Full wwPDB X-ray Structure Validation Report (i)

Aug 6, 2020 – 06:52 PM BST

PDB ID	:	4ACR
Title	:	Crystal structure of N-glycosylated, C-terminally truncated human glypican-1
Authors	:	Svensson, G.; Awad, W.; Mani, K.; Logan, D.T.
Deposited on	:	2011-12-17
Resolution	:	2.55 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 following versions of software and data (see references (1)) were used in the production of this report:

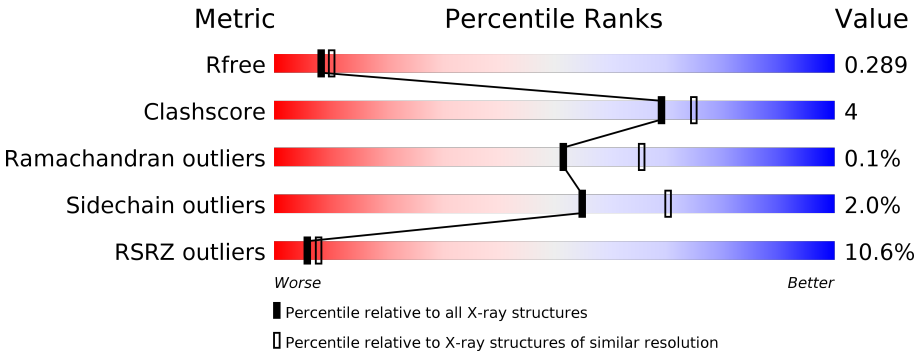
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.55 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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	478	9% 65% 7% 27%
1	B	478	8% 78% 10% 11%
1	C	478	10% 73% 9% 18%
1	D	478	8% 81% 8% • 10%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	B	500	X	-	-	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12812 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	347	Total	C	N	O	S	0	0	0
			2723	1696	500	506	21			
1	B	424	Total	C	N	O	S	0	0	0
			3324	2074	605	622	23			
1	C	393	Total	C	N	O	S	0	0	0
			3082	1921	562	577	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...

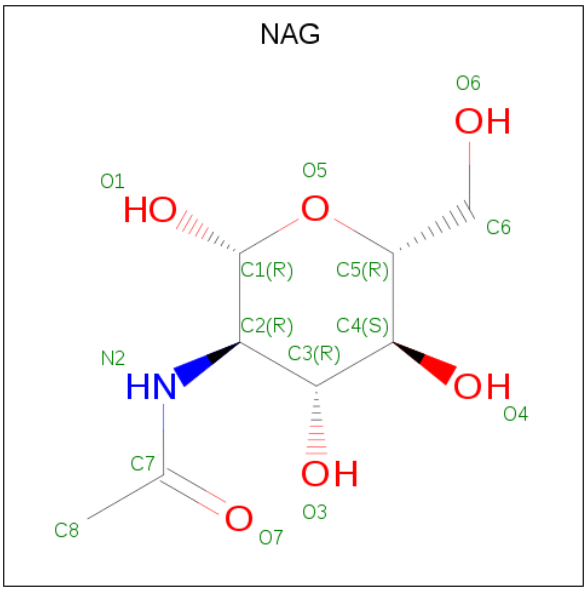
Chain	Residue	Modelled	Actual	Comment	Reference
A	23	LEU	-	expression tag	UNP P35052
B	2	ALA	-	expression tag	UNP P35052
B	3	PRO	-	expression tag	UNP P35052
B	4	GLN	-	expression tag	UNP P35052
B	5	LEU	-	expression tag	UNP P35052
B	6	HIS	-	expression tag	UNP P35052
B	7	HIS	-	expression tag	UNP P35052
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.

