Full wwPDB X-ray Structure Validation Report (i)

Aug 7, 2020 – 06:38 AM BST

PDB ID	:	4ADI
Title	:	Crystal structure of the Rubella virus envelope glycoprotein E1 in post-fusion form (crystal form I)
Authors	:	DuBois, R.M.; Vaney, M.C.; Tortorici, M.A.; Al Kurdi, R.; Barba-Spaeth, G.; Rey, F.A.
Deposited on	:	2011-12-26
Resolution	:	1.80 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	:	2.13.1
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 ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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.
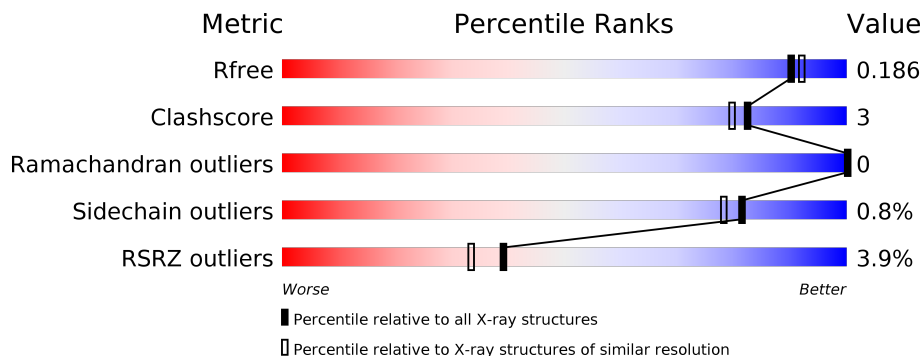
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	473	3% / 85% / 5% / 10%
1	B	473	4% / 85% / 5% / 10%
1	C	473	4% / 86% / • / 9%

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:

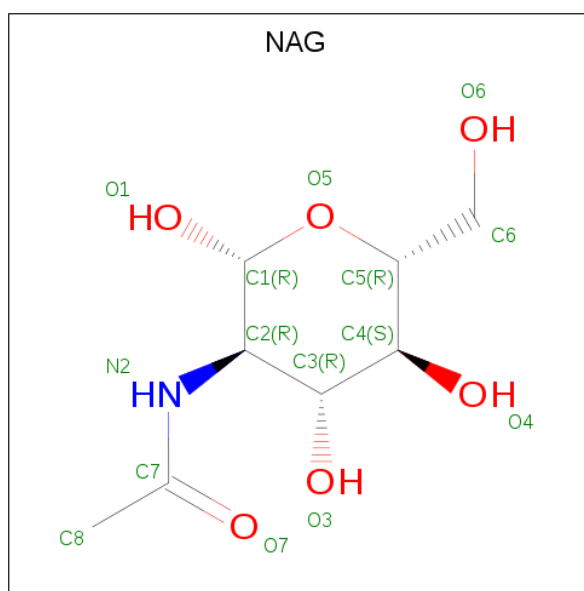| Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	A	1440	-	-	-	X

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 11662 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 E1 ENVELOPE GLYCOPROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	426	Total 3253	C 2060	N 565	O 605	S 23	0	12	1
1	B	425	Total 3259	C 2062	N 568	O 606	S 23	0	14	2
1	C	429	Total 3285	C 2075	N 578	O 609	S 23	0	13	3

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

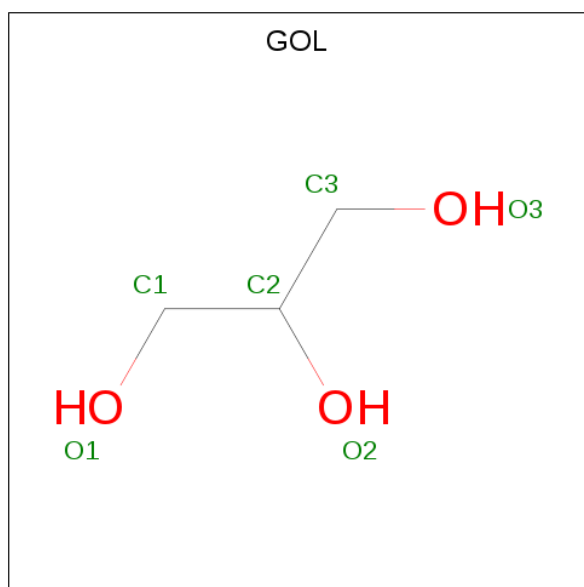| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 14	C 8	N 1	O 5	0	0
2	A	1	Total 14	C 8	N 1	O 5	0	0
2	B	1	Total 14	C 8	N 1	O 5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
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	C	1	Total 14	C 8	N 1	O 5	0	0

• Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 6	C 3	O 3	0	0
3	A	1	Total 6	C 3	O 3	0	0
3	A	1	Total 6	C 3	O 3	0	0
3	A	1	Total 6	C 3	O 3	0	0
3	A	1	Total 6	C 3	O 3	0	0
3	A	1	Total 6	C 3	O 3	0	0
3	B	1	Total 6	C 3	O 3	0	0
3	B	1	Total 6	C 3	O 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total 6	C 3	O 3	0	0
3	C	1	Total 6	C 3	O 3	0	0
3	C	1	Total 6	C 3	O 3	0	0
3	C	1	Total 6	C 3	O 3	0	0
3	C	1	Total 6	C 3	O 3	0	0
3	C	1	Total 6	C 3	O 3	0	0
3	C	1	Total 6	C 3	O 3	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total 1	Na 1	0	0
4	A	1	Total 1	Na 1	0	0
4	C	1	Total 1	Na 1	0	0

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).
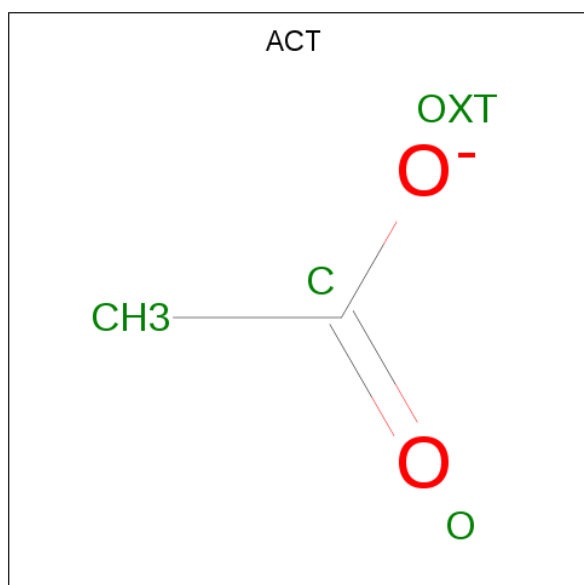
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total 4	C 2	O 2	0	0
5	A	1	Total 4	C 2	O 2	0	0
5	B	1	Total 4	C 2	O 2	0	0
5	B	1	Total 4	C 2	O 2	0	0
5	B	1	Total 4	C 2	O 2	0	0
5	B	1	Total 4	C 2	O 2	0	0
5	C	1	Total 4	C 2	O 2	0	0

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).
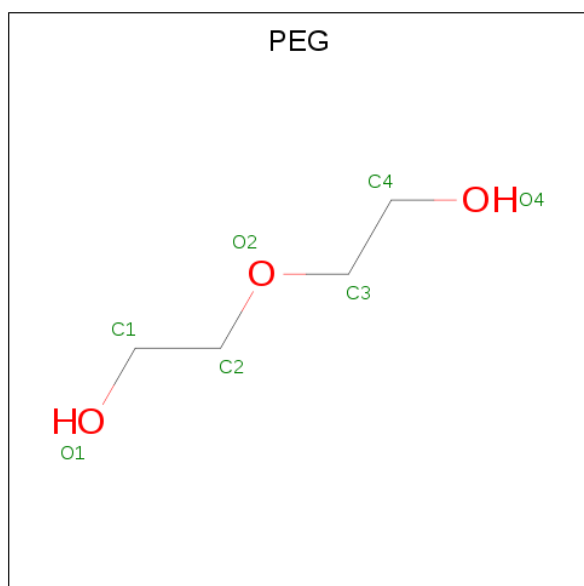
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 7	C 4	O 3	0	0
6	B	1	Total 7	C 4	O 3	0	0
6	C	1	Total 7	C 4	O 3	0	0
6	C	1	Total 7	C 4	O 3	0	0

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-galactopyranose (three-letter code: NGA)

