Full wwPDB X-ray Structure Validation Report (i)

Sep 24, 2023 – 07:22 AM EDT

PDB ID	:	5UN8
Title	:	Crystal Structure of human O-GlcNAcase in complex with glycopeptide p53
Authors	:	Li, B.; Jiang, J.
Deposited on	:	2017-01-30
Resolution	:	2.13 Å(reported)

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

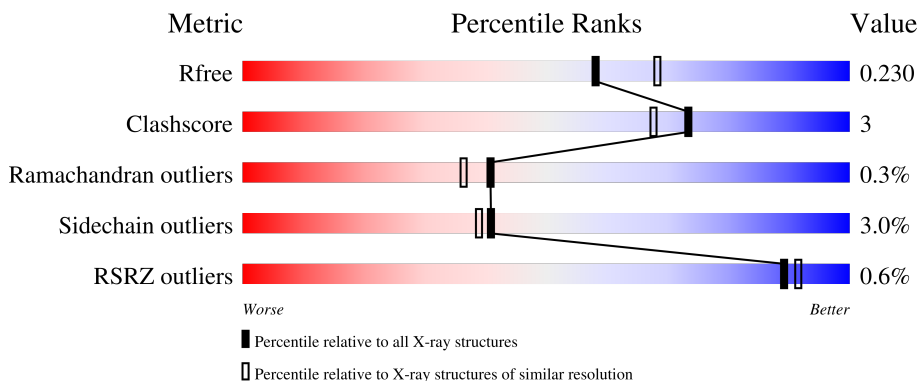|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.13 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	Percentile Ranks	Value
Rfree		0.230
Clashscore		3
Ramachandran outliers		0.3%
Sidechain outliers		3.0%
RSRZ outliers		0.6%

Worse ▬▬▬▬ *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2523 (2.16-2.12)
Clashscore	141614	2653 (2.16-2.12)
Ramachandran outliers	138981	2618 (2.16-2.12)
Sidechain outliers	138945	2617 (2.16-2.12)
RSRZ outliers	127900	2485 (2.16-2.12)

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	504	76% · 7% · 16%
1	B	504	76% · 8% · 15%
1	C	504	77% · 8% · 15%
1	D	504	74% · 11% · 15%
2	E	11	55% · 9% · 36%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	11	73% 27%
2	G	11	82% 18%
2	H	11	73% 27%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 15552 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 Protein O-GlcNAcase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	424	Total	C	N	O	S	0	0	0
			3480	2254	575	628	23			
1	B	426	Total	C	N	O	S	0	0	0
			3495	2263	578	631	23			
1	C	429	Total	C	N	O	S	0	0	0
			3520	2279	581	636	24			
1	D	429	Total	C	N	O	S	0	0	0
			3520	2279	581	636	24			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	59	HIS	-	expression tag	UNP O60502
A	175	ASN	ASP	conflict	UNP O60502
A	543	GLY	-	linker	UNP O60502
A	544	GLY	-	linker	UNP O60502
A	545	GLY	-	linker	UNP O60502
A	546	GLY	-	linker	UNP O60502
A	547	SER	-	linker	UNP O60502
A	548	GLY	-	linker	UNP O60502
A	549	GLY	-	linker	UNP O60502
A	550	GLY	-	linker	UNP O60502
A	551	GLY	-	linker	UNP O60502
A	552	SER	-	linker	UNP O60502
B	59	HIS	-	expression tag	UNP O60502
B	175	ASN	ASP	conflict	UNP O60502
B	543	GLY	-	linker	UNP O60502
B	544	GLY	-	linker	UNP O60502
B	545	GLY	-	linker	UNP O60502
B	546	GLY	-	linker	UNP O60502
B	547	SER	-	linker	UNP O60502
B	548	GLY	-	linker	UNP O60502
B	549	GLY	-	linker	UNP O60502

