W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB X-ray Structure Validation Report (i)

Feb 14, 2017 – 10:37 pm GMT

PDB ID : 1OA2

Title : COMPARISON OF FAMILY 12 GLYCOSIDE HYDROLASES AND RE-CRUITED SUBSTITUTIONS IMPORTANT FOR THERMAL STABILITY

Authors : Sandgren, M.; Gualfetti, P.J.; Shaw, A.; Gross, L.S.; Saldajeno, M.; Day, A.G.; Jones, T.A.; Mitchinson, C.

Deposited on : 2002-12-28

Resolution : 1.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

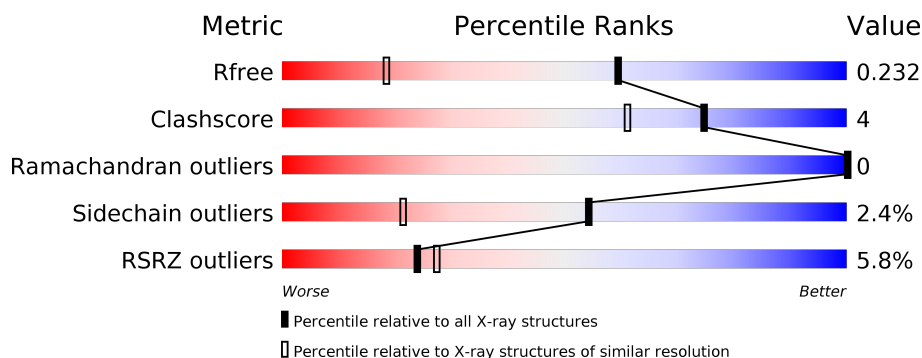|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 1.50 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		0.232
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		2.4%
RSRZ outliers		5.8%

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}	100719	2279 (1.50-1.50)
Clashscore	112137	2503 (1.50-1.50)
Ramachandran outliers	110173	2445 (1.50-1.50)
Sidechain outliers	110143	2443 (1.50-1.50)
RSRZ outliers	101464	2305 (1.50-1.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	218	2% — 93% — 5% •
1	B	218	91% — 7% •
1	C	218	2% — 88% — 11% •
1	D	218	13% — 89% — 10% •
1	E	218	17% — 90% — 10%
1	F	218	94% — 5%

WORLDWIDE PDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 10703 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 ENDO-BETA-1,4-GLUCANASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	218	Total	C	N	O	S	0	0	0
			1663	1047	278	333	5			
1	B	218	Total	C	N	O	S	0	0	0
			1663	1047	278	333	5			
1	C	218	Total	C	N	O	S	0	0	0
			1663	1047	278	333	5			
1	D	218	Total	C	N	O	S	0	0	0
			1663	1047	278	333	5			
1	E	218	Total	C	N	O	S	0	0	0
			1663	1047	278	333	5			
1	F	218	Total	C	N	O	S	0	0	0
			1663	1047	278	333	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	PCA	GLN	MODIFIED RESIDUE	UNP O00095
B	1	PCA	GLN	MODIFIED RESIDUE	UNP O00095
C	1	PCA	GLN	MODIFIED RESIDUE	UNP O00095
D	1	PCA	GLN	MODIFIED RESIDUE	UNP O00095
E	1	PCA	GLN	MODIFIED RESIDUE	UNP O00095
F	1	PCA	GLN	MODIFIED RESIDUE	UNP O00095
A	35	VAL	ALA	ENGINEERED MUTATION	UNP O00095
B	35	VAL	ALA	ENGINEERED MUTATION	UNP O00095
C	35	VAL	ALA	ENGINEERED MUTATION	UNP O00095
D	35	VAL	ALA	ENGINEERED MUTATION	UNP O00095
E	35	VAL	ALA	ENGINEERED MUTATION	UNP O00095
F	35	VAL	ALA	ENGINEERED MUTATION	UNP O00095

