Full wwPDB X-ray Structure Validation Report (i)

Dec 14, 2023 – 04:00 am GMT

PDB ID	:	4AD7
Title	:	Crystal structure of full-length N-glycosylated human glypican-1
Authors	:	Svensson, G.; Awad, W.; Mani, K.; Logan, D.T.
Deposited on	:	2011-12-22
Resolution	:	2.94 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

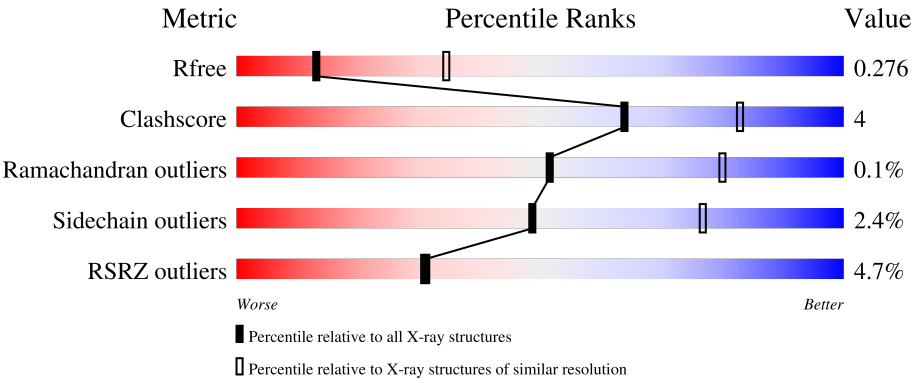
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.94 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	2969 (2.98-2.90)
Clashscore	141614	3218 (2.98-2.90)
Ramachandran outliers	138981	3122 (2.98-2.90)
Sidechain outliers	138945	3124 (2.98-2.90)
RSRZ outliers	127900	2902 (2.98-2.90)

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 of 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	528	5% 59% 9% 32%
1	B	528	3% 68% 12% • 19%
1	C	528	3% 66% 7% 27%
1	D	528	3% 70% 11% • 19%

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
2	NAG	B	1500	X	-	-	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 12754 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 GLYPICAN-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	358	Total	C	N	O	S	0	0	0
			2811	1751	512	526	22			
1	B	426	Total	C	N	O	S	0	0	0
			3341	2085	608	624	24			
1	C	385	Total	C	N	O	S	0	0	0
			3035	1894	554	565	22			
1	D	428	Total	C	N	O	S	0	0	0
			3357	2093	611	629	24			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	-	expression tag	UNP P35052
A	3	PRO	-	expression tag	UNP P35052
A	4	GLN	-	expression tag	UNP P35052
A	5	LEU	-	expression tag	UNP P35052
A	6	HIS	-	expression tag	UNP P35052
A	7	HIS	-	expression tag	UNP P35052
A	8	HIS	-	expression tag	UNP P35052
A	9	HIS	-	expression tag	UNP P35052
A	10	HIS	-	expression tag	UNP P35052
A	11	HIS	-	expression tag	UNP P35052
A	12	ASP	-	expression tag	UNP P35052
A	13	LEU	-	expression tag	UNP P35052
A	14	TYR	-	expression tag	UNP P35052
A	15	GLU	-	expression tag	UNP P35052
A	16	ASN	-	expression tag	UNP P35052
A	17	LEU	-	expression tag	UNP P35052
A	18	TYR	-	expression tag	UNP P35052
A	19	PHE	-	expression tag	UNP P35052
A	20	GLN	-	expression tag	UNP P35052
A	21	GLY	-	expression tag	UNP P35052
A	22	LYS	-	expression tag	UNP P35052

Continued on next page...

Continued from previous page...