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	550	GLY	-	linker	UNP O60502
B	551	GLY	-	linker	UNP O60502
B	552	SER	-	linker	UNP O60502
C	59	HIS	-	expression tag	UNP O60502
C	175	ASN	ASP	conflict	UNP O60502
C	543	GLY	-	linker	UNP O60502
C	544	GLY	-	linker	UNP O60502
C	545	GLY	-	linker	UNP O60502
C	546	GLY	-	linker	UNP O60502
C	547	SER	-	linker	UNP O60502
C	548	GLY	-	linker	UNP O60502
C	549	GLY	-	linker	UNP O60502
C	550	GLY	-	linker	UNP O60502
C	551	GLY	-	linker	UNP O60502
C	552	SER	-	linker	UNP O60502
D	59	HIS	-	expression tag	UNP O60502
D	175	ASN	ASP	conflict	UNP O60502
D	543	GLY	-	linker	UNP O60502
D	544	GLY	-	linker	UNP O60502
D	545	GLY	-	linker	UNP O60502
D	546	GLY	-	linker	UNP O60502
D	547	SER	-	linker	UNP O60502
D	548	GLY	-	linker	UNP O60502
D	549	GLY	-	linker	UNP O60502
D	550	GLY	-	linker	UNP O60502
D	551	GLY	-	linker	UNP O60502
D	552	SER	-	linker	UNP O60502

- Molecule 2 is a protein called P53 peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	7	Total 56	C 37	N 8	O 11	0	0	0
2	F	8	Total 63	C 42	N 9	O 12	0	0	0
2	G	9	Total 67	C 44	N 10	O 13	0	0	0
2	H	8	Total 63	C 42	N 9	O 12	0	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

NAG

O6
OH

O1   O5
HO   O
                              C6
       C1(R)   C5(R)
       C2(R)   C4(S)
N2   C3(R)
HN   OH
                      O4
C7
              OH
C8   O   O3
              O7

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	317	Total 317	O 317	0	0
4	B	313	Total 313	O 313	0	0
4	C	277	Total 277	O 277	0	0
4	D	289	Total 289	O 289	0	0
4	E	13	Total 13	O 13	0	0
4	F	8	Total 8	O 8	0	0
4	G	6	Total 6	O 6	0	0
4	H	9	Total 9	O 9	0	0

W O R L D W I D E
PDB
PROTEIN DATA BANK

3 Residue-property plots ⓘ

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: Protein O-GlcNAcase

Chain A:

• Molecule 1: Protein O-GlcNAcase

Chain B:

• Molecule 1: Protein O-GlcNAcase

Chain C:

LYS ASP SER GLU LYS I604 E605 M622 G623 M624 F625 T626 S629 M640 K648 S649 I650 S656 C663 H666 L672 R691 P694 ILE ASP GLY ALA ASN ASP LEU PHE PHE GLN

• Molecule 1: Protein O-GlcNAcase

%

Chain D:　　　74%　　　11%　　•　　15%

H59 M75 E76 K79 R84 P97 D100 Y101 R104 W107 L119 I123 P139 V151 S162 C166 R167 S168 F169 C181 I198 C220 Y225 P226 S231 P232 Y233 L234 R235 T236 V237 T250 V261 E262 K269 K272 W278

D287 Q288 K289 R290 Y296 R299 R306 T312 N313 P314 Y322 M336 ASN GLY VAL ARG LYS ASP VAL VAL MET THR ASP SER SER GLU ASP SER THR VAL SER ILE GLN ILE LYS LEU GLU ASN GLU GLY SER SER ASP ASP ILE GLU THR ASP VAL L373 Q377 L382

E386 F391 GLY VAL PRO HIS GLN TYR SER SER ARG GLY GLY GLY SER GLY GLY GLY S552 V592 SER VAL ASN CYS LYS LYS GLY LYS ASP SER GLU LYS I604 S609 M622 T626 S629 N630 M640 K648 S656 A669 I673 G683 R691