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	108	Total 108	O 108	0	0
3	B	168	Total 168	O 168	0	0
3	C	112	Total 112	O 112	0	0
3	D	59	Total 59	O 59	0	0
3	E	63	Total 63	O 63	0	0
3	F	131	Total 131	O 131	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.
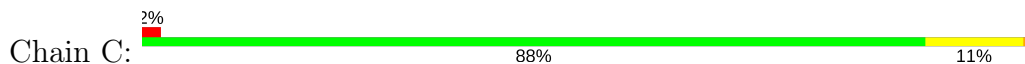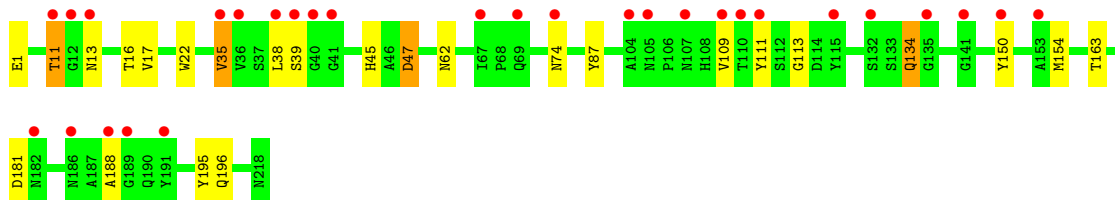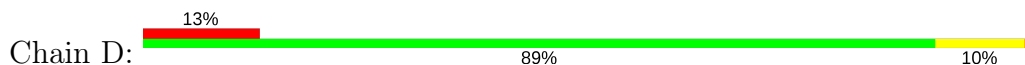
- Molecule 1: ENDO-BETA-1,4-GLUCANASE
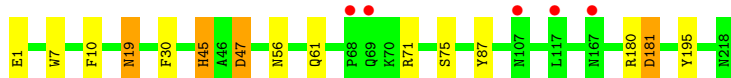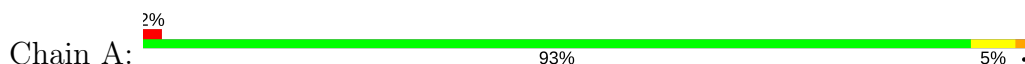
Chain A:

- Molecule 1: ENDO-BETA-1,4-GLUCANASE

Chain B:

- Molecule 1: ENDO-BETA-1,4-GLUCANASE

Chain C:

- Molecule 1: ENDO-BETA-1,4-GLUCANASE
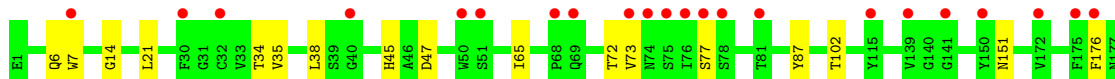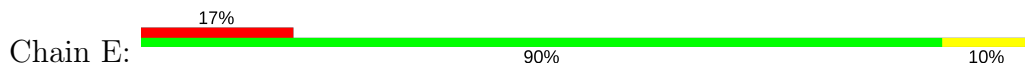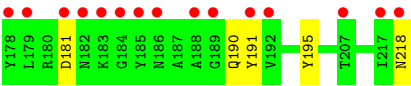
Chain D:

- Molecule 1: ENDO-BETA-1,4-GLUCANASE

Chain E:

Y178 L179 R180 D181 N182 K183 G184 Y185 N186 A187 A188 G189 Q190 Y191 V192 Y195 T207 I217 N218

• Molecule 1: ENDO-BETA-1,4-GLUCANASE

Chain F:

94%　　　　　　5%

E1 N19 W22 V33 H45 A46 D47 Q61 Q69 I76 Y87 K173 R180 D181 Y195 N218
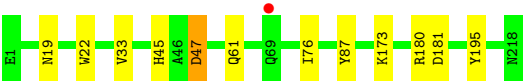
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	68.28Å　　71.29Å　　119.30Å 90.00°　　91.49°　　90.00°	Depositor
Resolution (Å)	20.00　–　1.50 19.88　–　1.50	Depositor EDS
% Data completeness (in resolution range)	94.3 (20.00-1.50) 94.3 (19.88-1.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	2.37 (at 1.50Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.205　,　0.223 0.213　,　0.232	Depositor DCC
R_{free} test set	5174 reflections (3.09%)	DCC
Wilson B-factor (Å^2)	16.3	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 40.5	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.000 for -k,-h,-l 0.000 for k,h,-l 0.026 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10703	wwPDB-VP
Average B, all atoms (Å^2)	20.0	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 31.00 % 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 1.1914e-03. 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: PCA, 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	0.54	0/1704	0.78	2/2329 (0.1%)
1	B	0.58	0/1704	0.80	3/2329 (0.1%)
1	C	0.53	0/1704	0.77	3/2329 (0.1%)
1	D	0.49	0/1704	0.73	2/2329 (0.1%)
1	E	0.49	0/1704	0.74	1/2329 (0.0%)
1	F	0.56	0/1704	0.77	1/2329 (0.0%)
All	All	0.53	0/10224	0.77	12/13974 (0.1%)

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	A	0	1

There are no bond length outliers.