(formula: $C_8H_{15}NO_6$).
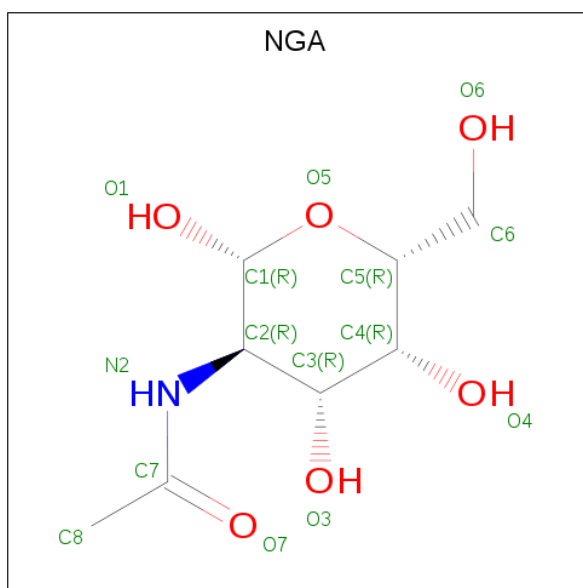
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total 14	C 8	N 1	O 5	0	0
7	A	1	Total 14	C 8	N 1	O 5	0	0
7	B	1	Total 14	C 8	N 1	O 5	0	0
7	B	1	Total 14	C 8	N 1	O 5	0	0
7	C	1	Total 14	C 8	N 1	O 5	0	0
7	C	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	586	Total 586	O 586	0	0
8	B	503	Total 503	O 503	0	0
8	C	458	Total 459	O 459	0	1

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: E1 ENVELOPE GLYCOPROTEIN
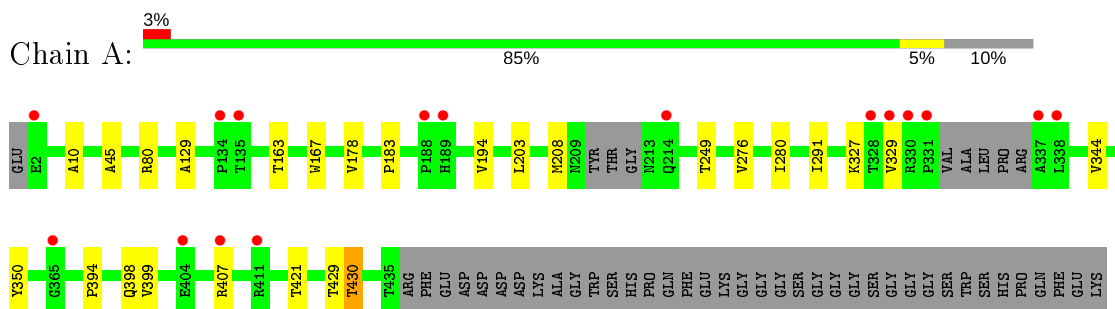
Chain A:

- Molecule 1: E1 ENVELOPE GLYCOPROTEIN

Chain B:

- Molecule 1: E1 ENVELOPE GLYCOPROTEIN

Chain C:
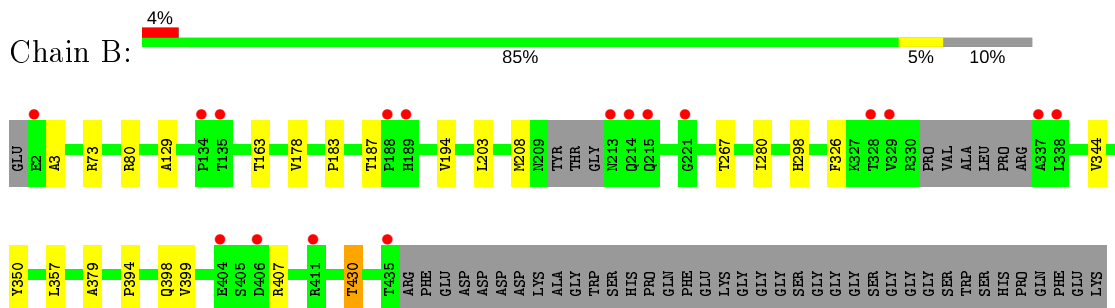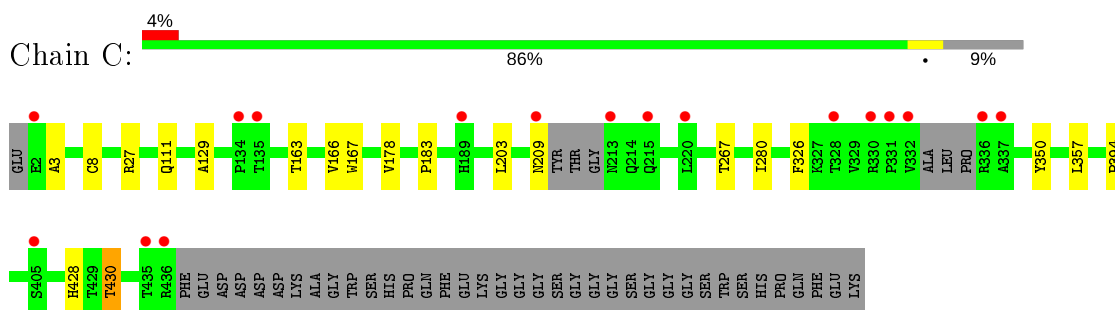
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	129.98Å 121.38Å 126.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.80 44.36 – 1.78	Depositor EDS
% Data completeness (in resolution range)	99.6 (20.00-1.80) 98.3 (44.36-1.78)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 1.78Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, R_{free}	0.167 , 0.183 0.169 , 0.186	Depositor DCC
R_{free} test set	18933 reflections (10.05%)	wwPDB-VP
Wilson B-factor (Å^2)	24.1	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 47.8	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.014 for l,-k,h	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	11662	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, NAG, NA, ACT, PEG, NGA

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	0.37	0/3390	0.59	0/4659
1	B	0.37	0/3396	0.59	0/4664
1	C	0.37	0/3417	0.59	0/4695
All	All	0.37	0/10203	0.59	0/14018

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 (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	3253	0	3097	20	0
1	B	3259	0	3094	19	0
1	C	3285	0	3114	16	0
2	A	28	0	26	0	0
2	B	28	0	26	0	0
2	C	28	0	26	0	0
3	A	36	0	48	3	0
3	B	18	0	24	1	0
3	C	36	0	48	1	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1	0	0	0	0
4	C	1	0	0	0	0
5	A	8	0	6	1	0
5	B	16	0	12	0	0
5	C	4	0	3	0	0
6	A	7	0	10	1	0
6	B	7	0	10	0	0
6	C	14	0	20	0	0
7	A	28	0	26	4	0
7	B	28	0	26	4	0
7	C	28	0	26	4	0
8	A	586	0	0	0	0
8	B	503	0	0	1	0
8	C	459	0	0	0	0
All	All	11662	0	9642	50	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 3.

