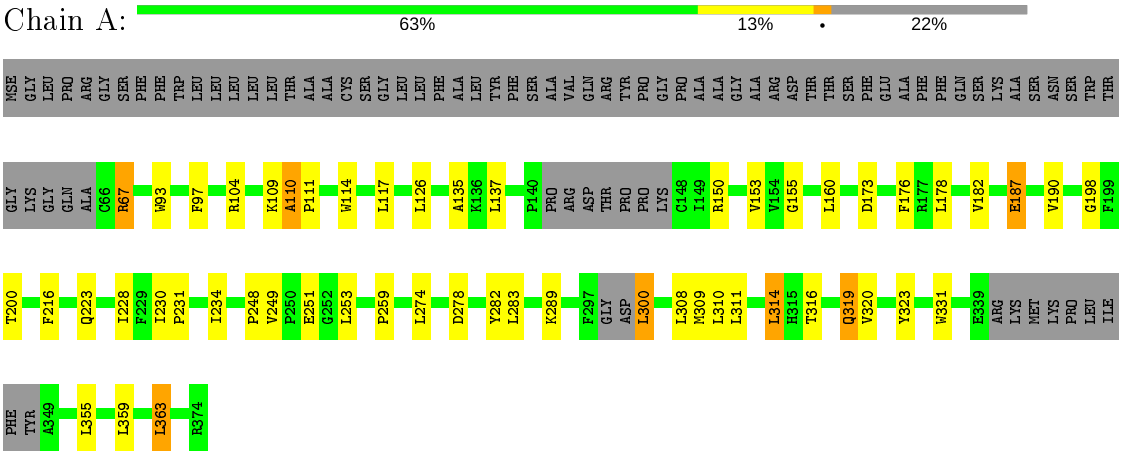
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C N O 14 8 1 5	0	0
2	A	1	Total C N O 14 8 1 5	0	0
2	E	1	Total C N O 14 8 1 5	0	0
2	F	1	Total C N O 14 8 1 5	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

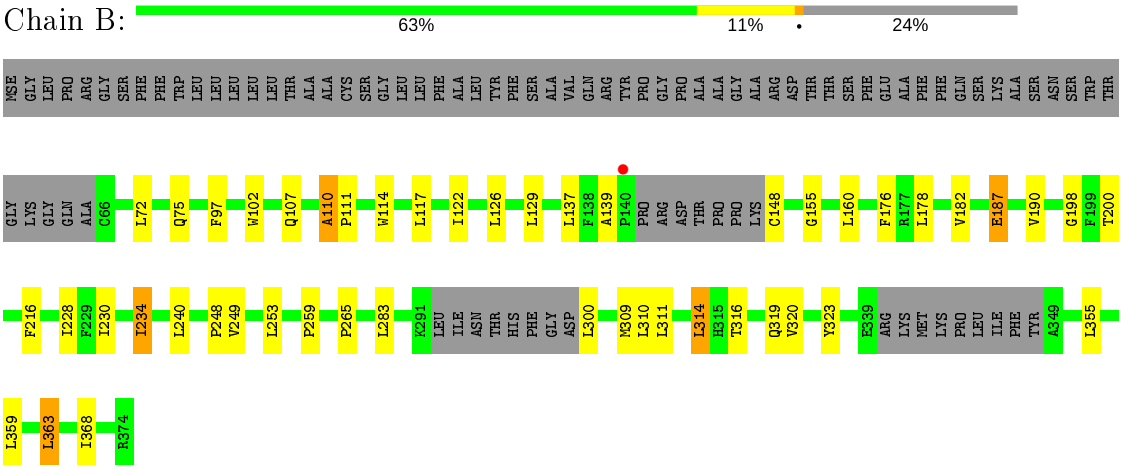
- Molecule 1: Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2

Chain A:



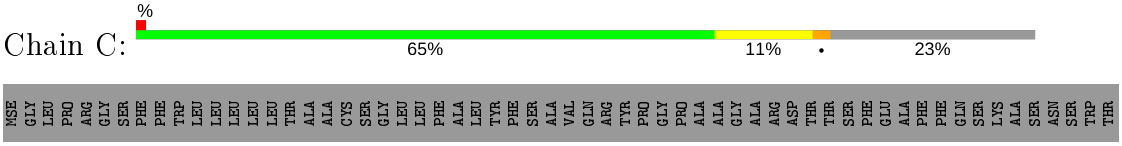
- Molecule 1: Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2

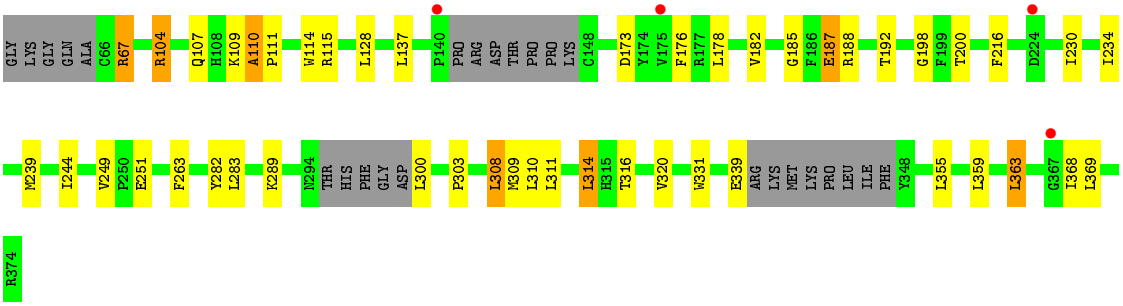
Chain B:



- Molecule 1: Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2

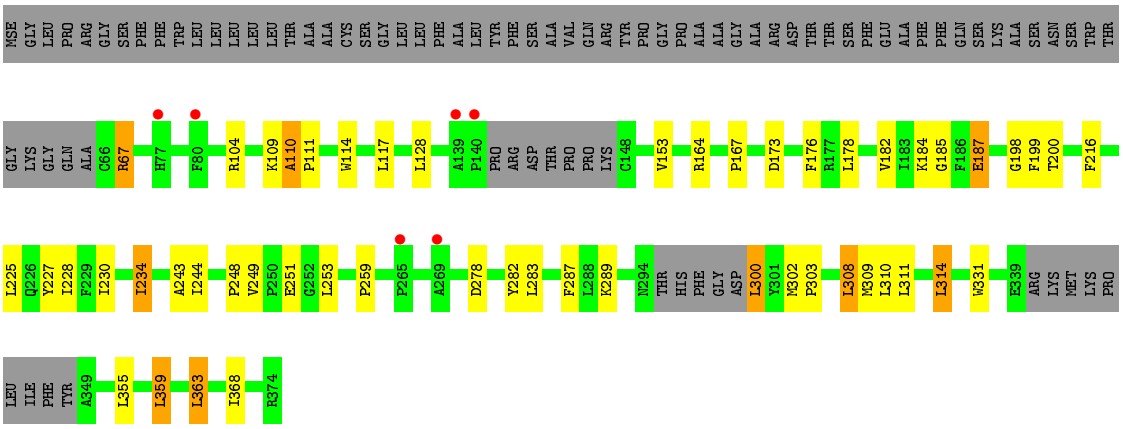
Chain C:





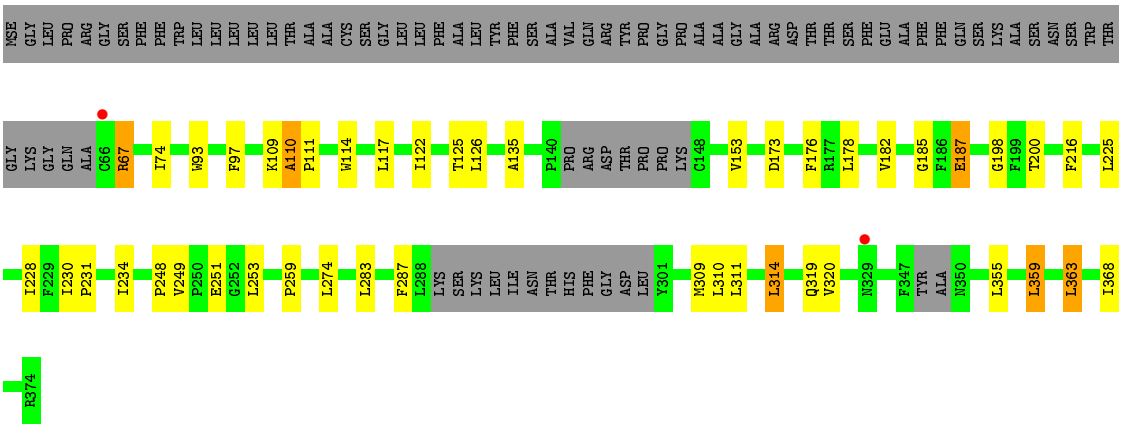
• Molecule 1: Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2

Chain D:



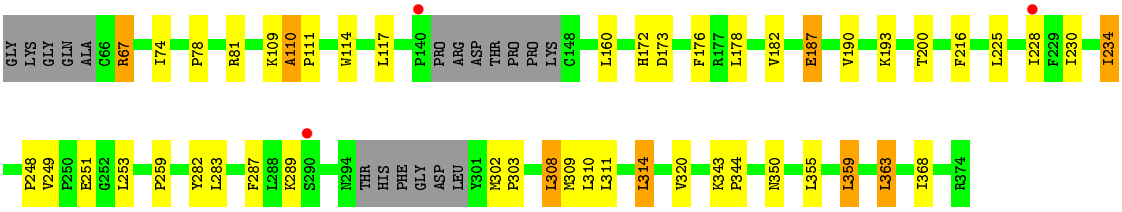
• Molecule 1: Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2

Chain E:



• Molecule 1: Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2

Chain F:


GLY LYS GLY GLN ALA C66 R67 I74 P78 R81 K109 A110 P111 W114 L117 P140 PRO ARG ASP THR PRO PRO LYS C148 L160 H172 D173 F176 R177 L178 V182 E187 V190 K193 T200 F216 L225 I228 F229 I230 I234
P248 V249 P250 E251 G252 L253 P259 Y282 L283 F287 L288 K289 S290 M294 THR HIS PHE GLY ASP LEU Y301 M302 P303 L308 M309 L310 L311 L314 V320 K343 P344 N350 L355 L359 L363 I368 R374

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	71.35Å 71.80Å 134.25Å 98.78° 101.92° 103.48°	Depositor
Resolution (Å)	44.28 – 3.10 44.28 – 3.10	Depositor EDS
% Data completeness (in resolution range)	93.2 (44.28-3.10) 93.2 (44.28-3.10)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.72 (at 3.12Å)	Xtriage
Refinement program	PHENIX, REFMAC, CNS	Depositor
R, R_{free}	0.201 , 0.252 0.201 , 0.252	Depositor DCC
R_{free} test set	4186 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å^2)	91.0	Xtriage
Anisotropy	0.307	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 63.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.021 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	13939	wwPDB-VP
Average B, all atoms (Å^2)	100.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.22% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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	0.24	0/2388	0.42	0/3232
1	B	0.25	0/2334	0.42	0/3158
1	C	0.24	0/2371	0.42	0/3209
1	D	0.24	0/2358	0.42	0/3191
1	E	0.24	0/2370	0.41	0/3206
1	F	0.24	0/2437	0.42	0/3297
All	All	0.24	0/14258	0.42	0/19293