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	181	ASP	CB-CG-OD2	6.33	124.00	118.30
1	C	47	ASP	CB-CG-OD2	6.01	123.71	118.30
1	A	47	ASP	CB-CG-OD2	6.00	123.70	118.30
1	B	47	ASP	CB-CG-OD2	5.92	123.62	118.30
1	F	47	ASP	CB-CG-OD2	5.90	123.61	118.30
1	A	181	ASP	CB-CG-OD2	5.80	123.52	118.30
1	D	181	ASP	CB-CG-OD2	5.71	123.44	118.30
1	C	126	ASP	CB-CG-OD2	5.67	123.40	118.30
1	B	181	ASP	CB-CG-OD2	5.24	123.01	118.30
1	C	99	ASP	CB-CG-OD1	5.15	122.94	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	47	ASP	CB-CG-OD2	5.05	122.84	118.30
1	B	126	ASP	CB-CG-OD2	5.04	122.83	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	45	HIS	Sidechain

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	1663	0	1517	10	0
1	B	1663	0	1517	10	0
1	C	1663	0	1517	15	0
1	D	1663	0	1517	20	0
1	E	1663	0	1517	14	0
1	F	1663	0	1517	8	0
2	A	14	0	13	0	0
2	B	14	0	13	0	0
2	C	14	0	13	0	0
2	D	14	0	13	0	0
2	E	14	0	13	0	0
2	F	14	0	13	0	0
3	A	108	0	0	0	0
3	B	168	0	0	1	0
3	C	112	0	0	1	0
3	D	59	0	0	0	0
3	E	63	0	0	1	0
3	F	131	0	0	0	0
All	All	10703	0	9180	72	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 (72) 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:E:35:VAL:HG11	1:E:38:LEU:HD21	1.51	0.92
1:E:45:HIS:HD2	1:E:47:ASP:OD1	1.61	0.83
1:A:180:ARG:HD2	1:A:181:ASP:OD1	1.81	0.81
1:E:35:VAL:HG11	1:E:38:LEU:CD2	2.15	0.77
1:A:45:HIS:HD2	1:A:47:ASP:OD1	1.70	0.73
1:E:45:HIS:CD2	1:E:47:ASP:OD1	2.44	0.70
1:A:71:ARG:NH2	1:A:75:SER:O	2.28	0.66
1:B:45:HIS:HD2	1:B:47:ASP:OD1	1.77	0.66
1:C:45:HIS:HD2	1:C:47:ASP:OD1	1.79	0.65
1:D:45:HIS:HD2	1:D:47:ASP:OD1	1.79	0.65
1:D:74:ASN:HD22	1:D:188:ALA:CA	2.10	0.65
1:D:74:ASN:HD22	1:D:188:ALA:HA	1.62	0.64
1:E:35:VAL:CG1	1:E:38:LEU:HD21	2.23	0.64
1:D:150:TYR:HD2	3:E:2014:HOH:O	1.80	0.64
1:D:17:VAL:HG23	1:D:38:LEU:HD21	1.80	0.64
1:F:45:HIS:HD2	1:F:47:ASP:OD1	1.81	0.63
1:D:113:GLY:O	1:D:154:MET:HG2	1.99	0.62
1:D:22:TRP:HB2	1:E:7:TRP:CZ2	2.36	0.60
1:C:132:SER:O	1:C:134:GLN:NE2	2.33	0.60
1:E:35:VAL:CG1	1:E:38:LEU:CD2	2.79	0.60
1:F:180:ARG:HD2	1:F:181:ASP:OD1	2.03	0.59
1:A:180:ARG:CD	1:A:181:ASP:OD1	2.51	0.57
1:C:76:ILE:HD12	1:C:217:ILE:CG2	2.35	0.57