All (50) 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:430[B]:THR:HB	7:C:4001:NGA:C1	1.86	1.05
1:B:430[B]:THR:HB	7:B:4001:NGA:C1	1.95	0.96
1:A:430[B]:THR:HB	7:A:4001:NGA:C1	2.00	0.90
1:B:73[B]:ARG:NH1	8:B:2106:HOH:O	2.17	0.77
1:A:249:THR:OG1	1:B:298:HIS:HE1	1.76	0.69
1:B:178[B]:VAL:HG13	1:B:183:PRO:HA	1.76	0.67
1:A:430[B]:THR:CG2	7:A:4001:NGA:C1	2.73	0.66
1:A:430[B]:THR:CB	7:A:4001:NGA:C1	2.68	0.65
1:C:430[B]:THR:CG2	7:C:4001:NGA:C1	2.75	0.64
1:A:430[B]:THR:HG22	7:A:4001:NGA:C1	2.26	0.64
1:A:80:ARG:HH22	1:B:267:THR:HG21	1.62	0.63
1:C:430[B]:THR:CB	7:C:4001:NGA:C1	2.61	0.62
1:B:430[B]:THR:CB	7:B:4001:NGA:C1	2.68	0.62
1:B:430[B]:THR:CG2	7:B:4001:NGA:C1	2.80	0.60
3:A:1436:GOL:H2	1:C:428:HIS:HD2	1.69	0.57
1:B:357:LEU:HB2	3:B:1436:GOL:H11	1.86	0.57
1:C:357:LEU:HB2	3:C:1440:GOL:H32	1.87	0.56
1:A:178[B]:VAL:HG13	1:A:183:PRO:HA	1.86	0.55
1:C:430[B]:THR:HG22	7:C:4001:NGA:C1	2.37	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:163[B]:THR:HG21	1:C:203:LEU:HG	1.89	0.55
1:A:163[B]:THR:HG21	1:A:203:LEU:HG	1.90	0.54
1:B:430[B]:THR:HG22	7:B:4001:NGA:C1	2.38	0.53
1:B:163[B]:THR:HG21	1:B:203:LEU:HG	1.90	0.53
1:A:45:ALA:HB3	3:A:1436:GOL:H12	1.91	0.53
1:A:10:ALA:HA	3:A:1438:GOL:H11	1.93	0.51
6:A:1445:PEG:H42	1:C:111:GLN:HG2	1.94	0.50
1:C:167:TRP:HB2	1:C:178[B]:VAL:HG23	1.94	0.50
1:B:178[B]:VAL:HG11	1:B:194:VAL:HB	1.95	0.48
1:B:398:GLN:HE21	1:B:407:ARG:HE	1.62	0.48
1:C:178[A]:VAL:HG13	1:C:183:PRO:HA	1.95	0.48
1:B:129:ALA:HA	1:B:280:ILE:HA	1.95	0.48
1:A:80:ARG:NH2	1:B:267:THR:HG21	2.29	0.47
1:A:398:GLN:HE21	1:A:407:ARG:HE	1.62	0.47
1:C:27:ARG:HG3	1:C:166:VAL:HB	1.97	0.47
1:A:327:LYS:HB3	1:A:329:VAL:HG23	1.98	0.46
1:A:350:TYR:CZ	1:A:394:PRO:HB3	2.52	0.45
1:C:350:TYR:CZ	1:C:394:PRO:HB3	2.52	0.45
1:A:276:VAL:HG12	1:A:291:ILE:HG12	1.98	0.44
1:B:350:TYR:CZ	1:B:394:PRO:HB3	2.52	0.44
1:A:421:THR:HA	5:A:1444:ACT:H3	2.00	0.43
1:B:3:ALA:HB3	1:B:326:PHE:HB2	2.00	0.43
1:A:178[B]:VAL:HG11	1:A:194:VAL:HB	2.01	0.43
1:A:344:VAL:HG11	1:A:399:VAL:HG11	2.01	0.42
1:B:379:ALA:HB3	1:C:8:CYS:HB3	2.01	0.42
1:B:80:ARG:HH22	1:C:267:THR:HG21	1.85	0.42
1:C:129:ALA:HA	1:C:280:ILE:HA	2.01	0.42
1:A:129:ALA:HA	1:A:280:ILE:HA	2.02	0.42
1:B:344:VAL:HG11	1:B:399:VAL:HG11	2.02	0.41
1:A:167:TRP:HB2	1:A:178[A]:VAL:HG23	2.03	0.41
1:C:3:ALA:HB3	1:C:326:PHE:HB2	2.02	0.41

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	432/473 (91%)	423 (98%)	9 (2%)	0	100	100
1	B	433/473 (92%)	425 (98%)	8 (2%)	0	100	100
1	C	436/473 (92%)	427 (98%)	9 (2%)	0	100	100
All	All	1301/1419 (92%)	1275 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	352/376 (94%)	348 (99%)	4 (1%)	73	68
1	B	352/376 (94%)	348 (99%)	4 (1%)	73	68
1	C	354/376 (94%)	351 (99%)	3 (1%)	81	78
All	All	1058/1128 (94%)	1047 (99%)	11 (1%)	81	71

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	208	MET
1	A	429	THR
1	A	430[A]	THR
1	A	430[B]	THR
1	B	187	THR
1	B	208	MET
1	B	430[A]	THR
1	B	430[B]	THR
1	C	209	ASN
1	C	430[A]	THR
1	C	430[B]	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	398	GLN
1	B	111	GLN
1	B	298	HIS
1	B	369	HIS
1	B	398	GLN
1	C	398	GLN
1	C	428	HIS

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 ⓘ

Of 41 ligands modelled in this entry, 3 are monoatomic - leaving 38 for Mogul analysis.