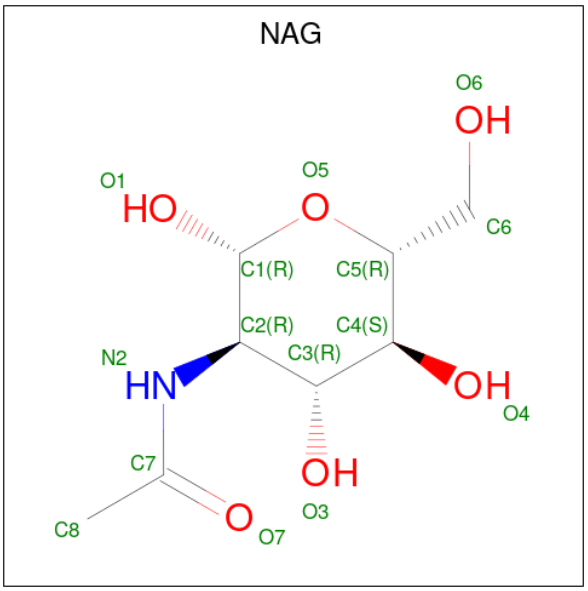
Chain	Residue	Modelled	Actual	Comment	Reference
A	23	LEU	-	expression tag	UNP P35052
B	2	ALA	-	expression tag	UNP P35052
B	3	PRO	-	expression tag	UNP P35052
B	4	GLN	-	expression tag	UNP P35052
B	5	LEU	-	expression tag	UNP P35052
B	6	HIS	-	expression tag	UNP P35052
B	7	HIS	-	expression tag	UNP P35052
B	8	HIS	-	expression tag	UNP P35052
B	9	HIS	-	expression tag	UNP P35052
B	10	HIS	-	expression tag	UNP P35052
B	11	HIS	-	expression tag	UNP P35052
B	12	ASP	-	expression tag	UNP P35052
B	13	LEU	-	expression tag	UNP P35052
B	14	TYR	-	expression tag	UNP P35052
B	15	GLU	-	expression tag	UNP P35052
B	16	ASN	-	expression tag	UNP P35052
B	17	LEU	-	expression tag	UNP P35052
B	18	TYR	-	expression tag	UNP P35052
B	19	PHE	-	expression tag	UNP P35052
B	20	GLN	-	expression tag	UNP P35052
B	21	GLY	-	expression tag	UNP P35052
B	22	LYS	-	expression tag	UNP P35052
B	23	LEU	-	expression tag	UNP P35052
C	2	ALA	-	expression tag	UNP P35052
C	3	PRO	-	expression tag	UNP P35052
C	4	GLN	-	expression tag	UNP P35052
C	5	LEU	-	expression tag	UNP P35052
C	6	HIS	-	expression tag	UNP P35052
C	7	HIS	-	expression tag	UNP P35052
C	8	HIS	-	expression tag	UNP P35052
C	9	HIS	-	expression tag	UNP P35052
C	10	HIS	-	expression tag	UNP P35052
C	11	HIS	-	expression tag	UNP P35052
C	12	ASP	-	expression tag	UNP P35052
C	13	LEU	-	expression tag	UNP P35052
C	14	TYR	-	expression tag	UNP P35052
C	15	GLU	-	expression tag	UNP P35052
C	16	ASN	-	expression tag	UNP P35052
C	17	LEU	-	expression tag	UNP P35052
C	18	TYR	-	expression tag	UNP P35052
C	19	PHE	-	expression tag	UNP P35052
C	20	GLN	-	expression tag	UNP P35052

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	21	GLY	-	expression tag	UNP P35052
C	22	LYS	-	expression tag	UNP P35052
C	23	LEU	-	expression tag	UNP P35052
D	2	ALA	-	expression tag	UNP P35052
D	3	PRO	-	expression tag	UNP P35052
D	4	GLN	-	expression tag	UNP P35052
D	5	LEU	-	expression tag	UNP P35052
D	6	HIS	-	expression tag	UNP P35052
D	7	HIS	-	expression tag	UNP P35052
D	8	HIS	-	expression tag	UNP P35052
D	9	HIS	-	expression tag	UNP P35052
D	10	HIS	-	expression tag	UNP P35052
D	11	HIS	-	expression tag	UNP P35052
D	12	ASP	-	expression tag	UNP P35052
D	13	LEU	-	expression tag	UNP P35052
D	14	TYR	-	expression tag	UNP P35052
D	15	GLU	-	expression tag	UNP P35052
D	16	ASN	-	expression tag	UNP P35052
D	17	LEU	-	expression tag	UNP P35052
D	18	TYR	-	expression tag	UNP P35052
D	19	PHE	-	expression tag	UNP P35052
D	20	GLN	-	expression tag	UNP P35052
D	21	GLY	-	expression tag	UNP P35052
D	22	LYS	-	expression tag	UNP P35052
D	23	LEU	-	expression tag	UNP P35052

- 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	B	1	Total C N O 14 8 1 5	0	0
2	B	1	Total C N O 14 8 1 5	0	0
2	C	1	Total C N O 14 8 1 5	0	0
2	D	1	Total C N O 14 8 1 5	0	0
2	D	1	Total C N O 14 8 1 5	0	0

- Molecule 3 is water.