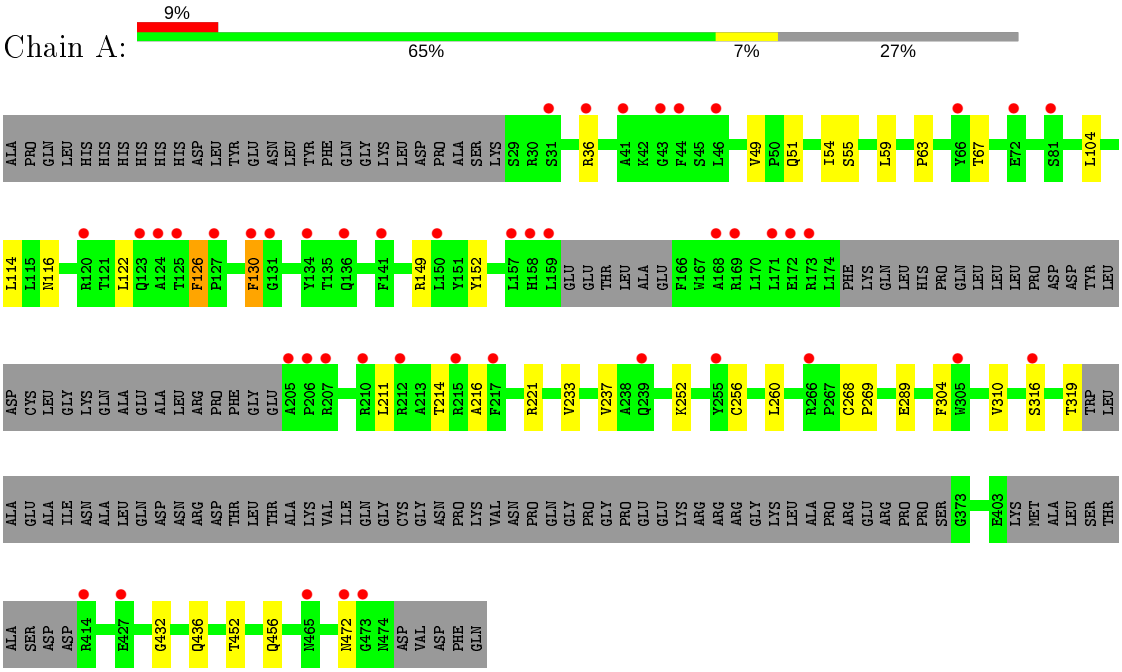
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	35	Total O 35 35	0	0
3	B	55	Total O 55 55	0	0
3	C	70	Total O 70 70	0	0
3	D	82	Total O 82 82	0	0

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.

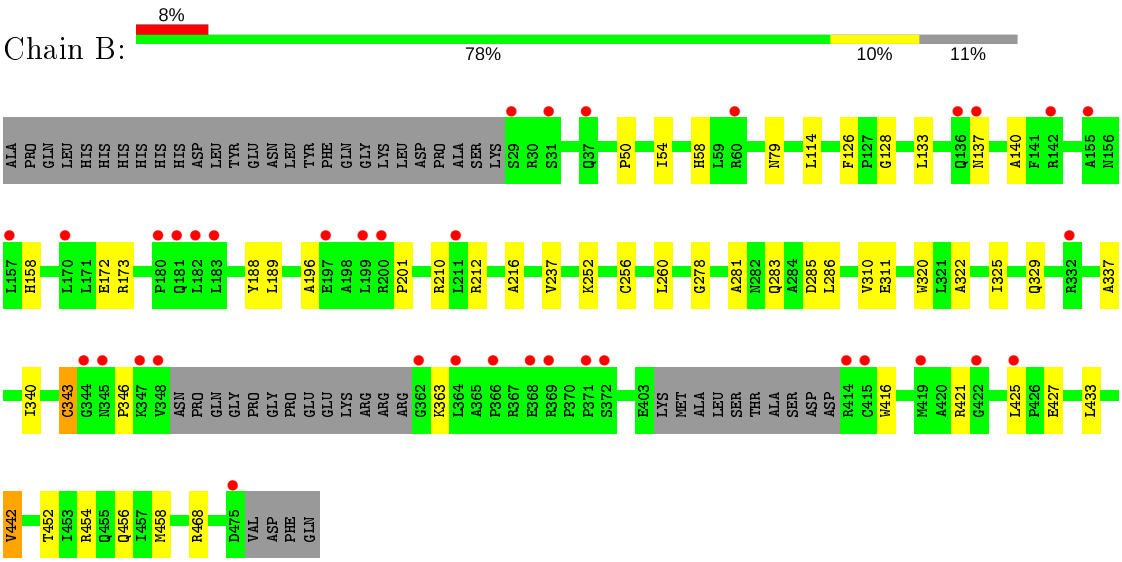
- Molecule 1: GLYPICAN-1

Chain A:



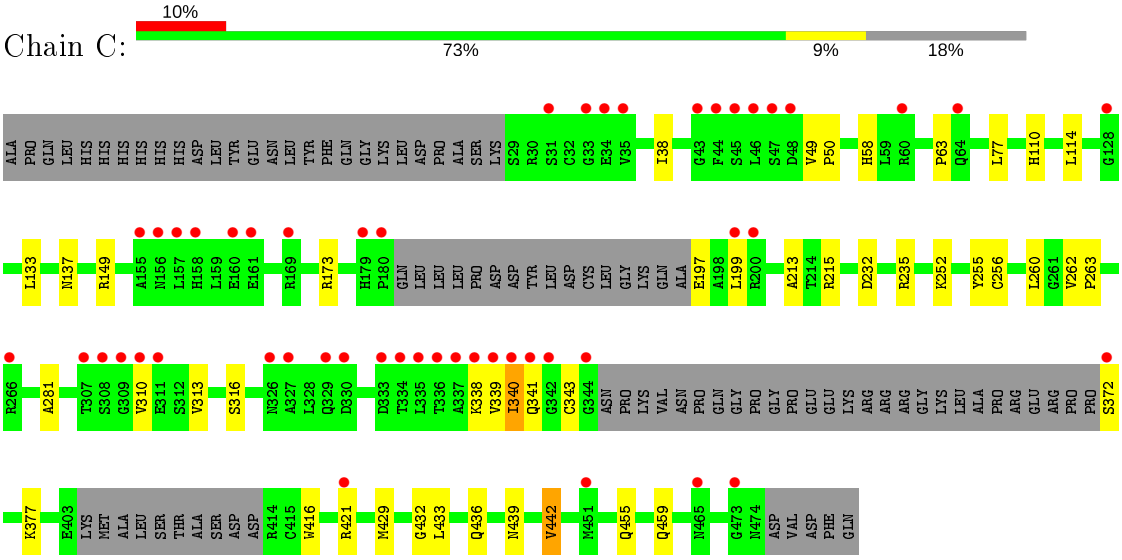
- Molecule 1: GLYPICAN-1

Chain B:



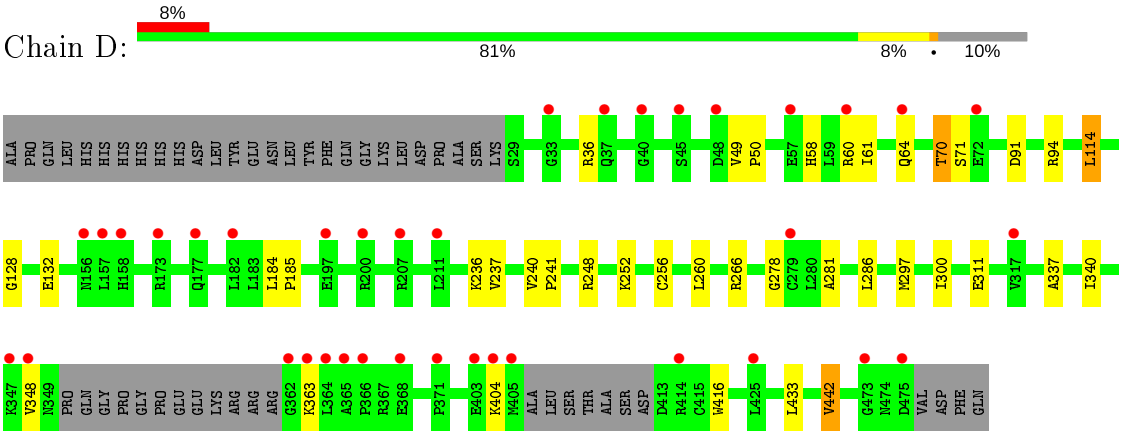
• Molecule 1: GLYPICAN-1

Chain C:



• Molecule 1: GLYPICAN-1

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	47.17Å 168.63Å 147.76Å 90.00° 94.59° 90.00°	Depositor
Resolution (Å)	29.71 – 2.55 29.71 – 2.55	Depositor EDS
% Data completeness (in resolution range)	99.7 (29.71-2.55) 99.8 (29.71-2.55)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.18 (at 2.54Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.251 , 0.292 0.250 , 0.289	Depositor DCC
R_{free} test set	3757 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	45.9	Xtriage
Anisotropy	0.825	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 53.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	12812	wwPDB-VP
Average B, all atoms (Å^2)	73.0	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 35.39 % 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 5.8764e-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	0.22	0/2769	0.39	0/3739
1	B	0.22	0/3385	0.40	0/4580
1	C	0.23	0/3136	0.42	0/4239
1	D	0.23	0/3418	0.41	0/4623
All	All	0.22	0/12708	0.40	0/17181

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	2723	0	2681	21	0
1	B	3324	0	3277	28	0
1	C	3082	0	3030	25	0
1	D	3357	0	3309	19	0
2	A	14	0	13	0	0
2	B	28	0	26	1	0
2	C	14	0	13	0	0
2	D	28	0	26	0	0
3	A	35	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	55	0	0	1	0
3	C	70	0	0	2	0
3	D	82	0	0	1	0
All	All	12812	0	12375	93	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 (93) 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:D:128:GLY:HA3	1:D:363:LYS:HD3	1.69	0.74
1:C:63:PRO:HG2	1:C:260:LEU:HD11	1.71	0.71
1:C:137:ASN:OD1	1:C:173:ARG:NH1	2.27	0.67
1:A:116:ASN:OD1	1:A:149:ARG:NH1	2.28	0.66
1:D:114:LEU:HD21	1:D:311:GLU:HG3	1.77	0.65
1:B:260:LEU:HD12	1:B:416:TRP:HH2	1.62	0.64
1:B:133:LEU:O	1:B:137:ASN:ND2	2.33	0.62
1:A:216:ALA:HA	1:A:310:VAL:HG22	1.84	0.59
1:A:122:LEU:O	1:A:126:PHE:HB2	2.05	0.56
1:B:340:ILE:HG22	1:B:346:PRO:HD2	1.87	0.56
1:D:91:ASP:OD1	1:D:94:ARG:NH2	2.39	0.55
1:C:260:LEU:HD12	1:C:416:TRP:HH2	1.70	0.55
1:C:313:VAL:HA	1:C:316:SER:HB2	1.88	0.55
1:C:372:SER:OG	1:C:377:LYS:NZ	2.40	0.55
1:C:262:VAL:HG11	1:C:421:ARG:HE	1.73	0.54
1:D:260:LEU:HD12	1:D:416:TRP:HH2	1.72	0.54
1:B:421:ARG:NH2	3:B:2021:HOH:O	2.42	0.52
1:C:232:ASP:OD1	1:C:235:ARG:NH1	2.43	0.51
1:D:61:ILE:HG23	1:D:248:ARG:HG2	1.90	0.51
1:A:316:SER:OG	1:A:319:THR:OG1	2.23	0.51
1:C:262:VAL:HG13	1:C:421:ARG:HH21	1.75	0.51
1:A:211:LEU:O	1:A:214:THR:OG1	2.28	0.50
1:C:215:ARG:NH1	3:C:2020:HOH:O	2.44	0.50
1:D:281:ALA:HB1	1:D:433:LEU:HA	1.94	0.50
1:D:236:LYS:NZ	3:D:2028:HOH:O	2.45	0.50
1:B:140:ALA:HB2	1:B:173:ARG:HH12	1.77	0.49
1:B:128:GLY:HA3	1:B:363:LYS:HD2	1.94	0.49
1:C:133:LEU:O	1:C:137:ASN:ND2	2.44	0.49
1:B:188:TYR:CE2	1:B:343:CYS:HB3	2.47	0.49
1:D:248:ARG:HH12	1:D:442:VAL:HA	1.78	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:266:ARG:HD2	1:D:404:LYS:HD2	1.94	0.49