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	GOL	B	1436	-	5,5,5	0.07	0	5,5,5	0.17	0
2	NAG	A	2001	1	14,14,15	0.26	0	17,19,21	0.46	0
3	GOL	C	1442	-	5,5,5	0.04	0	5,5,5	0.19	0
2	NAG	B	1001	1	14,14,15	0.29	0	17,19,21	0.55	0
5	ACT	B	1440	-	1,3,3	4.72	1 (100%)	0,3,3	0.00	-
2	NAG	C	2001	1	14,14,15	0.27	0	17,19,21	0.43	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
7	NGA	C	4001	1	14,14,15	0.29	0	17,19,21	0.79	1 (5%)
5	ACT	B	1443	-	1,3,3	4.66	1 (100%)	0,3,3	0.00	-
3	GOL	B	1438	-	5,5,5	0.05	0	5,5,5	0.22	0
7	NGA	B	4001	1	14,14,15	0.30	0	17,19,21	0.96	2 (11%)
3	GOL	A	1436	-	5,5,5	0.06	0	5,5,5	0.33	0
3	GOL	C	1438	-	5,5,5	0.03	0	5,5,5	0.07	0
3	GOL	B	1437	-	5,5,5	0.06	0	5,5,5	0.22	0
6	PEG	A	1445	-	6,6,6	0.13	0	5,5,5	0.10	0
7	NGA	A	3001	1	14,14,15	0.27	0	17,19,21	0.76	2 (11%)
3	GOL	A	1437	-	5,5,5	0.06	0	5,5,5	0.19	0
3	GOL	C	1441	-	5,5,5	0.05	0	5,5,5	0.21	0
3	GOL	A	1441	-	5,5,5	0.05	0	5,5,5	0.14	0
5	ACT	B	1441	-	1,3,3	4.72	1 (100%)	0,3,3	0.00	-
3	GOL	C	1440	-	5,5,5	0.08	0	5,5,5	0.16	0
2	NAG	B	2001	1	14,14,15	0.28	0	17,19,21	0.44	0
3	GOL	A	1438	-	5,5,5	0.07	0	5,5,5	0.10	0
7	NGA	A	4001	1	14,14,15	0.30	0	17,19,21	0.92	2 (11%)
3	GOL	A	1440	-	5,5,5	0.05	0	5,5,5	0.20	0
6	PEG	B	1444	-	6,6,6	0.08	0	5,5,5	0.03	0
7	NGA	B	3001	1	14,14,15	0.26	0	17,19,21	2.35	2 (11%)
2	NAG	A	1001	1	14,14,15	0.27	0	17,19,21	0.53	0
6	PEG	C	1445	-	6,6,6	0.21	0	5,5,5	0.06	0
3	GOL	A	1439	-	5,5,5	0.05	0	5,5,5	0.19	0
6	PEG	C	1446	-	6,6,6	0.08	0	5,5,5	0.03	0
7	NGA	C	3001	1	14,14,15	0.26	0	17,19,21	0.84	2 (11%)
5	ACT	A	1443	-	1,3,3	4.69	1 (100%)	0,3,3	0.00	-
5	ACT	A	1444	-	1,3,3	4.38	1 (100%)	0,3,3	0.00	-
5	ACT	C	1444	-	1,3,3	4.86	1 (100%)	0,3,3	0.00	-
5	ACT	B	1442	-	1,3,3	4.69	1 (100%)	0,3,3	0.00	-
3	GOL	C	1437	-	5,5,5	0.06	0	5,5,5	0.39	0
2	NAG	C	1001	1	14,14,15	0.26	0	17,19,21	0.55	0
3	GOL	C	1439	-	5,5,5	0.05	0	5,5,5	0.19	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
3	GOL	B	1436	-	-	2/4/4/4	-

Continued on next page...