1:D:13:ASN:HB2	1:D:39:SER:HA	1.87	0.55
1:E:73:VAL:HG11	1:E:176:PHE:HB3	1.89	0.55
1:C:40:GLY:N	3:C:2032:HOH:O	2.40	0.54
1:A:56:ASN:ND2	1:B:56:ASN:HD21	2.06	0.53
1:F:19:ASN:ND2	1:F:61:GLN:HE21	2.08	0.51
1:C:91:ASN:ND2	1:D:163:THR:HG23	2.25	0.51
1:C:71:ARG:HB2	1:C:76:ILE:HD11	1.93	0.50
1:E:6:GLN:HG3	1:E:21:LEU:HB2	1.93	0.50
1:E:77:SER:N	1:E:218:ASN:O	2.44	0.49
1:B:151:ASN:C	1:B:151:ASN:HD22	2.16	0.48
1:C:74:ASN:HD22	1:C:188:ALA:CA	2.26	0.48
1:B:19:ASN:ND2	1:B:61:GLN:HE21	2.12	0.48
1:A:7:TRP:CZ2	1:B:22:TRP:HB2	2.49	0.47
1:E:14:GLY:O	1:E:65:ILE:HA	2.14	0.47
1:C:73:VAL:O	1:C:173:LYS:NZ	2.47	0.47
1:B:19:ASN:HD21	1:B:61:GLN:HE21	1.63	0.47
1:F:19:ASN:HD22	1:F:33:VAL:HG13	1.80	0.47
1:C:19:ASN:HD22	1:C:33:VAL:HG13	1.80	0.46
1:E:72:THR:HA	1:E:190:GLN:O	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:13:ASN:HB2	1:D:38:LEU:O	2.16	0.46
1:B:177:ASN:ND2	3:B:2139:HOH:O	2.25	0.45
1:D:74:ASN:ND2	1:D:188:ALA:N	2.64	0.45
1:A:19:ASN:ND2	1:A:61:GLN:HE21	2.14	0.44
1:B:173:LYS:HA	1:B:173:LYS:HD2	1.81	0.43
1:D:62:ASN:ND2	1:D:196:GLN:HE21	2.16	0.43
1:C:22:TRP:O	1:F:22:TRP:HB2	2.19	0.43
1:C:74:ASN:HD22	1:C:188:ALA:N	2.17	0.43
1:F:76:ILE:O	1:F:173:LYS:NZ	2.52	0.42
1:B:62:ASN:ND2	1:B:196:GLN:HE21	2.18	0.42
1:C:180:ARG:HD2	1:C:181:ASP:OD1	2.20	0.42
1:D:109:VAL:HG21	1:D:111:TYR:CE2	2.54	0.42
1:D:35:VAL:HG21	1:D:38:LEU:HG	2.01	0.42
1:D:13:ASN:CB	1:D:39:SER:HA	2.50	0.42
1:E:102:THR:HA	1:E:191:TYR:O	2.20	0.41
1:F:45:HIS:CD2	1:F:47:ASP:OD1	2.68	0.41
1:C:74:ASN:ND2	1:C:188:ALA:N	2.69	0.41
1:E:34:THR:HB	1:E:45:HIS:CE1	2.55	0.41
1:D:74:ASN:ND2	1:D:188:ALA:CA	2.79	0.41
1:D:109:VAL:HG21	1:D:111:TYR:CZ	2.55	0.41
1:A:19:ASN:HD21	1:A:61:GLN:HE21	1.68	0.41
1:D:11:THR:HG22	1:D:16:THR:OG1	2.20	0.41
1:A:30:PHE:CD1	1:A:30:PHE:C	2.94	0.41
1:F:19:ASN:HD21	1:F:61:GLN:HE21	1.68	0.41
1:A:1:PCA:HG2	1:A:10:PHE:CE1	2.55	0.41
1:D:134:GLN:HE21	1:D:134:GLN:CA	2.34	0.40
1:C:102:THR:HA	1:C:191:TYR:O	2.22	0.40
1:B:102:THR:HA	1:B:191:TYR:O	2.21	0.40
1:C:72:THR:O	1:C:76:ILE:HG12	2.21	0.40
1:D:1:PCA:CD	1:D:35:VAL:HG13	2.52	0.40

There are no symmetry-related clashes.