1:B:188:TYR:CZ	1:B:343:CYS:HB3	2.48	0.48
1:C:149:ARG:NH1	3:C:2012:HOH:O	2.42	0.48
1:D:237:VAL:HG21	1:D:286:LEU:HD11	1.96	0.48
1:B:278:GLY:HA3	1:B:442:VAL:HG11	1.96	0.48
1:C:38:ILE:HD12	1:C:263:PRO:HB3	1.94	0.48
1:C:439:ASN:HD22	1:C:442:VAL:H	1.62	0.48
1:C:455:GLN:O	1:C:459:GLN:HG2	2.14	0.48
1:A:59:LEU:HD11	1:A:67:THR:HG21	1.97	0.47
1:B:281:ALA:HB1	1:B:433:LEU:HA	1.96	0.47
1:A:51:GLN:N	1:A:51:GLN:OE1	2.47	0.47
1:C:213:ALA:HA	1:C:313:VAL:HG21	1.96	0.47
1:C:439:ASN:ND2	1:C:442:VAL:H	2.14	0.46
1:C:281:ALA:HB1	1:C:433:LEU:HA	1.97	0.46
1:A:63:PRO:HG2	1:A:260:LEU:HD21	1.97	0.46
1:B:425:LEU:H	1:B:425:LEU:HD22	1.81	0.46
1:A:36:ARG:HG3	1:A:49:VAL:HG11	1.98	0.45
1:C:252:LYS:HA	1:C:256:CYS:SG	2.56	0.45
1:B:337:ALA:O	1:B:340:ILE:HG13	2.16	0.45
1:A:126:PHE:CE2	1:A:130:PHE:HD2	2.35	0.45
1:A:216:ALA:HB1	1:A:310:VAL:HG13	1.98	0.45
1:B:79:ASN:OD1	2:B:500:NAG:N2	2.50	0.45
1:D:337:ALA:O	1:D:340:ILE:HG13	2.17	0.45
1:A:252:LYS:HA	1:A:256:CYS:SG	2.57	0.44
1:A:304:PHE:HA	1:A:310:VAL:HB	2.00	0.44
1:A:54:ILE:HG22	1:A:55:SER:H	1.82	0.44
1:B:158:HIS:O	1:B:158:HIS:ND1	2.50	0.44
1:B:283:GLN:N	1:B:283:GLN:OE1	2.49	0.44
1:B:126:PHE:HE1	1:B:322:ALA:HB2	1.82	0.44
1:D:36:ARG:HG3	1:D:49:VAL:HG11	1.99	0.44
1:C:199:LEU:HD23	1:C:199:LEU:HA	1.72	0.44
1:B:212:ARG:HH11	1:B:320:TRP:HZ2	1.65	0.44
1:B:216:ALA:HB1	1:B:310:VAL:HA	2.00	0.44
1:B:172:GLU:HG2	1:B:189:LEU:HD13	2.00	0.43
1:B:114:LEU:HD13	1:B:311:GLU:HA	1.99	0.43
1:C:77:LEU:HD13	1:C:255:TYR:HB3	2.00	0.43
1:D:240:VAL:HA	1:D:241:PRO:HD3	1.85	0.43
1:A:233:VAL:HG13	1:A:456:GLN:HB3	2.01	0.43
1:D:297:MET:O	1:D:300:ILE:HG22	2.19	0.43
1:B:252:LYS:HA	1:B:256:CYS:SG	2.58	0.43
1:B:285:ASP:OD2	1:B:454:ARG:NH2	2.52	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:432:GLY:O	1:C:436:GLN:HG2	2.19	0.43
1:B:325:ILE:O	1:B:329:GLN:HG2	2.19	0.42
1:C:49:VAL:HA	1:C:50:PRO:HD3	1.87	0.42
1:D:252:LYS:HA	1:D:256:CYS:SG	2.58	0.42
1:A:432:GLY:O	1:A:436:GLN:NE2	2.49	0.42
1:C:340:ILE:HG23	1:C:341:GLN:H	1.85	0.42
1:B:50:PRO:HB3	1:B:54:ILE:HD11	2.02	0.42
1:A:152:TYR:O	1:A:221:ARG:HG3	2.19	0.42
1:B:196:ALA:O	1:B:201:PRO:HD3	2.20	0.41
1:D:184:LEU:HA	1:D:185:PRO:HD2	1.81	0.41
1:A:237:VAL:HG23	1:A:456:GLN:HG3	2.02	0.41
1:B:237:VAL:HG21	1:B:286:LEU:HD11	2.02	0.41
1:D:278:GLY:HA3	1:D:442:VAL:HG11	2.01	0.41
1:A:316:SER:O	1:A:319:THR:OG1	2.39	0.41
1:B:188:TYR:CE2	1:B:346:PRO:HD3	2.56	0.41
1:A:452:THR:O	1:A:456:GLN:HG2	2.21	0.41
1:B:452:THR:O	1:B:456:GLN:HG2	2.21	0.41
1:A:268:CYS:HA	1:A:269:PRO:HD3	1.92	0.41
1:D:49:VAL:HA	1:D:50:PRO:HD3	1.78	0.40
1:C:429:MET:HG3	1:C:439:ASN:OD1	2.22	0.40
1:D:70:THR:HG22	1:D:71:SER:H	1.85	0.40
1:C:110:HIS:O	1:C:114:LEU:HB2	2.22	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	337/478 (70%)	332 (98%)	5 (2%)	0	100	100
1	B	418/478 (87%)	409 (98%)	9 (2%)	0	100	100
1	C	385/478 (80%)	377 (98%)	6 (2%)	2 (0%)	29	40