Full wwPDB X-ray Structure Validation Report

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	C	1442	-	-	1/4/4/4	-
2	NAG	B	1001	1	-	0/6/23/26	0/1/1/1
7	NGA	C	4001	1	-	0/6/23/26	0/1/1/1
3	GOL	B	1438	-	-	0/4/4/4	-
7	NGA	B	4001	1	-	0/6/23/26	0/1/1/1
3	GOL	A	1436	-	-	2/4/4/4	-
3	GOL	C	1438	-	-	2/4/4/4	-
3	GOL	B	1437	-	-	0/4/4/4	-
6	PEG	A	1445	-	-	2/4/4/4	-
7	NGA	A	3001	1	-	0/6/23/26	0/1/1/1
3	GOL	A	1437	-	-	0/4/4/4	-
3	GOL	C	1441	-	-	0/4/4/4	-
3	GOL	A	1441	-	-	0/4/4/4	-
2	NAG	C	2001	1	-	0/6/23/26	0/1/1/1
2	NAG	B	2001	1	-	0/6/23/26	0/1/1/1
3	GOL	A	1438	-	-	1/4/4/4	-
7	NGA	A	4001	1	-	0/6/23/26	0/1/1/1
2	NAG	A	2001	1	-	0/6/23/26	0/1/1/1
6	PEG	B	1444	-	-	0/4/4/4	-
7	NGA	B	3001	1	-	2/6/23/26	0/1/1/1
2	NAG	A	1001	1	-	0/6/23/26	0/1/1/1
6	PEG	C	1445	-	-	2/4/4/4	-
3	GOL	A	1439	-	-	0/4/4/4	-
6	PEG	C	1446	-	-	1/4/4/4	-
7	NGA	C	3001	1	-	0/6/23/26	0/1/1/1
3	GOL	C	1440	-	-	3/4/4/4	-
3	GOL	A	1440	-	-	0/4/4/4	-
3	GOL	C	1437	-	-	2/4/4/4	-
2	NAG	C	1001	1	-	0/6/23/26	0/1/1/1
3	GOL	C	1439	-	-	3/4/4/4	-

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	1444	ACT	CH3-C	4.86	1.54	1.48
5	B	1440	ACT	CH3-C	4.72	1.54	1.48
5	B	1441	ACT	CH3-C	4.72	1.54	1.48
5	A	1443	ACT	CH3-C	4.69	1.54	1.48
5	B	1442	ACT	CH3-C	4.69	1.54	1.48
5	B	1443	ACT	CH3-C	4.66	1.54	1.48
5	A	1444	ACT	CH3-C	4.38	1.54	1.48

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	B	3001	NGA	C1-C2-N2	8.67	125.29	110.49
7	B	3001	NGA	O5-C1-C2	-3.91	105.12	111.29
7	A	4001	NGA	C1-C2-N2	2.63	114.98	110.49
7	B	4001	NGA	C1-C2-N2	2.59	114.92	110.49
7	B	4001	NGA	O5-C1-C2	-2.55	107.25	111.29
7	C	3001	NGA	O5-C1-C2	-2.36	107.57	111.29
7	A	4001	NGA	O5-C1-C2	-2.20	107.82	111.29
7	C	3001	NGA	C1-C2-N2	2.17	114.19	110.49
7	C	4001	NGA	C1-C2-N2	2.13	114.13	110.49
7	A	3001	NGA	O5-C1-C2	-2.01	108.12	111.29
7	A	3001	NGA	C1-C2-N2	2.00	113.91	110.49

There are no chirality outliers.

All (23) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	1436	GOL	O1-C1-C2-C3
3	A	1436	GOL	O1-C1-C2-O2
3	A	1436	GOL	O1-C1-C2-C3
7	B	3001	NGA	C8-C7-N2-C2
7	B	3001	NGA	O7-C7-N2-C2
6	A	1445	PEG	O1-C1-C2-O2
3	C	1438	GOL	O1-C1-C2-C3
3	C	1440	GOL	C1-C2-C3-O3
3	C	1437	GOL	C1-C2-C3-O3
3	C	1439	GOL	O1-C1-C2-C3
3	C	1439	GOL	C1-C2-C3-O3
3	C	1440	GOL	O2-C2-C3-O3
3	B	1436	GOL	O1-C1-C2-O2
6	C	1445	PEG	C1-C2-O2-C3
3	C	1437	GOL	O2-C2-C3-O3
6	C	1446	PEG	C4-C3-O2-C2
3	C	1440	GOL	O1-C1-C2-C3
3	C	1439	GOL	O1-C1-C2-O2
6	A	1445	PEG	C1-C2-O2-C3
3	C	1442	GOL	O1-C1-C2-C3
3	A	1438	GOL	O1-C1-C2-C3
6	C	1445	PEG	C4-C3-O2-C2
3	C	1438	GOL	O1-C1-C2-O2

There are no ring outliers.