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

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	2325	0	2281	27	0
1	B	2273	0	2232	23	0
1	C	2309	0	2269	22	0
1	D	2297	0	2260	29	0
1	E	2307	0	2270	23	0
1	F	2372	0	2344	22	0
2	A	28	0	26	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	14	0	13	1	0
2	F	14	0	13	1	0
All	All	13939	0	13708	140	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (140) 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:A:182:VAL:HG22	1:A:216:PHE:HB2	1.65	0.78
1:B:182:VAL:HG22	1:B:216:PHE:HB2	1.67	0.75
1:C:182:VAL:HG22	1:C:216:PHE:HB2	1.67	0.74
1:D:182:VAL:HG22	1:D:216:PHE:HB2	1.73	0.70
1:E:228:ILE:HD12	1:E:309:MSE:HE1	1.72	0.70
1:B:228:ILE:HD12	1:B:309:MSE:HE1	1.75	0.68
1:A:228:ILE:HD12	1:A:309:MSE:HE1	1.77	0.66
1:A:223:GLN:OE1	1:C:188:ARG:NH2	2.33	0.62
1:D:104:ARG:NH2	1:D:278:ASP:OD2	2.32	0.61
1:F:228:ILE:HD12	1:F:309:MSE:HE1	1.81	0.61
1:B:187:GLU:HG3	1:C:331:TRP:CE2	2.36	0.61
1:E:117:LEU:HG	1:E:234:ILE:HG23	1.82	0.61
1:D:185:GLY:N	1:D:187:GLU:OE1	2.32	0.61
2:E:401:NAG:H83	2:E:401:NAG:H3	1.83	0.60
1:D:117:LEU:HG	1:D:234:ILE:HG23	1.84	0.59
1:F:200:THR:HG22	1:F:230:ILE:HB	1.83	0.59
2:F:401:NAG:H3	2:F:401:NAG:H83	1.83	0.59
1:D:110:ALA:HB1	1:D:111:PRO:HD2	1.85	0.59
1:C:110:ALA:HB1	1:C:111:PRO:HD2	1.85	0.59
1:A:200:THR:HG22	1:A:230:ILE:HB	1.85	0.58
1:B:200:THR:HG22	1:B:230:ILE:HB	1.85	0.58
1:E:176:PHE:HD2	1:E:309:MSE:HE3	1.69	0.58
1:D:200:THR:HG22	1:D:230:ILE:HB	1.87	0.57
1:E:110:ALA:HB1	1:E:111:PRO:HD2	1.88	0.56
1:E:182:VAL:HG22	1:E:216:PHE:HB2	1.87	0.56
1:B:110:ALA:HB1	1:B:111:PRO:HD2	1.88	0.56
1:F:343:LYS:HD3	1:F:344:PRO:HD2	1.88	0.56
1:A:187:GLU:HG3	1:D:331:TRP:CZ2	2.42	0.55
1:A:110:ALA:HB1	1:A:111:PRO:HD2	1.88	0.55
1:B:283:LEU:HD22	1:B:311:LEU:HD13	1.89	0.55
1:F:182:VAL:HG22	1:F:216:PHE:HB2	1.89	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:283:LEU:HD22	1:A:311:LEU:HD13	1.89	0.55
1:E:363:LEU:HD12	1:E:368:ILE:HB	1.89	0.55
1:F:110:ALA:HB1	1:F:111:PRO:HD2	1.88	0.54
1:C:200:THR:HG22	1:C:230:ILE:HB	1.90	0.54
1:D:287:PHE:HB3	1:D:359:LEU:HD11	1.90	0.53
1:A:187:GLU:HG3	1:D:331:TRP:CE2	2.43	0.53
1:D:67:ARG:HD3	1:D:67:ARG:H	1.74	0.53
1:C:283:LEU:HD22	1:C:311:LEU:HD13	1.91	0.53
1:B:265:PRO:HB2	1:D:167:PRO:HG2	1.91	0.53
1:B:97:PHE:HB2	1:B:126:LEU:HD23	1.90	0.53
1:F:248:PRO:HA	1:F:259:PRO:HD3	1.90	0.53