L692 L693 P694 ILE ASP GLY ALA ASN ASP LEU PHE PHE GLN

• Molecule 2: P53 peptide

Chain E:　　　55%　　　9%　　　36%

GLN LEU W146 P152 PRO GLY

• Molecule 2: P53 peptide

Chain F:　　　73%　　　27%

GLN LEU W146 P153 GLY

• Molecule 2: P53 peptide

Chain G:　　　82%　　　18%

GLN LEU W146 G154

• Molecule 2: P53 peptide

Chain H:　　　73%　　　27%

GLN LEU W146 P153 GLY

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	89.91Å 95.39Å 149.32Å 90.00° 96.91° 90.00°	Depositor
Resolution (Å)	148.24 – 2.13 49.41 – 2.13	Depositor EDS
% Data completeness (in resolution range)	96.8 (148.24-2.13) 96.8 (49.41-2.13)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.85 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.184 , 0.229 0.185 , 0.230	Depositor DCC
R_{free} test set	6796 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	24.3	Xtriage
Anisotropy	0.053	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 58.7	EDS
L-test for twinning[2]	$< |L| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	15552	wwPDB-VP
Average B, all atoms (Å^2)	30.0	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 39.09 % 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 3.3407e-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	1.05	4/3574 (0.1%)	0.98	9/4838 (0.2%)
1	B	1.01	0/3589	0.92	7/4857 (0.1%)
1	C	1.00	1/3615 (0.0%)	0.91	7/4893 (0.1%)
1	D	0.98	1/3615 (0.0%)	0.96	10/4893 (0.2%)
2	E	1.23	0/59	0.97	0/83
2	F	1.15	0/67	1.02	0/95
2	G	1.25	0/71	0.84	0/100
2	H	1.22	0/67	0.92	0/95
All	All	1.01	6/14657 (0.0%)	0.94	33/19854 (0.2%)

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	663	CYS	CB-SG	12.03	2.02	1.82
1	A	663	CYS	CA-CB	6.01	1.67	1.53
1	A	663	CYS	N-CA	5.82	1.57	1.46
1	D	629	SER	CB-OG	-5.10	1.35	1.42
1	C	663	CYS	CA-CB	5.10	1.65	1.53
1	A	553	VAL	N-CA	5.09	1.56	1.46

All (33) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(º)	Ideal(º)
1	A	663	CYS	CA-CB-SG	19.58	149.24	114.00
1	D	691	ARG	NE-CZ-NH2	-13.86	113.37	120.30
1	B	83	ARG	NE-CZ-NH1	7.62	124.11	120.30
1	D	640	MET	CG-SD-CE	-7.46	88.26	100.20
1	B	83	ARG	NE-CZ-NH2	-6.98	116.81	120.30
1	C	100	ASP	CB-CG-OD2	6.89	124.50	118.30
1	D	691	ARG	NE-CZ-NH1	6.87	123.74	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	640	MET	CG-SD-CE	-6.28	90.15	100.20
1	A	83	ARG	NE-CZ-NH1	6.26	123.43	120.30
1	A	663	CYS	N-CA-C	-6.22	94.19	111.00
1	D	104	ARG	NE-CZ-NH2	-6.09	117.25	120.30
1	B	127	ARG	NE-CZ-NH1	6.08	123.34	120.30
1	C	624	MET	CG-SD-CE	-5.97	90.64	100.20
1	C	691	ARG	NE-CZ-NH1	5.97	123.29	120.30
1	A	83	ARG	NE-CZ-NH2	-5.92	117.34	120.30
1	C	650	ILE	CG1-CB-CG2	-5.91	98.39	111.40
1	C	663	CYS	CB-CA-C	5.89	122.19	110.40
1	D	287	ASP	CB-CG-OD1	5.82	123.53	118.30
1	D	691	ARG	CG-CD-NE	-5.75	99.72	111.80
1	C	691	ARG	NE-CZ-NH2	-5.69	117.46	120.30
1	A	290	ARG	NE-CZ-NH2	-5.62	117.49	120.30
1	D	84	ARG	NE-CZ-NH2	-5.58	117.51	120.30
1	A	167	ARG	NE-CZ-NH2	5.54	123.07	120.30
1	B	84	ARG	NE-CZ-NH2	-5.52	117.54	120.30
1	D	167	ARG	NE-CZ-NH2	5.42	123.01	120.30
1	A	78	ARG	NE-CZ-NH1	-5.30	117.65	120.30
1	A	663	CYS	N-CA-CB	5.18	119.92	110.60
1	A	156	ARG	NE-CZ-NH1	5.14	122.87	120.30
1	B	167	ARG	NE-CZ-NH2	5.09	122.85	120.30
1	B	167	ARG	NE-CZ-NH1	-5.04	117.78	120.30
1	D	693	LEU	CB-CG-CD1	-5.04	102.44	111.00
1	D	290	ARG	NE-CZ-NH2	-5.03	117.78	120.30
1	B	308	LYS	CD-CE-NZ	5.00	123.21	111.70