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	422/478 (88%)	410 (97%)	12 (3%)	0	100 100
All	All	1562/1912 (82%)	1528 (98%)	32 (2%)	2 (0%)	51 65

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	340	ILE
1	C	339	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	292/403 (72%)	286 (98%)	6 (2%)	53 68
1	B	356/403 (88%)	349 (98%)	7 (2%)	55 70
1	C	329/403 (82%)	323 (98%)	6 (2%)	59 74
1	D	360/403 (89%)	352 (98%)	8 (2%)	52 66
All	All	1337/1612 (83%)	1310 (98%)	27 (2%)	55 70

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	104	LEU
1	A	114	LEU
1	A	126	PHE
1	A	130	PHE
1	A	289	GLU
1	A	472	ASN
1	B	58	HIS
1	B	210	ARG
1	B	343	CYS
1	B	427	GLU
1	B	442	VAL
1	B	458	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	468	ARG
1	C	58	HIS
1	C	197	GLU
1	C	310	VAL
1	C	338	LYS
1	C	343	CYS
1	C	442	VAL
1	D	58	HIS
1	D	60	ARG
1	D	64	GLN
1	D	70	THR
1	D	114	LEU
1	D	132	GLU
1	D	348	VAL
1	D	442	VAL

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

Mol	Chain	Res	Type
1	B	58	HIS

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	B	500	1	14,14,15	0.58	0	17,19,21	0.87	1 (5%)
2	NAG	D	500	1	14,14,15	0.58	0	17,19,21	0.67	0
2	NAG	A	501	1	14,14,15	0.45	0	17,19,21	1.39	2 (11%)
2	NAG	C	501	1	14,14,15	0.49	0	17,19,21	0.85	1 (5%)
2	NAG	D	501	1	14,14,15	0.50	0	17,19,21	0.85	1 (5%)
2	NAG	B	501	1	14,14,15	0.50	0	17,19,21	0.86	1 (5%)