1:A:331:TRP:CZ2	1:F:187:GLU:HG3	2.44	0.52
1:D:359:LEU:HD22	1:D:363:LEU:HD22	1.91	0.52
1:E:314:LEU:HD22	1:E:320:VAL:HG11	1.91	0.52
1:A:97:PHE:HB2	1:A:126:LEU:HD23	1.90	0.52
1:E:109:LYS:HG2	1:E:110:ALA:H	1.75	0.52
1:A:248:PRO:HA	1:A:259:PRO:HD3	1.92	0.52
1:F:117:LEU:HG	1:F:234:ILE:HG23	1.91	0.52
1:F:363:LEU:HD12	1:F:368:ILE:HB	1.90	0.52
1:A:67:ARG:HD3	1:A:67:ARG:H	1.74	0.52
1:B:176:PHE:HD2	1:B:309:MSE:HE3	1.75	0.52
1:A:176:PHE:HD2	1:A:309:MSE:HE3	1.75	0.51
1:C:128:LEU:HD23	1:C:244:ILE:HB	1.92	0.51
1:E:200:THR:HG22	1:E:230:ILE:HB	1.92	0.51
1:A:104:ARG:NH2	1:A:278:ASP:OD2	2.42	0.51
1:C:137:LEU:HD12	1:C:316:THR:HG21	1.92	0.51
1:F:160:LEU:HD13	1:F:190:VAL:HG12	1.93	0.51
1:F:303:PRO:HB2	1:F:308:LEU:HD13	1.92	0.50
1:F:283:LEU:HD22	1:F:311:LEU:HD13	1.94	0.50
1:A:160:LEU:HD13	1:A:190:VAL:HG12	1.92	0.50
1:D:248:PRO:HA	1:D:259:PRO:HD3	1.92	0.50
1:B:155:GLY:N	1:B:323:TYR:O	2.44	0.50
1:C:109:LYS:HD2	1:C:115:ARG:HH11	1.77	0.50
1:E:97:PHE:HB2	1:E:126:LEU:HD23	1.94	0.50
1:F:67:ARG:H	1:F:67:ARG:HD3	1.76	0.49
1:A:231:PRO:HB3	1:A:274:LEU:HD11	1.94	0.49
1:E:67:ARG:H	1:E:67:ARG:HD3	1.77	0.49
1:B:363:LEU:HD12	1:B:368:ILE:HB	1.93	0.49
1:B:160:LEU:HD13	1:B:190:VAL:HG12	1.95	0.49
1:C:185:GLY:N	1:C:187:GLU:OE1	2.44	0.49
1:C:303:PRO:HB2	1:C:308:LEU:HD13	1.95	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:67:ARG:H	1:C:67:ARG:HD3	1.78	0.49
1:D:303:PRO:HB2	1:D:308:LEU:HD13	1.94	0.48
1:B:314:LEU:HD22	1:B:320:VAL:HG11	1.96	0.48
1:C:314:LEU:HD22	1:C:320:VAL:HG11	1.95	0.48
1:E:153:VAL:HA	1:E:176:PHE:HB2	1.96	0.48
1:D:153:VAL:HA	1:D:176:PHE:HB2	1.96	0.48
1:F:109:LYS:HG2	1:F:110:ALA:H	1.79	0.47
1:C:176:PHE:HD2	1:C:309:MSE:HE3	1.79	0.47
1:A:153:VAL:HA	1:A:176:PHE:HB2	1.96	0.47
1:D:111:PRO:HG3	1:D:302:MSE:SE	2.64	0.47
1:E:234:ILE:H	1:E:234:ILE:HG13	1.49	0.47
1:F:287:PHE:HB3	1:F:359:LEU:HD11	1.97	0.47
1:D:283:LEU:HD22	1:D:311:LEU:HD13	1.97	0.46
1:A:110:ALA:HB2	1:A:300:LEU:HG	1.97	0.46
1:B:117:LEU:HG	1:B:234:ILE:HG23	1.96	0.46
1:C:363:LEU:HD12	1:C:368:ILE:HB	1.98	0.46
1:D:363:LEU:HD12	1:D:368:ILE:HB	1.97	0.46
1:E:74:ILE:HG12	1:E:368:ILE:HD13	1.97	0.46
1:E:248:PRO:HA	1:E:259:PRO:HD3	1.98	0.46
1:B:102:TRP:HZ3	1:B:122:ILE:HD12	1.80	0.46
1:E:231:PRO:HB3	1:E:274:LEU:HD11	1.97	0.46
1:E:198:GLY:HA3	1:E:309:MSE:SE	2.65	0.46
1:E:283:LEU:HD22	1:E:311:LEU:HD13	1.98	0.46
1:B:187:GLU:HG3	1:C:331:TRP:CZ2	2.51	0.46