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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	3480	0	3401	19	1
1	B	3495	0	3419	30	1
1	C	3520	0	3446	21	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	3520	0	3446	31	0
2	E	56	0	47	1	0
2	F	63	0	54	0	0
2	G	67	0	57	0	0
2	H	63	0	54	0	0
3	E	14	0	13	0	0
3	F	14	0	13	0	0
3	G	14	0	13	0	0
3	H	14	0	13	0	0
4	A	317	0	0	4	0
4	B	313	0	0	8	0
4	C	277	0	0	7	0
4	D	289	0	0	5	1
4	E	13	0	0	0	0
4	F	8	0	0	0	0
4	G	6	0	0	0	0
4	H	9	0	0	0	0
All	All	15552	0	13976	94	2

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 3.

All (94) 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:663:CYS:CB	1:A:663:CYS:SG	2.02	1.46
1:C:231:SER:HB2	4:C:801:HOH:O	1.45	1.13
1:D:261:VAL:HG13	4:D:809:HOH:O	1.68	0.92
1:D:306:ARG:HH11	1:D:306:ARG:HG3	1.40	0.87
1:C:663:CYS:SG	1:C:666:HIS:NE2	2.52	0.82
1:A:288:GLN:OE1	1:C:289:LYS:HG2	1.81	0.81
1:C:231:SER:CB	4:C:801:HOH:O	2.09	0.79
1:B:603:LYS:N	4:B:801:HOH:O	2.16	0.78
1:B:59:HIS:N	4:B:802:HOH:O	2.18	0.77
1:D:373:LEU:N	4:D:803:HOH:O	2.22	0.72
1:C:148:PRO:HA	4:C:995:HOH:O	1.88	0.72
1:C:381:LYS:HE2	1:C:381:LYS:HA	1.73	0.69
1:A:156:ARG:NH2	4:A:801:HOH:O	2.25	0.67
1:A:695:ILE:HG12	1:A:695:ILE:O	1.95	0.67
1:B:581:GLU:CD	1:B:613:LYS:HZ1	2.00	0.65
1:B:635:THR:HG21	4:B:816:HOH:O	1.98	0.64