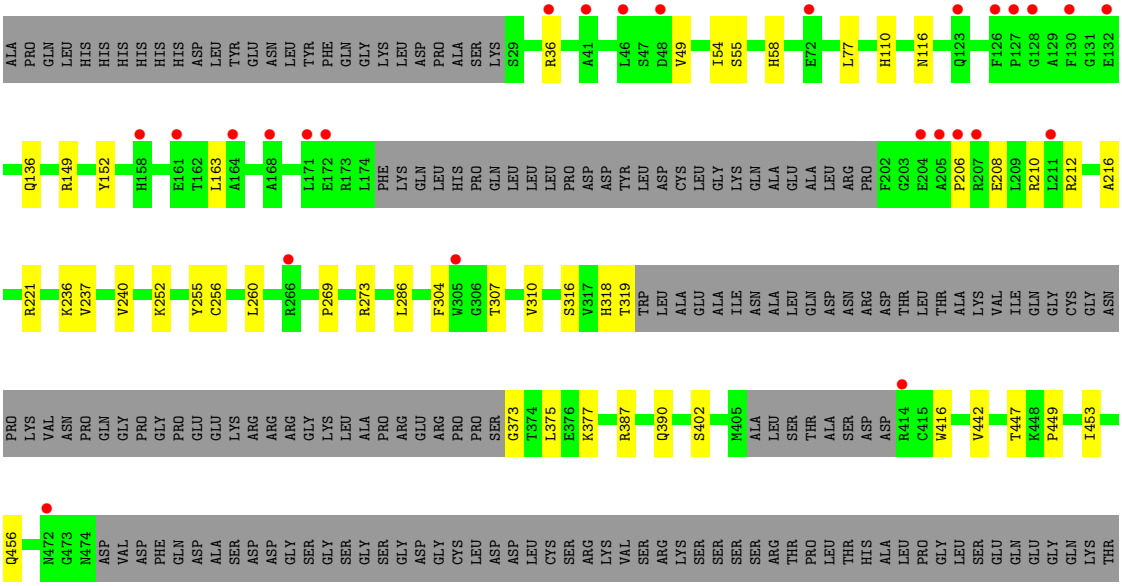
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	18	Total O 18 18	0	0
3	B	39	Total O 39 39	0	0
3	C	30	Total O 30 30	0	0
3	D	39	Total O 39 39	0	0

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.

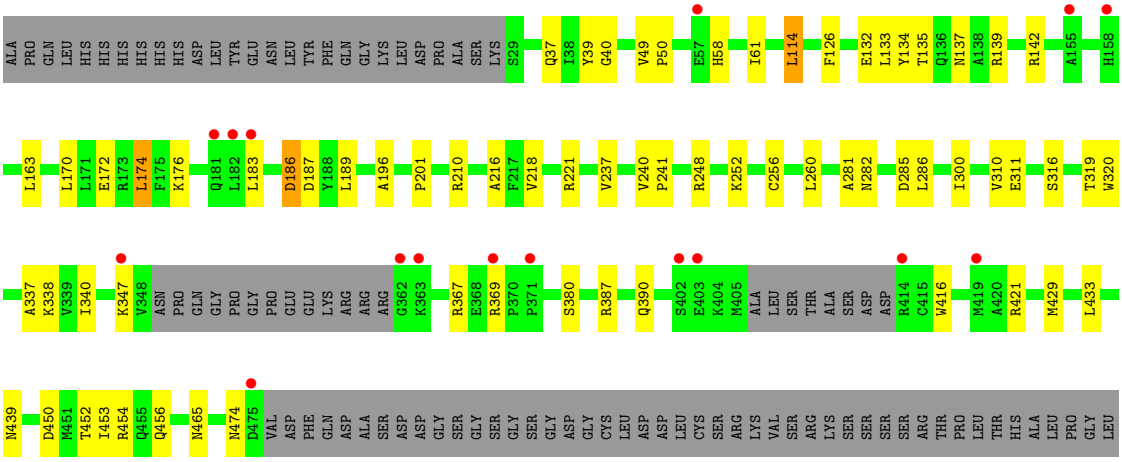
• Molecule 1: GLYPICAN-1

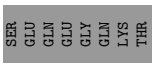
Chain A:



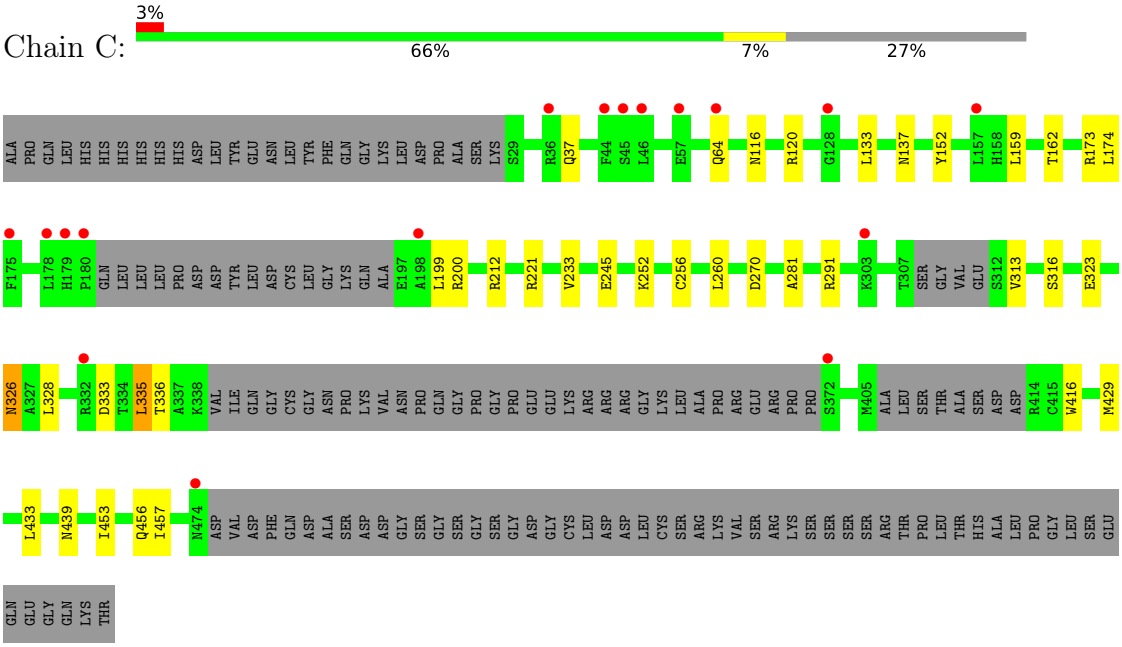
• Molecule 1: GLYPICAN-1

Chain B:

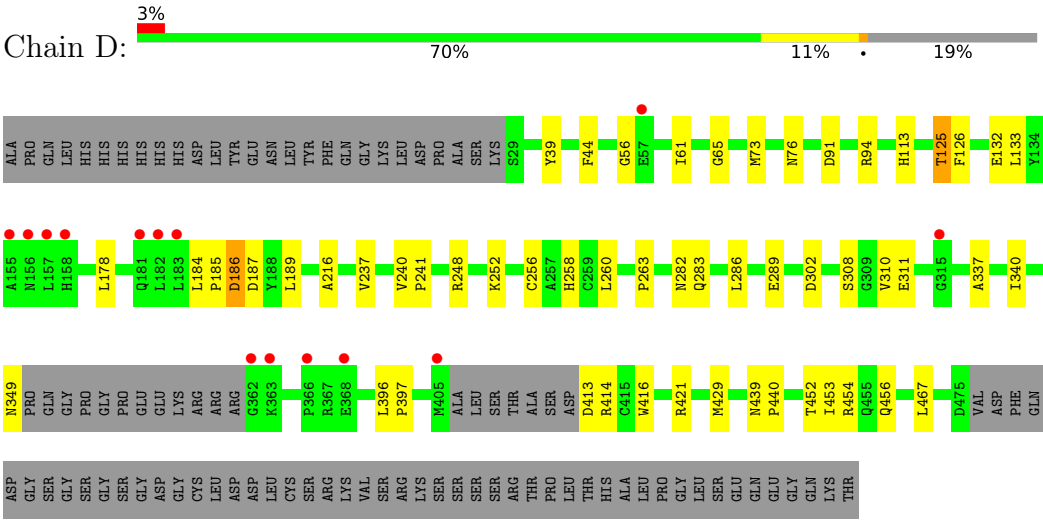




• Molecule 1: GLYPICAN-1



• Molecule 1: GLYPICAN-1



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	47.25Å 169.20Å 151.71Å 90.00° 94.96° 90.00°	Depositor
Resolution (Å)	29.76 – 2.94 29.76 – 2.95	Depositor EDS
% Data completeness (in resolution range)	99.0 (29.76-2.94) 99.2 (29.76-2.95)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.15 (at 2.95Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: DEV_938)	Depositor
R, R_{free}	0.218 , 0.278 0.216 , 0.276	Depositor DCC
R_{free} test set	2533 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	52.5	Xtriage
Anisotropy	0.945	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 67.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	12754	wwPDB-VP
Average B, all atoms (Å^2)	74.0	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 39.33 % 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 3.2232e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	0.23	0/2859	0.41	0/3860
1	B	0.25	0/3402	0.43	0/4601
1	C	0.24	0/3088	0.41	0/4171
1	D	0.24	0/3418	0.41	0/4623
All	All	0.24	0/12767	0.42	0/17255