1:E:287:PHE:HB3	1:E:359:LEU:HD11	1.97	0.46
1:F:111:PRO:HG3	1:F:302:MSE:SE	2.66	0.46
1:B:248:PRO:HA	1:B:259:PRO:HD3	1.99	0.45
1:D:128:LEU:HD23	1:D:244:ILE:HB	1.98	0.45
1:B:265:PRO:HG3	1:D:164:ARG:HG2	1.99	0.45
1:F:172:HIS:O	1:F:193:LYS:NZ	2.40	0.45
1:F:176:PHE:HD2	1:F:309:MSE:HE3	1.81	0.45
1:B:198:GLY:HA3	1:B:309:MSE:SE	2.67	0.45
1:D:198:GLY:HA3	1:D:309:MSE:SE	2.67	0.45
1:D:243:ALA:HA	1:D:259:PRO:HB3	1.98	0.44
1:E:185:GLY:N	1:E:187:GLU:OE1	2.38	0.44
1:A:137:LEU:HD12	1:A:316:THR:HG21	2.00	0.44
1:D:110:ALA:HB2	1:D:300:LEU:HG	1.98	0.44
1:A:117:LEU:HG	1:A:234:ILE:HG23	2.00	0.44
1:B:234:ILE:HG13	1:B:234:ILE:H	1.50	0.43
1:D:109:LYS:HG2	1:D:110:ALA:H	1.83	0.43
1:A:109:LYS:HG2	1:A:110:ALA:H	1.83	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:234:ILE:HG13	1:D:234:ILE:H	1.49	0.43
1:C:198:GLY:HA3	1:C:309:MSE:SE	2.68	0.43
1:B:363:LEU:HA	1:B:363:LEU:HD13	1.87	0.43
1:A:150:ARG:NE	1:A:319:GLN:HG2	2.34	0.42
1:F:314:LEU:HD13	1:F:314:LEU:HA	1.84	0.42
1:F:314:LEU:HD22	1:F:320:VAL:HG11	2.00	0.42
1:D:199:PHE:HE1	1:D:227:TYR:HB3	1.84	0.42
1:F:78:PRO:HB3	1:F:81:ARG:HH12	1.84	0.42
1:C:187:GLU:HB3	1:C:192:THR:HG23	2.02	0.42
1:C:363:LEU:HA	1:C:363:LEU:HD13	1.91	0.42
1:A:155:GLY:N	1:A:323:TYR:O	2.47	0.42
1:D:228:ILE:HD12	1:D:309:MSE:HE1	2.01	0.42
1:E:93:TRP:CE2	1:E:135:ALA:HB2	2.55	0.42
1:A:198:GLY:HA3	1:A:309:MSE:SE	2.70	0.41
1:B:137:LEU:HD12	1:B:316:THR:HG21	2.02	0.41
1:C:239:MSE:HE2	1:C:263:PHE:HE2	1.85	0.41
1:C:104:ARG:O	1:C:107:GLN:HG2	2.20	0.41
1:A:314:LEU:HD22	1:A:320:VAL:HG11	2.03	0.41
1:D:314:LEU:HA	1:D:314:LEU:HD13	1.85	0.41
1:C:363:LEU:HB3	1:C:369:LEU:HB3	2.03	0.41
1:D:184:LYS:HD2	1:D:184:LYS:HA	1.88	0.41
1:E:359:LEU:HA	1:E:359:LEU:HD23	1.84	0.41
1:F:74:ILE:HG12	1:F:368:ILE:HD13	2.02	0.41
1:B:129:LEU:HD11	1:B:240:LEU:HD21	2.02	0.40
1:E:122:ILE:HA	1:E:125:THR:HG22	2.03	0.40
1:A:363:LEU:HD12	1:A:363:LEU:HA	1.90	0.40
1:A:93:TRP:CE2	1:A:135:ALA:HB2	2.57	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	283/374 (76%)	264 (93%)	18 (6%)	1 (0%)	34	69
1	B	277/374 (74%)	255 (92%)	20 (7%)	2 (1%)	22	57
1	C	281/374 (75%)	261 (93%)	19 (7%)	1 (0%)	34	69
1	D	280/374 (75%)	258 (92%)	21 (8%)	1 (0%)	34	69
1	E	280/374 (75%)	265 (95%)	14 (5%)	1 (0%)	34	69
1	F	290/374 (78%)	266 (92%)	23 (8%)	1 (0%)	41	73
All	All	1691/2244 (75%)	1569 (93%)	115 (7%)	7 (0%)	34	69