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:664:ARG:HG2	1:B:665:SER:H	1.66	0.60
1:B:75:MET:SD	4:B:1061:HOH:O	2.56	0.60
1:B:664:ARG:HG2	1:B:665:SER:N	2.16	0.60
1:B:79:LYS:HD2	4:B:1061:HOH:O	2.01	0.60
1:B:306:ARG:HD3	4:B:803:HOH:O	2.01	0.59
1:C:75:MET:HB3	1:C:79:LYS:NZ	2.19	0.58
1:B:581:GLU:HG3	1:B:613:LYS:HZ2	1.68	0.57
1:C:257:LYS:HB2	1:C:569:TYR:CD2	2.39	0.57
1:B:88:TRP:O	1:B:333:LYS:HE2	2.06	0.56
1:A:695:ILE:HD11	1:C:672:LEU:HD11	1.88	0.56
1:C:604:ILE:N	4:C:808:HOH:O	2.39	0.56
1:A:159:ASP:OD2	1:A:205:TYR:OH	2.21	0.55
1:D:76:GLU:CD	1:D:76:GLU:H	2.11	0.54
1:D:306:ARG:HH11	1:D:306:ARG:CG	2.16	0.54
1:A:141:LEU:HD13	4:A:930:HOH:O	2.07	0.53
1:B:59:HIS:N	4:B:810:HOH:O	2.40	0.53
1:B:105:MET:CE	1:D:630:ASN:HD21	2.22	0.52
1:C:89:GLU:HB2	1:C:333:LYS:HD3	1.90	0.52
1:C:235:ARG:HH12	1:C:272:LYS:HE2	1.75	0.52
1:C:552:SER:HA	4:C:868:HOH:O	2.11	0.51
1:D:232:PRO:HD2	4:D:868:HOH:O	2.10	0.51
1:B:149:LYS:NZ	4:B:809:HOH:O	2.38	0.51
1:D:296:TYR:CZ	1:D:299:ARG:HD3	2.46	0.51
1:B:581:GLU:CD	1:B:613:LYS:NZ	2.65	0.50
1:D:552:SER:HA	4:D:893:HOH:O	2.12	0.50
1:D:261:VAL:HG22	1:D:262:GLU:OE2	2.12	0.49
1:B:393:VAL:HG22	1:B:394:PRO:HD2	1.94	0.49
1:D:235:ARG:HA	4:D:824:HOH:O	2.12	0.49
1:A:235:ARG:NH1	4:A:810:HOH:O	2.46	0.48
1:C:261:VAL:HG13	4:C:899:HOH:O	2.13	0.48
1:C:381:LYS:HA	1:C:381:LYS:CE	2.41	0.48
1:A:175:ASN:HA	2:E:146:TRP:CE3	2.50	0.47
1:D:262:GLU:H	1:D:262:GLU:CD	2.18	0.47
1:A:107:TRP:O	1:A:157:LYS:NZ	2.43	0.46
1:A:96:ALA:N	1:A:97:PRO:HD3	2.30	0.46
1:B:610:ARG:HA	1:B:613:LYS:HE2	1.97	0.46
1:B:105:MET:CE	1:D:630:ASN:ND2	2.79	0.46
1:B:581:GLU:HG3	1:B:613:LYS:NZ	2.30	0.46
1:B:700:ASP:OD2	1:D:691:ARG:HD2	2.16	0.45
1:C:559:GLN:NE2	4:C:810:HOH:O	2.47	0.45
1:A:196:VAL:HG11	1:A:240:LYS:HB2	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:610:ARG:O	1:B:613:LYS:HE2	2.16	0.45
1:B:683:GLY:HA2	1:D:289:LYS:HB3	1.98	0.45
1:D:306:ARG:HG3	1:D:306:ARG:NH1	2.19	0.45
1:A:101:TYR:CE1	1:A:106:PHE:CD2	3.05	0.45
1:D:97:PRO:HG2	1:D:100:ASP:HB2	1.99	0.45
1:D:269:LYS:O	1:D:272:LYS:HD2	2.16	0.44
1:A:375:SER:HB3	1:A:378:MET:HB2	1.98	0.44
1:C:650:ILE:HD12	1:C:650:ILE:HA	1.85	0.44
1:D:162:SER:HB3	1:D:169:PHE:HZ	1.82	0.44
1:B:105:MET:HE3	1:B:141:LEU:HD13	2.01	0.43
1:A:88:TRP:O	1:A:333:LYS:HE3	2.19	0.43
1:B:289:LYS:HD3	1:D:683:GLY:HA2	1.99	0.43
1:A:625:PHE:HD1	1:A:644:VAL:HG12	1.84	0.42
1:B:192:ALA:O	1:B:196:VAL:HG13	2.19	0.42
1:B:634:ARG:HA	1:B:634:ARG:HD2	1.87	0.42
1:C:296:TYR:CZ	1:C:299:ARG:HD3	2.54	0.42
1:D:151:VAL:HG22	1:D:198:ILE:HD12	2.01	0.42
1:D:622:MET:O	1:D:626:THR:HG23	2.18	0.42
1:D:669:ALA:HB1	1:D:673:ILE:HD11	2.02	0.42
1:A:604:ILE:N	4:A:818:HOH:O	2.53	0.42
1:D:231:SER:HB3	1:D:234:LEU:HB2	2.02	0.42
1:D:119:LEU:O	1:D:123:ILE:HG13	2.20	0.42
1:D:220:CYS:HB3	1:D:250:THR:OG1	2.20	0.42
1:D:123:ILE:HD13	1:D:166:CYS:HB2	2.02	0.41
1:C:555:LEU:O	1:C:559:GLN:HG3	2.20	0.41
1:D:322:TYR:OH	1:D:386:GLU:HG2	2.20	0.41
1:A:200:ASN:O	1:A:204:GLN:HG3	2.20	0.41
1:B:168:SER:HB3	1:B:211:THR:HB	2.02	0.41
1:B:105:MET:CE	1:B:105:MET:HA	2.51	0.41
1:A:220:CYS:HB3	1:A:250:THR:OG1	2.20	0.41
1:B:172:LEU:HD22	1:B:215:CYS:HB2	2.03	0.41
1:D:107:TRP:CD1	1:D:139:PRO:HA	2.56	0.41
1:B:220:CYS:HB3	1:B:250:THR:OG1	2.20	0.40
1:C:375:SER:HB3	1:C:378:MET:HB2	2.03	0.40
1:C:622:MET:O	1:C:626:THR:HG23	2.20	0.40
1:D:75:MET:HB3	1:D:79:LYS:NZ	2.36	0.40
1:D:225:TYR:HA	1:D:226:PRO:HA	1.85	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:663:CYS:SG	1:C:663:CYS:SG[2_856]	1.53	0.67
1:B:663:CYS:SG	4:D:801:HOH:O[2_755]	2.13	0.07