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	B	500	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	D	500	1	-	2/6/23/26	0/1/1/1
2	NAG	A	501	1	-	4/6/23/26	0/1/1/1
2	NAG	C	501	1	-	1/6/23/26	0/1/1/1
2	NAG	D	501	1	-	2/6/23/26	0/1/1/1
2	NAG	B	501	1	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	501	NAG	C1-O5-C5	4.06	117.69	112.19
2	C	501	NAG	C1-O5-C5	2.32	115.33	112.19
2	A	501	NAG	C3-C4-C5	2.25	114.26	110.24
2	B	500	NAG	O5-C1-C2	-2.22	107.78	111.29
2	B	501	NAG	C1-O5-C5	2.20	115.17	112.19
2	D	501	NAG	C1-O5-C5	2.05	114.97	112.19

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	500	NAG	C1

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	500	NAG	C8-C7-N2-C2
2	D	500	NAG	O7-C7-N2-C2
2	A	501	NAG	C3-C2-N2-C7
2	A	501	NAG	C8-C7-N2-C2
2	A	501	NAG	O7-C7-N2-C2
2	B	501	NAG	C8-C7-N2-C2
2	B	501	NAG	O7-C7-N2-C2
2	D	501	NAG	C8-C7-N2-C2
2	D	501	NAG	O7-C7-N2-C2
2	C	501	NAG	O5-C5-C6-O6
2	A	501	NAG	O5-C5-C6-O6
2	B	501	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	500	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	347/478 (72%)	0.74	45 (12%) 3 4	41, 81, 136, 152	0
1	B	424/478 (88%)	0.53	36 (8%) 10 12	27, 67, 125, 159	0
1	C	393/478 (82%)	0.66	50 (12%) 3 4	28, 67, 131, 154	0
1	D	428/478 (89%)	0.38	37 (8%) 10 12	26, 60, 114, 146	0
All	All	1592/1912 (83%)	0.57	168 (10%) 6 8	26, 68, 128, 159	0