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	39	TYR	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	2811	0	2763	23	0
1	B	3341	0	3299	36	0
1	C	3035	0	2989	19	0
1	D	3357	0	3308	28	0
2	A	14	0	13	0	0
2	B	28	0	26	0	0
2	C	14	0	13	0	0
2	D	28	0	26	0	0
3	A	18	0	0	0	0
3	B	39	0	0	2	0
3	C	30	0	0	0	0
3	D	39	0	0	1	0
All	All	12754	0	12437	105	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 (105) 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:C:313:VAL:HA	1:C:316:SER:HB3	1.72	0.71
1:D:61:ILE:HG23	1:D:248:ARG:HG2	1.71	0.71
1:D:414:ARG:HH21	1:D:421:ARG:HD3	1.58	0.67
1:D:454:ARG:NH2	3:D:2018:HOH:O	2.29	0.65
1:B:133:LEU:O	1:B:137:ASN:ND2	2.30	0.64
1:C:133:LEU:O	1:C:137:ASN:ND2	2.26	0.64
1:A:212:ARG:NH2	1:A:307:THR:O	2.30	0.63
1:B:114:LEU:HD21	1:B:311:GLU:HG3	1.80	0.63
1:B:260:LEU:HD12	1:B:416:TRP:HH2	1.63	0.62
1:D:125:THR:HG22	1:D:126:PHE:HD1	1.64	0.62
1:D:133:LEU:O	1:D:137:ASN:ND2	2.34	0.61
1:A:36:ARG:HG3	1:A:49:VAL:HG11	1.82	0.60
1:A:260:LEU:HD12	1:A:416:TRP:HH2	1.65	0.60
1:B:176:LYS:HE3	1:B:189:LEU:HD11	1.84	0.59
1:C:137:ASN:OD1	1:C:173:ARG:NH1	2.36	0.59
1:A:216:ALA:HA	1:A:310:VAL:HG22	1.85	0.58
1:B:285:ASP:OD2	1:B:454:ARG:NH2	2.39	0.56
1:D:91:ASP:OD1	1:D:94:ARG:NH2	2.39	0.56
1:B:172:GLU:HG3	1:B:189:LEU:HD13	1.87	0.55
1:D:216:ALA:HB1	1:D:310:VAL:HA	1.88	0.55
1:B:37:GLN:O	1:B:40:GLY:HA3	2.07	0.55
1:B:163:LEU:HD12	1:B:210:ARG:HG3	1.89	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:335:LEU:O	1:C:336:THR:OG1	2.25	0.54
1:A:373:GLY:HA2	1:A:377:LYS:HE2	1.91	0.53
1:B:429:MET:HG3	1:B:439:ASN:HD22	1.74	0.52
1:A:316:SER:OG	1:A:319:THR:OG1	2.27	0.52
1:B:282:ASN:HB3	1:B:453:ILE:HD12	1.92	0.52
1:A:110:HIS:CE1	1:A:375:LEU:HB2	2.45	0.52
1:D:337:ALA:O	1:D:340:ILE:HG13	2.11	0.51
1:B:319:THR:HG21	1:B:367:ARG:HB3	1.91	0.51
1:D:282:ASN:HB3	1:D:453:ILE:HD12	1.92	0.50
1:B:474:ASN:ND2	3:B:2039:HOH:O	2.37	0.49
1:C:233:VAL:HG13	1:C:456:GLN:HB3	1.94	0.49
1:A:252:LYS:HA	1:A:256:CYS:SG	2.53	0.48
1:B:186:ASP:HA	1:B:189:LEU:HB2	1.95	0.48
1:C:281:ALA:HB1	1:C:433:LEU:HA	1.95	0.48
1:D:260:LEU:HD12	1:D:416:TRP:HH2	1.79	0.48
1:D:132:GLU:HA	1:D:135:THR:OG1	2.13	0.47
1:B:132:GLU:HA	1:B:135:THR:OG1	2.14	0.47
1:B:421:ARG:NH2	3:B:2018:HOH:O	2.46	0.47
1:A:54:ILE:HG22	1:A:55:SER:H	1.79	0.47
1:A:304:PHE:HA	1:A:310:VAL:HB	1.97	0.47
1:C:323:GLU:O	1:C:326:ASN:ND2	2.47	0.47
1:D:348:VAL:HG12	1:D:349:ASN:H	1.80	0.46
1:D:429:MET:HG3	1:D:439:ASN:HD22	1.79	0.46
1:C:159:LEU:O	1:C:162:THR:OG1	2.25	0.46
1:C:252:LYS:HA	1:C:256:CYS:SG	2.55	0.46
1:C:429:MET:HG3	1:C:439:ASN:HD22	1.81	0.46
1:D:186:ASP:HA	1:D:189:LEU:HB2	1.96	0.46
1:A:77:LEU:HD13	1:A:255:TYR:HB3	1.98	0.46
1:B:237:VAL:HG21	1:B:286:LEU:HD11	1.97	0.46