5.3   Torsion angles ⓘ

5.3.1   Protein backbone ⓘ

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	414/504 (82%)	403 (97%)	9 (2%)	2 (0%)	29	22
1	B	416/504 (82%)	401 (96%)	14 (3%)	1 (0%)	47	45
1	C	421/504 (84%)	406 (96%)	14 (3%)	1 (0%)	47	45
1	D	421/504 (84%)	404 (96%)	16 (4%)	1 (0%)	47	45
2	E	5/11 (46%)	5 (100%)	0	0	100	100
2	F	6/11 (54%)	5 (83%)	1 (17%)	0	100	100
2	G	7/11 (64%)	6 (86%)	1 (14%)	0	100	100
2	H	6/11 (54%)	6 (100%)	0	0	100	100
All	All	1696/2060 (82%)	1636 (96%)	55 (3%)	5 (0%)	41	36

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	314	PRO
1	A	664	ARG
1	B	314	PRO
1	C	314	PRO
1	D	314	PRO

5.3.2   Protein sidechains ⓘ

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	375/442 (85%)	365 (97%)	10 (3%)	44	43
1	B	377/442 (85%)	365 (97%)	12 (3%)	39	37
1	C	381/442 (86%)	368 (97%)	13 (3%)	37	34
1	D	381/442 (86%)	370 (97%)	11 (3%)	42	40
2	E	7/10 (70%)	7 (100%)	0	100	100
2	F	8/10 (80%)	8 (100%)	0	100	100
2	G	8/10 (80%)	8 (100%)	0	100	100
2	H	8/10 (80%)	8 (100%)	0	100	100
All	All	1545/1808 (86%)	1499 (97%)	46 (3%)	41	39