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	110	ALA
1	B	110	ALA
1	C	110	ALA
1	D	110	ALA
1	E	110	ALA
1	F	110	ALA
1	B	139	ALA

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	248/308 (80%)	230 (93%)	18 (7%)	14	43
1	B	242/308 (79%)	225 (93%)	17 (7%)	15	45
1	C	246/308 (80%)	227 (92%)	19 (8%)	13	41
1	D	245/308 (80%)	226 (92%)	19 (8%)	12	40
1	E	246/308 (80%)	231 (94%)	15 (6%)	18	49
1	F	253/308 (82%)	234 (92%)	19 (8%)	13	42
All	All	1480/1848 (80%)	1373 (93%)	107 (7%)	14	44

All (107) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	67	ARG
1	A	114	TRP
1	A	173	ASP
1	A	178	LEU
1	A	187	GLU
1	A	249	VAL
1	A	251	GLU
1	A	253	LEU
1	A	282	TYR
1	A	289	LYS
1	A	300	LEU
1	A	308	LEU
1	A	310	LEU
1	A	314	LEU
1	A	319	GLN
1	A	355	LEU
1	A	359	LEU
1	A	363	LEU
1	B	72	LEU
1	B	75	GLN
1	B	107	GLN
1	B	114	TRP
1	B	148	CYS
1	B	178	LEU
1	B	187	GLU
1	B	234	ILE
1	B	249	VAL
1	B	253	LEU
1	B	300	LEU
1	B	310	LEU
1	B	314	LEU
1	B	319	GLN
1	B	355	LEU
1	B	359	LEU
1	B	363	LEU
1	C	67	ARG
1	C	104	ARG
1	C	114	TRP
1	C	173	ASP
1	C	178	LEU
1	C	187	GLU
1	C	234	ILE
1	C	249	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	251	GLU
1	C	282	TYR
1	C	289	LYS
1	C	300	LEU
1	C	308	LEU
1	C	310	LEU
1	C	314	LEU
1	C	339	GLU
1	C	355	LEU
1	C	359	LEU
1	C	363	LEU
1	D	67	ARG
1	D	114	TRP
1	D	173	ASP
1	D	178	LEU
1	D	187	GLU
1	D	225	LEU
1	D	234	ILE
1	D	249	VAL
1	D	251	GLU
1	D	253	LEU
1	D	282	TYR
1	D	289	LYS
1	D	300	LEU
1	D	308	LEU
1	D	310	LEU
1	D	314	LEU
1	D	355	LEU
1	D	359	LEU
1	D	363	LEU
1	E	67	ARG
1	E	114	TRP
1	E	173	ASP
1	E	178	LEU
1	E	187	GLU
1	E	225	LEU
1	E	249	VAL
1	E	251	GLU
1	E	253	LEU
1	E	310	LEU
1	E	314	LEU
1	E	319	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	355	LEU
1	E	359	LEU
1	E	363	LEU
1	F	67	ARG
1	F	114	TRP
1	F	173	ASP
1	F	178	LEU
1	F	187	GLU
1	F	225	LEU
1	F	234	ILE
1	F	249	VAL
1	F	251	GLU
1	F	253	LEU
1	F	282	TYR
1	F	289	LYS
1	F	308	LEU
1	F	310	LEU
1	F	314	LEU
1	F	350	ASN
1	F	355	LEU
1	F	359	LEU
1	F	363	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	319	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