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	216/218 (99%)	210 (97%)	6 (3%)	0	100	100
1	B	216/218 (99%)	214 (99%)	2 (1%)	0	100	100
1	C	216/218 (99%)	211 (98%)	5 (2%)	0	100	100
1	D	216/218 (99%)	211 (98%)	5 (2%)	0	100	100
1	E	216/218 (99%)	211 (98%)	5 (2%)	0	100	100
1	F	216/218 (99%)	210 (97%)	6 (3%)	0	100	100
All	All	1296/1308 (99%)	1267 (98%)	29 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	173/173 (100%)	170 (98%)	3 (2%)	66	36
1	B	173/173 (100%)	166 (96%)	7 (4%)	36	7
1	C	173/173 (100%)	168 (97%)	5 (3%)	48	15
1	D	173/173 (100%)	168 (97%)	5 (3%)	48	15
1	E	173/173 (100%)	170 (98%)	3 (2%)	66	36
1	F	173/173 (100%)	171 (99%)	2 (1%)	75	52
All	All	1038/1038 (100%)	1013 (98%)	25 (2%)	54	21

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19	ASN
1	A	87	TYR
1	A	195	TYR
1	B	19	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	74	ASN
1	B	87	TYR
1	B	111	TYR
1	B	151	ASN
1	B	173	LYS
1	B	195	TYR
1	C	19	ASN
1	C	87	TYR
1	C	134	GLN
1	C	151	ASN
1	C	195	TYR
1	D	11	THR
1	D	35	VAL
1	D	87	TYR
1	D	134	GLN
1	D	195	TYR
1	E	87	TYR
1	E	151	ASN
1	E	195	TYR
1	F	87	TYR
1	F	195	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (47) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	19	ASN
1	A	20	ASN
1	A	45	HIS
1	A	62	ASN
1	A	142	GLN
1	A	151	ASN
1	A	155	GLN
1	A	162	GLN
1	B	19	ASN
1	B	20	ASN
1	B	45	HIS
1	B	56	ASN
1	B	62	ASN
1	B	64	GLN
1	B	151	ASN
1	B	155	GLN
1	B	182	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	19	ASN
1	C	20	ASN
1	C	45	HIS
1	C	62	ASN
1	C	74	ASN
1	C	138	ASN
1	C	142	GLN
1	C	151	ASN
1	D	13	ASN
1	D	20	ASN
1	D	45	HIS
1	D	56	ASN
1	D	62	ASN
1	D	74	ASN
1	D	134	GLN
1	D	138	ASN
1	D	142	GLN
1	D	155	GLN
1	D	190	GLN
1	E	20	ASN
1	E	45	HIS
1	E	56	ASN
1	E	151	ASN
1	F	19	ASN
1	F	20	ASN
1	F	45	HIS
1	F	49	GLN
1	F	54	GLN
1	F	69	GLN
1	F	134	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

6 non-standard protein/DNA/RNA residues 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	#$|Z| > 2$
1	PCA	A	1	1	8,8,9	1.86	1 (12%)	9,10,12	2.24	5 (55%)
1	PCA	B	1	1	8,8,9	1.79	3 (37%)	9,10,12	2.35	5 (55%)
1	PCA	C	1	1	8,8,9	1.99	3 (37%)	9,10,12	1.71	3 (33%)
1	PCA	D	1	1	8,8,9	1.84	2 (25%)	9,10,12	2.05	5 (55%)
1	PCA	E	1	1	8,8,9	1.68	2 (25%)	9,10,12	1.97	4 (44%)
1	PCA	F	1	1	8,8,9	1.63	2 (25%)	9,10,12	1.94	2 (22%)

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
1	PCA	A	1	1	-	0/0/11/13	0/1/1/1
1	PCA	B	1	1	-	0/0/11/13	0/1/1/1
1	PCA	C	1	1	-	0/0/11/13	0/1/1/1
1	PCA	D	1	1	-	0/0/11/13	0/1/1/1
1	PCA	E	1	1	-	0/0/11/13	0/1/1/1
1	PCA	F	1	1	-	0/0/11/13	0/1/1/1

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1	PCA	CA-N	2.17	1.49	1.46
1	C	1	PCA	CA-N	2.27	1.49	1.46
1	E	1	PCA	CA-N	2.45	1.49	1.46
1	F	1	PCA	CA-N	2.46	1.49	1.46
1	B	1	PCA	CA-C	2.48	1.53	1.50
1	D	1	PCA	CA-N	2.85	1.49	1.46
1	C	1	PCA	CA-C	2.91	1.54	1.50
1	F	1	PCA	CD-N	3.65	1.45	1.34
1	B	1	PCA	CD-N	3.72	1.45	1.34
1	E	1	PCA	CD-N	3.82	1.45	1.34
1	D	1	PCA	CD-N	3.89	1.45	1.34
1	C	1	PCA	CD-N	4.13	1.46	1.34
1	A	1	PCA	CD-N	4.34	1.47	1.34