All (168) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	124	ALA	7.1
1	B	157	LEU	5.7
1	A	125	THR	5.6
1	C	179	HIS	5.5
1	C	337	ALA	5.5
1	B	197	GLU	5.4
1	C	199	LEU	5.3
1	B	348	VAL	5.3
1	C	336	THR	5.2
1	A	158	HIS	5.2
1	A	472	ASN	5.1
1	B	183	LEU	5.1
1	B	136	GLN	5.0
1	C	311	GLU	5.0
1	C	341	GLN	4.9
1	B	181	GLN	4.9
1	A	305	TRP	4.6
1	A	168	ALA	4.6
1	D	157	LEU	4.5
1	A	130	PHE	4.5
1	C	180	PRO	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	169	ARG	4.4
1	C	307	THR	4.4
1	C	326	ASN	4.3
1	C	329	GLN	4.2
1	B	414	ARG	4.1
1	D	347	LYS	4.1
1	D	405	MET	4.1
1	B	29	SER	4.0
1	A	473	GLY	3.9
1	A	171	LEU	3.9
1	D	366	PRO	3.9
1	B	37	GLN	3.9
1	A	266	ARG	3.8
1	B	182	LEU	3.8
1	C	310	VAL	3.7
1	C	342	GLY	3.6
1	A	465	ASN	3.6
1	A	207	ARG	3.6
1	A	206	PRO	3.5
1	A	173	ARG	3.5
1	A	43	GLY	3.5
1	C	308	SER	3.4
1	C	33	GLY	3.4
1	A	131	GLY	3.4
1	A	120	ARG	3.4
1	D	348	VAL	3.3
1	C	161	GLU	3.3
1	C	45	SER	3.3
1	D	363	LYS	3.3
1	C	200	ARG	3.2
1	B	347	LYS	3.2
1	A	205	ALA	3.2
1	C	421	ARG	3.2
1	C	333	ASP	3.1
1	B	366	PRO	3.1
1	C	43	GLY	3.1
1	D	57	GLU	3.1
1	A	150	LEU	3.1
1	B	419	MET	3.1
1	C	44	PHE	3.1
1	D	64	GLN	3.1
1	C	335	LEU	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	339	VAL	3.1
1	B	142	ARG	3.0
1	A	123	GLN	3.0
1	C	158	HIS	3.0
1	A	31	SER	3.0
1	C	31	SER	3.0
1	B	371	PRO	2.9
1	C	46	LEU	2.9
1	B	372	SER	2.9
1	C	266	ARG	2.9
1	B	200	ARG	2.9
1	C	60	ARG	2.9
1	A	134	TYR	2.9
1	D	197	GLU	2.8
1	D	414	ARG	2.8
1	D	211	LEU	2.8
1	D	33	GLY	2.8
1	C	330	ASP	2.8
1	B	31	SER	2.8
1	A	172	GLU	2.8
1	D	364	LEU	2.8
1	C	64	GLN	2.8
1	C	473	GLY	2.8
1	C	309	GLY	2.7
1	C	344	GLY	2.7
1	D	317	VAL	2.7
1	D	156	ASN	2.7
1	A	46	LEU	2.6
1	A	210	ARG	2.6
1	D	473	GLY	2.6
1	C	128	GLY	2.6
1	B	170	LEU	2.6
1	C	169	ARG	2.6
1	B	332	ARG	2.6
1	D	362	GLY	2.6
1	A	141	PHE	2.5
1	C	451	MET	2.5
1	D	60	ARG	2.5
1	A	41	ALA	2.5
1	B	211	LEU	2.5
1	A	136	GLN	2.5
1	B	369	ARG	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	160	GLU	2.5
1	D	45	SER	2.5
1	C	156	ASN	2.4
1	A	127	PRO	2.4
1	D	48	ASP	2.4
1	B	344	GLY	2.4
1	D	279	CYS	2.4
1	C	372	SER	2.4
1	B	422	GLY	2.4
1	B	415	CYS	2.4
1	B	199	LEU	2.4
1	C	47	SER	2.4
1	D	40	GLY	2.4
1	D	158	HIS	2.4
1	A	157	LEU	2.4
1	D	368	GLU	2.3
1	B	60	ARG	2.3
1	B	475	ASP	2.3
1	C	327	ALA	2.3
1	C	338	LYS	2.3
1	B	137	ASN	2.3
1	C	155	ALA	2.3
1	C	334	THR	2.3
1	C	48	ASP	2.3
1	D	72	GLU	2.3
1	A	215	ARG	2.2
1	C	465	ASN	2.2
1	B	362	GLY	2.2
1	A	36	ARG	2.2
1	A	427	GLU	2.2
1	A	217	PHE	2.2
1	D	182	LEU	2.2
1	A	81	SER	2.2
1	A	414	ARG	2.2
1	B	180	PRO	2.2
1	C	157	LEU	2.2
1	A	212	ARG	2.2
1	A	44	PHE	2.2
1	D	207	ARG	2.2
1	C	340	ILE	2.1
1	D	37	GLN	2.1
1	A	159	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	403	GLU	2.1
1	C	35	VAL	2.1
1	A	72	GLU	2.1
1	B	364	LEU	2.1
1	B	155	ALA	2.1
1	D	404	LYS	2.1
1	A	316	SER	2.1
1	C	34	GLU	2.1
1	D	200	ARG	2.1
1	D	475	ASP	2.0
1	A	255	TYR	2.0
1	A	239	GLN	2.0
1	B	345	ASN	2.0
1	A	66	TYR	2.0
1	D	425	LEU	2.0
1	B	368	GLU	2.0
1	D	371	PRO	2.0
1	D	177	GLN	2.0
1	D	365	ALA	2.0
1	D	173	ARG	2.0
1	B	425	LEU	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	A	501	14/15	0.77	0.18	107,116,118,118	0
2	NAG	B	500	14/15	0.81	0.19	80,92,97,99	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	D	500	14/15	0.82	0.23	115,118,120,120	0
2	NAG	B	501	14/15	0.87	0.23	88,98,103,107	0
2	NAG	D	501	14/15	0.88	0.23	75,85,96,102	0
2	NAG	C	501	14/15	0.91	0.17	60,66,77,85	0

6.5 Other polymers

There are no such residues in this entry.