4 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	$\#\|Z\| > 2$
2	NAG	A	501	1	14,14,15	0.27	0	17,19,21	0.38	0
2	NAG	E	401	1	14,14,15	0.62	1 (7%)	17,19,21	1.30	1 (5%)
2	NAG	F	401	1	14,14,15	0.56	0	17,19,21	1.30	2 (11%)
2	NAG	A	502	1	14,14,15	0.34	0	17,19,21	0.54	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
2	NAG	A	501	1	-	2/6/23/26	0/1/1/1
2	NAG	E	401	1	-	3/6/23/26	0/1/1/1
2	NAG	F	401	1	-	5/6/23/26	0/1/1/1
2	NAG	A	502	1	-	1/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	401	NAG	C1-C2	2.06	1.55	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	401	NAG	C2-N2-C7	4.30	129.02	122.90
2	E	401	NAG	C2-N2-C7	4.23	128.93	122.90
2	F	401	NAG	C1-C2-N2	2.09	114.06	110.49

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	401	NAG	O5-C5-C6-O6
2	F	401	NAG	C4-C5-C6-O6
2	E	401	NAG	C8-C7-N2-C2
2	E	401	NAG	O7-C7-N2-C2
2	F	401	NAG	C8-C7-N2-C2
2	F	401	NAG	O7-C7-N2-C2
2	A	501	NAG	O5-C5-C6-O6
2	A	501	NAG	C4-C5-C6-O6
2	F	401	NAG	C3-C2-N2-C7
2	A	502	NAG	C3-C2-N2-C7
2	E	401	NAG	C3-C2-N2-C7

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	401	NAG	1	0
2	F	401	NAG	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	287/374 (76%)	-0.44	0 100 100	46, 73, 130, 228	0
1	B	281/374 (75%)	-0.38	1 (0%) 92 84	45, 74, 123, 229	0
1	C	285/374 (76%)	-0.17	4 (1%) 75 56	56, 102, 160, 237	0
1	D	284/374 (75%)	-0.13	6 (2%) 63 43	50, 102, 172, 231	0
1	E	284/374 (75%)	-0.16	2 (0%) 87 75	59, 111, 174, 268	0
1	F	292/374 (78%)	-0.21	3 (1%) 82 67	56, 106, 168, 226	0
All	All	1713/2244 (76%)	-0.25	16 (0%) 84 69	45, 93, 167, 268	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	140	PRO	4.0
1	D	140	PRO	3.6
1	D	77	HIS	3.4
1	E	329	ASN	3.2
1	B	140	PRO	2.9
1	C	367	GLY	2.7
1	C	224	ASP	2.6
1	D	265	PRO	2.4
1	F	228	ILE	2.3
1	D	139	ALA	2.3
1	D	269	ALA	2.2
1	F	290	SER	2.2
1	E	66	CYS	2.2
1	D	80	PHE	2.2
1	F	140	PRO	2.0
1	C	175	VAL	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	F	401	14/15	0.70	0.22	133,171,182,193	0
2	NAG	E	401	14/15	0.73	0.23	135,157,168,182	0
2	NAG	A	502	14/15	0.86	0.18	150,163,173,177	0
2	NAG	A	501	14/15	0.92	0.14	123,136,160,171	0

6.5 Other polymers (i)

There are no such residues in this entry.