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1	PCA	CB-CA-C	-3.66	107.67	112.70
1	B	1	PCA	CA-N-CD	-3.50	101.61	113.58
1	A	1	PCA	CA-N-CD	-3.39	101.96	113.58
1	E	1	PCA	CB-CA-C	-3.26	108.22	112.70
1	F	1	PCA	CB-CA-C	-3.22	108.27	112.70
1	D	1	PCA	CA-N-CD	-3.04	103.18	113.58
1	A	1	PCA	OE-CD-CG	-3.03	121.29	126.86
1	F	1	PCA	CA-N-CD	-2.98	103.37	113.58
1	A	1	PCA	CB-CA-C	-2.94	108.66	112.70
1	C	1	PCA	OE-CD-CG	-2.66	121.96	126.86
1	D	1	PCA	O-C-CA	-2.65	118.98	125.15
1	B	1	PCA	O-C-CA	-2.61	119.07	125.15
1	E	1	PCA	CA-N-CD	-2.53	104.93	113.58
1	D	1	PCA	OE-CD-CG	-2.45	122.34	126.86
1	C	1	PCA	CA-N-CD	-2.39	105.39	113.58
1	E	1	PCA	O-C-CA	-2.20	120.02	125.15
1	C	1	PCA	CG-CD-N	2.09	114.28	108.33
1	B	1	PCA	CG-CD-N	2.17	114.50	108.33
1	A	1	PCA	CB-CA-N	2.28	109.86	103.30
1	E	1	PCA	CB-CA-N	2.32	109.95	103.30
1	A	1	PCA	CG-CD-N	2.32	114.92	108.33
1	B	1	PCA	CB-CA-N	2.36	110.07	103.30
1	D	1	PCA	CG-CD-N	2.51	115.47	108.33
1	D	1	PCA	CB-CA-N	2.58	110.70	103.30

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1	PCA	1	0
1	D	1	PCA	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

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	301	1	14,14,15	0.66	0	15,19,21	1.62	2 (13%)
2	NAG	B	301	1	14,14,15	0.63	0	15,19,21	1.33	2 (13%)
2	NAG	C	301	1	14,14,15	0.69	0	15,19,21	1.27	1 (6%)
2	NAG	D	301	1	14,14,15	0.66	0	15,19,21	0.81	0
2	NAG	E	301	1	14,14,15	0.69	0	15,19,21	1.28	2 (13%)
2	NAG	F	301	1	14,14,15	0.79	1 (7%)	15,19,21	1.05	2 (13%)

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	301	1	-	0/6/23/26	0/1/1/1
2	NAG	B	301	1	-	0/6/23/26	0/1/1/1
2	NAG	C	301	1	-	0/6/23/26	0/1/1/1
2	NAG	D	301	1	-	0/6/23/26	0/1/1/1
2	NAG	E	301	1	-	0/6/23/26	0/1/1/1
2	NAG	F	301	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	301	NAG	C1-C2	2.29	1.55	1.52

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	301	NAG	O5-C1-C2	-3.97	105.95	111.47
2	C	301	NAG	O5-C1-C2	-3.37	106.79	111.47

Continued on next page...