1:A:236:LYS:HD2	1:A:456:GLN:NE2	2.31	0.46
1:C:116:ASN:HB3	1:C:120:ARG:HH12	1.80	0.45
1:A:58:HIS:H	1:A:58:HIS:HD1	1.63	0.45
1:C:453:ILE:O	1:C:457:ILE:HG13	2.17	0.45
1:D:240:VAL:HA	1:D:241:PRO:HD3	1.78	0.45
1:B:216:ALA:HB1	1:B:310:VAL:HA	1.99	0.45
1:D:439:ASN:HA	1:D:440:PRO:HD3	1.76	0.45
1:B:61:ILE:HG23	1:B:248:ARG:HB2	1.99	0.44
1:D:452:THR:O	1:D:456:GLN:HG2	2.17	0.44
1:D:56:GLY:HA3	1:D:65:GLY:O	2.17	0.44
1:D:258:HIS:CD2	1:D:263:PRO:HB3	2.53	0.44
1:A:237:VAL:HG21	1:A:286:LEU:HD11	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:212:ARG:HB3	1:C:313:VAL:HG11	2.00	0.44
1:D:252:LYS:HA	1:D:256:CYS:SG	2.58	0.44
1:C:328:LEU:HD23	1:C:328:LEU:HA	1.84	0.43
1:C:291:ARG:NH2	1:D:302:ASP:OD2	2.30	0.43
1:A:163:LEU:HD12	1:A:210:ARG:HG3	2.01	0.43
1:A:318:HIS:ND1	1:A:319:THR:HG23	2.33	0.43
1:A:152:TYR:O	1:A:221:ARG:HG3	2.18	0.43
1:A:269:PRO:O	1:A:273:ARG:HB2	2.17	0.43
1:B:139:ARG:HG3	1:B:142:ARG:HH21	1.84	0.43
1:B:240:VAL:HA	1:B:241:PRO:HD3	1.83	0.43
1:A:110:HIS:HE1	1:A:375:LEU:HB2	1.82	0.43
1:A:240:VAL:HG21	1:A:453:ILE:HD11	2.01	0.43
1:B:316:SER:HB2	1:B:320:TRP:CZ3	2.53	0.43
1:B:139:ARG:HG3	1:B:142:ARG:NH2	2.34	0.42
1:B:196:ALA:O	1:B:201:PRO:HD3	2.18	0.42
1:C:260:LEU:HD12	1:C:416:TRP:HH2	1.83	0.42
1:D:44:PHE:CZ	1:D:76:ASN:HB3	2.55	0.42
1:B:281:ALA:HB1	1:B:433:LEU:HA	2.01	0.42
1:B:252:LYS:HA	1:B:256:CYS:SG	2.59	0.42
1:B:452:THR:O	1:B:456:GLN:HG2	2.20	0.42
1:B:337:ALA:O	1:B:340:ILE:HG13	2.20	0.42
1:B:338:LYS:HB3	1:B:338:LYS:HE2	1.84	0.42
1:D:283:GLN:OE1	1:D:283:GLN:N	2.47	0.42
1:B:49:VAL:HA	1:B:50:PRO:HD3	1.79	0.42
1:D:39:TYR:CZ	1:D:73:MET:HG2	2.55	0.42
1:D:178:LEU:HD23	1:D:178:LEU:HA	1.92	0.41
1:B:126:PHE:HB3	1:B:134:TYR:CG	2.55	0.41
1:A:447:THR:O	1:A:449:PRO:HD3	2.20	0.41
1:D:237:VAL:HG21	1:D:286:LEU:HD11	2.02	0.41
1:B:218:VAL:HG12	1:B:221:ARG:HH12	1.85	0.41
1:B:390:GLN:HE21	1:B:390:GLN:HB3	1.66	0.41
1:D:396:LEU:HB2	1:D:397:PRO:HD3	2.03	0.41
1:B:300:ILE:HD12	1:B:300:ILE:HA	1.93	0.41
1:C:199:LEU:HD12	1:C:199:LEU:HA	1.86	0.41
1:B:170:LEU:HG	1:B:174:LEU:HD13	2.03	0.41
1:C:245:GLU:CD	1:C:245:GLU:H	2.24	0.41
1:D:184:LEU:HA	1:D:185:PRO:HD2	1.79	0.41
1:A:387:ARG:O	1:A:390:GLN:HG2	2.20	0.41
1:B:183:LEU:HB3	1:B:347:LYS:HG3	2.03	0.40
1:B:450:ASP:O	1:B:454:ARG:HG2	2.22	0.40
1:C:152:TYR:O	1:C:221:ARG:HG3	2.21	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:116:ASN:OD1	1:A:149:ARG:NH1	2.54	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	350/528 (66%)	339 (97%)	10 (3%)	1 (0%)	41 69
1	B	420/528 (80%)	410 (98%)	10 (2%)	0	100 100
1	C	375/528 (71%)	363 (97%)	12 (3%)	0	100 100
1	D	422/528 (80%)	412 (98%)	10 (2%)	0	100 100
All	All	1567/2112 (74%)	1524 (97%)	42 (3%)	1 (0%)	51 80