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	65	VAL
1	A	101	TYR
1	A	314	PRO
1	A	583	GLN
1	A	613	LYS
1	A	631	CYS
1	A	648	LYS
1	A	663	CYS
1	A	665	SER
1	A	695	ILE
1	B	101	TYR
1	B	141	LEU
1	B	189	SER
1	B	197	SER
1	B	210	GLU
1	B	278	TRP
1	B	314	PRO
1	B	377	GLN
1	B	603	LYS
1	B	613	LYS
1	B	648	LYS
1	B	664	ARG
1	C	97	PRO
1	C	100	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	101	TYR
1	C	191	PHE
1	C	210	GLU
1	C	278	TRP
1	C	373	LEU
1	C	381	LYS
1	C	605	GLU
1	C	629	SER
1	C	648	LYS
1	C	656	SER
1	C	663	CYS
1	D	100	ASP
1	D	101	TYR
1	D	235	ARG
1	D	278	TRP
1	D	306	ARG
1	D	312	THR
1	D	377	GLN
1	D	382	LEU
1	D	609	SER
1	D	648	LYS
1	D	656	SER

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	377	GLN
1	D	630	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

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	#$|Z| > 2$
3	NAG	G	200	2	14,14,15	3.44	8 (57%)	17,19,21	3.22	8 (47%)
3	NAG	H	200	2	14,14,15	3.26	4 (28%)	17,19,21	3.39	8 (47%)
3	NAG	F	200	2	14,14,15	3.34	5 (35%)	17,19,21	2.61	8 (47%)
3	NAG	E	200	2	14,14,15	3.86	7 (50%)	17,19,21	2.51	7 (41%)

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	NAG	G	200	2	-	1/6/23/26	0/1/1/1
3	NAG	H	200	2	-	2/6/23/26	0/1/1/1
3	NAG	F	200	2	-	2/6/23/26	0/1/1/1
3	NAG	E	200	2	-	2/6/23/26	0/1/1/1

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	200	NAG	C1-C2	11.89	1.70	1.52
3	F	200	NAG	C3-C2	9.26	1.72	1.52
3	H	200	NAG	C1-C2	8.18	1.64	1.52
3	H	200	NAG	C3-C2	7.16	1.67	1.52
3	G	200	NAG	C1-C2	6.81	1.62	1.52
3	G	200	NAG	C3-C2	6.64	1.66	1.52
3	F	200	NAG	C1-C2	5.03	1.59	1.52
3	E	200	NAG	C8-C7	4.98	1.60	1.50
3	G	200	NAG	O5-C5	4.49	1.52	1.43
3	G	200	NAG	O5-C1	-3.88	1.37	1.43
3	G	200	NAG	C8-C7	3.44	1.57	1.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	200	NAG	C4-C5	3.42	1.60	1.53
3	H	200	NAG	C6-C5	3.22	1.62	1.51
3	F	200	NAG	C8-C7	3.11	1.57	1.50
3	E	200	NAG	C3-C2	3.01	1.58	1.52
3	E	200	NAG	O5-C1	2.99	1.48	1.43
3	G	200	NAG	C4-C5	2.93	1.59	1.53
3	G	200	NAG	C2-N2	-2.56	1.42	1.46
3	H	200	NAG	C4-C3	2.53	1.58	1.52
3	G	200	NAG	C6-C5	2.42	1.60	1.51
3	F	200	NAG	O5-C5	2.34	1.48	1.43
3	E	200	NAG	C4-C5	2.29	1.57	1.53
3	E	200	NAG	C6-C5	2.17	1.59	1.51
3	E	200	NAG	O4-C4	2.07	1.47	1.43