Full wwPDB X-ray Structure Validation Report

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	301	NAG	O5-C1-C2	-3.21	107.01	111.47
2	B	301	NAG	O5-C1-C2	-2.60	107.85	111.47
2	F	301	NAG	O5-C1-C2	-2.11	108.54	111.47
2	F	301	NAG	C2-N2-C7	-2.06	119.94	122.94
2	E	301	NAG	C1-O5-C5	2.94	116.22	112.17
2	B	301	NAG	C1-O5-C5	3.36	116.80	112.17
2	A	301	NAG	C1-O5-C5	4.05	117.75	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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(Å²)	Q<0.9
1	A	217/218 (99%)	0.16	5 (2%)	61	66	12, 18, 26, 31	0
1	B	217/218 (99%)	-0.07	0	100	100	9, 15, 23, 27	0
1	C	217/218 (99%)	0.13	4 (1%)	69	74	12, 19, 26, 31	0
1	D	217/218 (99%)	0.87	29 (13%)	4	4	12, 22, 36, 40	0
1	E	217/218 (99%)	1.06	37 (17%)	2	2	12, 26, 37, 41	0
1	F	217/218 (99%)	0.11	1 (0%)	90	93	10, 17, 25, 29	0
All	All	1302/1308 (99%)	0.38	76 (5%)	24	27	9, 19, 32, 41	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	13	ASN	7.6
1	D	150	TYR	5.7
1	D	109	VAL	5.5
1	E	188	ALA	5.0
1	E	184	GLY	4.5
1	E	76	ILE	4.5
1	D	111	TYR	4.2
1	D	69	GLN	3.9
1	E	189	GLY	3.9
1	E	77	SER	3.9
1	E	150	TYR	3.9
1	E	141	GLY	3.8
1	D	39	SER	3.7
1	D	38	LEU	3.6
1	D	41	GLY	3.5
1	D	11	THR	3.4
1	D	36	VAL	3.3
1	E	40	GLY	3.3
1	E	218	ASN	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	107	ASN	3.2
1	E	217	ILE	3.1
1	E	186	ASN	3.1
1	E	185	TYR	3.0
1	D	74	ASN	3.0
1	E	32	CYS	2.9
1	E	7	TRP	2.9
1	D	188	ALA	2.9
1	E	178	TYR	2.9
1	D	132	SER	2.8
1	D	135	GLY	2.8
1	D	105	ASN	2.8
1	D	12	GLY	2.7
1	E	50	TRP	2.7
1	D	153	ALA	2.7
1	E	69	GLN	2.7
1	C	141	GLY	2.6
1	D	67	ILE	2.6
1	D	104	ALA	2.5
1	D	141	GLY	2.5
1	E	74	ASN	2.5
1	F	69	GLN	2.5
1	E	175	PHE	2.5
1	E	176	PHE	2.5
1	D	186	ASN	2.4
1	E	181	ASP	2.4
1	E	182	ASN	2.4
1	D	115	TYR	2.4
1	D	110	THR	2.4
1	E	179	LEU	2.4
1	E	68	PRO	2.4
1	E	172	VAL	2.4
1	A	68	PRO	2.4
1	E	192	VAL	2.4
1	E	191	TYR	2.4
1	D	40	GLY	2.3
1	A	69	GLN	2.3
1	E	78	SER	2.3
1	E	115	TYR	2.3
1	C	150	TYR	2.3
1	D	189	GLY	2.3
1	A	167	ASN	2.2

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	191	TYR	2.2
1	C	68	PRO	2.2
1	E	81	THR	2.2
1	E	207	THR	2.2
1	E	73	VAL	2.2
1	E	139	VAL	2.2
1	E	51	SER	2.1
1	A	117	LEU	2.1
1	E	183	LYS	2.1
1	C	128	GLY	2.1
1	D	182	ASN	2.1
1	E	30	PHE	2.1
1	E	75	SER	2.0
1	A	107	ASN	2.0
1	D	35	VAL	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
1	PCA	F	1	8/9	0.95	0.07	-	17,18,18,18	0
1	PCA	E	1	8/9	0.80	0.14	-	28,28,29,30	0
1	PCA	C	1	8/9	0.94	0.08	-	18,19,20,22	0
1	PCA	A	1	8/9	0.92	0.08	-	21,22,22,24	0
1	PCA	B	1	8/9	0.92	0.10	-	16,18,19,19	0
1	PCA	D	1	8/9	0.89	0.13	-	26,27,27,28	0

6.3 Carbohydrates ⓘ

There are no carbohydrates 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. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å²)	Q<0.9
2	NAG	F	301	14/15	0.97	0.09	0.33	10,11,13,13	0
2	NAG	C	301	14/15	0.94	0.08	-0.51	14,17,19,20	0
2	NAG	E	301	14/15	0.97	0.08	-0.60	12,13,14,15	0
2	NAG	D	301	14/15	0.95	0.08	-1.07	16,18,21,22	0
2	NAG	A	301	14/15	0.96	0.06	-1.40	12,14,19,19	0
2	NAG	B	301	14/15	0.98	0.06	-1.59	11,13,14,15	0

6.5 Other polymers ⓘ

There are no such residues in this entry.