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	206	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	301/444 (68%)	297 (99%)	4 (1%)	69 88

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	358/444 (81%)	349 (98%)	9 (2%)	47 76
1	C	324/444 (73%)	316 (98%)	8 (2%)	47 76
1	D	360/444 (81%)	349 (97%)	11 (3%)	40 71
All	All	1343/1776 (76%)	1311 (98%)	32 (2%)	49 77

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	136	GLN
1	A	208	GLU
1	A	402	SER
1	A	442	VAL
1	B	58	HIS
1	B	114	LEU
1	B	174	LEU
1	B	186	ASP
1	B	187	ASP
1	B	369	ARG
1	B	380	SER
1	B	387	ARG
1	B	465	ASN
1	C	37	GLN
1	C	64	GLN
1	C	174	LEU
1	C	200	ARG
1	C	270	ASP
1	C	326	ASN
1	C	333	ASP
1	C	335	LEU
1	D	113	HIS
1	D	125	THR
1	D	186	ASP
1	D	187	ASP
1	D	289	GLU
1	D	308	SER
1	D	311	GLU
1	D	345	ASN
1	D	348	VAL
1	D	413	ASP
1	D	467	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	390	GLN
1	C	97	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)

6 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	1501	1	14,14,15	0.53	0	17,19,21	0.86	0
2	NAG	C	1501	1	14,14,15	0.61	0	17,19,21	1.32	4 (23%)
2	NAG	D	1501	1	14,14,15	0.51	0	17,19,21	0.86	1 (5%)
2	NAG	D	1500	1	14,14,15	0.65	0	17,19,21	1.06	1 (5%)
2	NAG	B	1501	1	14,14,15	0.50	0	17,19,21	0.89	1 (5%)
2	NAG	B	1500	1	14,14,15	0.65	0	17,19,21	0.75	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	1501	1	-	5/6/23/26	0/1/1/1
2	NAG	C	1501	1	-	2/6/23/26	0/1/1/1
2	NAG	D	1501	1	-	4/6/23/26	0/1/1/1
2	NAG	D	1500	1	-	4/6/23/26	0/1/1/1
2	NAG	B	1501	1	-	0/6/23/26	0/1/1/1
2	NAG	B	1500	1	1/1/5/7	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	1501	NAG	O5-C5-C6	2.80	111.60	107.20
2	C	1501	NAG	C1-O5-C5	2.64	115.77	112.19
2	D	1500	NAG	C4-C3-C2	2.42	114.56	111.02
2	C	1501	NAG	C4-C3-C2	-2.29	107.66	111.02
2	B	1501	NAG	C1-O5-C5	2.27	115.27	112.19
2	C	1501	NAG	O5-C1-C2	-2.08	108.00	111.29
2	D	1501	NAG	C2-N2-C7	2.00	125.75	122.90