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	200	NAG	C1-O5-C5	9.14	124.57	112.19
3	G	200	NAG	C1-O5-C5	8.07	123.13	112.19
3	G	200	NAG	O3-C3-C2	-5.70	97.68	109.47
3	F	200	NAG	C1-O5-C5	5.69	119.90	112.19
3	H	200	NAG	C3-C4-C5	-5.36	100.67	110.24
3	H	200	NAG	O5-C1-C2	5.23	119.54	111.29
3	F	200	NAG	C3-C4-C5	-4.97	101.38	110.24
3	G	200	NAG	C3-C4-C5	-4.79	101.69	110.24
3	E	200	NAG	C3-C4-C5	-4.69	101.87	110.24
3	E	200	NAG	C8-C7-N2	4.51	123.74	116.10
3	E	200	NAG	O3-C3-C2	-3.97	101.24	109.47
3	H	200	NAG	O3-C3-C2	-3.89	101.42	109.47
3	G	200	NAG	C2-N2-C7	3.89	128.44	122.90
3	E	200	NAG	C1-O5-C5	3.84	117.39	112.19
3	F	200	NAG	C8-C7-N2	3.43	121.91	116.10
3	F	200	NAG	O5-C1-C2	3.39	116.64	111.29
3	F	200	NAG	O3-C3-C2	-3.31	102.61	109.47
3	G	200	NAG	C8-C7-N2	3.27	121.64	116.10
3	H	200	NAG	O5-C5-C6	-3.18	102.21	107.20
3	E	200	NAG	O5-C1-C2	2.99	116.02	111.29
3	E	200	NAG	O5-C5-C4	-2.83	103.93	110.83
3	H	200	NAG	C2-N2-C7	2.82	126.92	122.90
3	G	200	NAG	O5-C1-C2	2.69	115.53	111.29
3	E	200	NAG	O7-C7-C8	-2.64	117.16	122.06
3	G	200	NAG	C1-C2-N2	2.54	114.82	110.49

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	200	NAG	O7-C7-C8	-2.48	117.45	122.06
3	F	200	NAG	O5-C5-C4	-2.44	104.89	110.83
3	H	200	NAG	O7-C7-C8	-2.29	117.80	122.06
3	H	200	NAG	C8-C7-N2	2.27	119.94	116.10
3	G	200	NAG	O7-C7-C8	-2.19	117.99	122.06
3	F	200	NAG	O3-C3-C4	-2.18	105.32	110.35

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	H	200	NAG	C1-C2-N2-C7
3	E	200	NAG	C1-C2-N2-C7
3	G	200	NAG	C3-C2-N2-C7
3	E	200	NAG	C3-C2-N2-C7
3	F	200	NAG	C3-C2-N2-C7
3	H	200	NAG	C3-C2-N2-C7
3	F	200	NAG	C1-C2-N2-C7

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, 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	424/504 (84%)	-0.25	3 (0%)	87	90	10, 25, 47, 71	0
1	B	426/504 (84%)	-0.19	5 (1%)	79	83	10, 26, 51, 74	0
1	C	429/504 (85%)	-0.26	0	100	100	12, 27, 51, 70	0
1	D	429/504 (85%)	-0.20	3 (0%)	87	90	12, 29, 54, 75	0
2	E	7/11 (63%)	-0.57	0	100	100	22, 28, 31, 34	0
2	F	8/11 (72%)	-0.23	0	100	100	23, 29, 38, 54	0
2	G	9/11 (81%)	-0.27	0	100	100	24, 30, 38, 50	0
2	H	8/11 (72%)	-0.36	0	100	100	26, 34, 39, 39	0
All	All	1740/2060 (84%)	-0.23	11 (0%)	89	91	10, 27, 51, 75	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	604	ILE	3.6
1	A	605	GLU	3.2
1	A	553	VAL	3.2
1	A	664	ARG	3.0
1	D	694	PRO	2.6
1	B	679	TRP	2.4
1	D	237	VAL	2.3
1	B	182	ALA	2.3
1	D	181	CYS	2.3
1	B	694	PRO	2.1
1	B	178	HIS	2.1

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

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

WORLDWIDE PDB PROTEIN DATA BANK

6.3　Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4　Ligands ⓘ

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, 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	B-factors(Å^2)	Q<0.9
3	NAG	E	200	14/15	0.98	0.12	12,16,18,19	0
3	NAG	F	200	14/15	0.98	0.13	12,16,21,22	0
3	NAG	G	200	14/15	0.98	0.12	13,15,20,21	0
3	NAG	H	200	14/15	0.98	0.13	15,17,23,24	0

6.5　Other polymers ⓘ

There are no such residues in this entry.