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	1500	NAG	C1

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1501	NAG	C3-C2-N2-C7
2	A	1501	NAG	C8-C7-N2-C2
2	A	1501	NAG	O7-C7-N2-C2
2	D	1501	NAG	C3-C2-N2-C7
2	D	1501	NAG	O7-C7-N2-C2
2	D	1501	NAG	C8-C7-N2-C2
2	D	1500	NAG	C8-C7-N2-C2
2	D	1500	NAG	O7-C7-N2-C2
2	C	1501	NAG	O5-C5-C6-O6
2	B	1500	NAG	C8-C7-N2-C2
2	D	1500	NAG	O5-C5-C6-O6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	A	1501	NAG	C1-C2-N2-C7
2	B	1500	NAG	O7-C7-N2-C2
2	A	1501	NAG	O5-C5-C6-O6
2	C	1501	NAG	C8-C7-N2-C2
2	D	1501	NAG	C4-C5-C6-O6
2	D	1500	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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	358/528 (67%)	0.20	26 (7%)	15	13	33, 75, 142, 171	0
1	B	426/528 (80%)	-0.14	16 (3%)	40	39	24, 64, 110, 146	0
1	C	385/528 (72%)	0.08	17 (4%)	34	33	34, 75, 134, 174	0
1	D	428/528 (81%)	-0.06	16 (3%)	41	40	28, 66, 117, 150	0
All	All	1597/2112 (75%)	0.01	75 (4%)	31	31	24, 69, 127, 174	0

All (75) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	372	SER	5.6
1	C	474	ASN	5.3
1	A	205	ALA	5.0
1	A	127	PRO	4.9
1	D	366	PRO	4.4
1	A	207	ARG	4.0
1	A	48	ASP	4.0
1	C	45	SER	3.9
1	B	181	GLN	3.9
1	B	155	ALA	3.6
1	B	182	LEU	3.5
1	D	158	HIS	3.5
1	D	348	VAL	3.5
1	A	204	GLU	3.4
1	A	211	LEU	3.4
1	B	57	GLU	3.4
1	B	403	GLU	3.2
1	D	57	GLU	3.2
1	B	347	LYS	3.2
1	A	130	PHE	3.1
1	C	36	ARG	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	362	GLY	3.0
1	C	46	LEU	3.0
1	C	157	LEU	3.0
1	D	368	GLU	3.0
1	A	472	ASN	3.0
1	D	182	LEU	2.9
1	C	179	HIS	2.9
1	C	44	PHE	2.8
1	A	206	PRO	2.8
1	B	371	PRO	2.8
1	D	181	GLN	2.8
1	C	57	GLU	2.7
1	D	155	ALA	2.7
1	C	198	ALA	2.7
1	A	123	GLN	2.7
1	D	157	LEU	2.6
1	A	305	TRP	2.6
1	B	419	MET	2.6
1	D	405	MET	2.6
1	B	362	GLY	2.5
1	C	178	LEU	2.5
1	D	183	LEU	2.5
1	D	363	LYS	2.5
1	C	303	LYS	2.5
1	B	183	LEU	2.5
1	A	36	ARG	2.4
1	A	266	ARG	2.4
1	A	158	HIS	2.4
1	A	72	GLU	2.4
1	B	363	LYS	2.3
1	A	132	GLU	2.3
1	C	332	ARG	2.3
1	A	172	GLU	2.3
1	A	46	LEU	2.2
1	A	128	GLY	2.2
1	A	168	ALA	2.2
1	B	158	HIS	2.2
1	C	180	PRO	2.2
1	C	128	GLY	2.1
1	A	126	PHE	2.1
1	A	41	ALA	2.1
1	A	164	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	347	LYS	2.1
1	B	402	SER	2.1
1	D	315	GLY	2.1
1	D	156	ASN	2.1
1	A	161	GLU	2.1
1	C	175	PHE	2.1
1	C	64	GLN	2.0
1	A	171	LEU	2.0
1	B	369	ARG	2.0
1	A	414	ARG	2.0
1	B	414	ARG	2.0
1	B	475	ASP	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	D	1500	14/15	0.75	0.25	103,125,137,142	0
2	NAG	B	1500	14/15	0.84	0.20	77,104,117,121	0
2	NAG	D	1501	14/15	0.86	0.24	65,88,119,140	0
2	NAG	A	1501	14/15	0.89	0.15	89,99,102,104	0
2	NAG	C	1501	14/15	0.92	0.21	66,75,87,92	0
2	NAG	B	1501	14/15	0.93	0.22	77,87,102,117	0

6.5 Other polymers (i)

There are no such residues in this entry.