9 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1436	GOL	1	0
7	C	4001	NGA	4	0
7	B	4001	NGA	4	0
3	A	1436	GOL	2	0
6	A	1445	PEG	1	0
3	C	1440	GOL	1	0
3	A	1438	GOL	1	0
7	A	4001	NGA	4	0
5	A	1444	ACT	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 (i)

6.1 Protein, DNA and RNA chains (i)

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	426/473 (90%)	-0.03	16 (3%)	40	35	18, 26, 48, 74	0
1	B	425/473 (89%)	-0.03	17 (4%)	38	32	18, 25, 46, 66	0
1	C	429/473 (90%)	-0.06	17 (3%)	38	32	17, 24, 43, 79	0
All	All	1280/1419 (90%)	-0.04	50 (3%)	39	33	17, 25, 46, 79	0

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	331	PRO	6.1
1	A	337	ALA	5.5
1	B	337	ALA	5.3
1	A	189	HIS	5.3
1	A	328	THR	4.6
1	B	189	HIS	4.5
1	A	330	ARG	4.5
1	A	329	VAL	4.5
1	C	332	VAL	4.4
1	C	213	ASN	4.4
1	C	337	ALA	4.1
1	C	336	ARG	3.8
1	A	135	THR	3.4
1	B	328	THR	3.3
1	C	331	PRO	3.3
1	A	188	PRO	3.2
1	A	2	GLU	3.1
1	C	436	ARG	3.1
1	B	135	THR	3.1
1	B	435	THR	3.0
1	A	365	GLY	3.0
1	C	2	GLU	3.0
1	B	2	GLU	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	338	LEU	2.8
1	A	404	GLU	2.8
1	B	338	LEU	2.8
1	B	213	ASN	2.8
1	A	134	PRO	2.7
1	C	134	PRO	2.7
1	B	221	GLY	2.7
1	C	435	THR	2.7
1	B	214	GLN	2.6
1	B	134	PRO	2.6
1	B	404	GLU	2.6
1	B	215	GLN	2.5
1	B	329	VAL	2.5
1	B	188	PRO	2.5
1	C	328	THR	2.5
1	C	135	THR	2.4
1	A	407	ARG	2.3
1	C	209	ASN	2.3
1	C	189	HIS	2.3
1	C	330	ARG	2.3
1	A	214	GLN	2.2
1	A	411	ARG	2.2
1	B	406	ASP	2.2
1	C	220	LEU	2.1
1	B	411	ARG	2.1
1	C	405	SER	2.1
1	C	215	GLN	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, 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
7	NGA	B	3001	14/15	0.37	0.33	41,42,44,45	14
6	PEG	A	1445	7/7	0.41	0.33	61,63,66,67	0
2	NAG	B	1001	14/15	0.56	0.36	67,69,71,71	0
2	NAG	A	1001	14/15	0.57	0.38	69,70,71,71	0
7	NGA	C	3001	14/15	0.58	0.23	46,46,47,47	14
7	NGA	A	4001	14/15	0.61	0.24	33,39,43,43	14
6	PEG	C	1446	7/7	0.63	0.23	70,72,75,75	0
7	NGA	C	4001	14/15	0.63	0.25	34,41,43,43	14
7	NGA	A	3001	14/15	0.64	0.24	44,45,46,47	14
3	GOL	A	1440	6/6	0.64	0.44	67,67,67,68	0
2	NAG	C	1001	14/15	0.66	0.32	61,65,67,68	0
3	GOL	B	1437	6/6	0.67	0.32	51,56,58,59	0
5	ACT	B	1442	4/4	0.67	0.23	58,58,59,62	0
3	GOL	A	1439	6/6	0.68	0.31	64,65,65,65	0
5	ACT	A	1444	4/4	0.68	0.33	76,76,76,78	0
5	ACT	C	1444	4/4	0.68	0.15	56,58,58,58	0
7	NGA	B	4001	14/15	0.69	0.26	29,35,39,41	14
5	ACT	B	1440	4/4	0.71	0.15	82,83,83,84	0
3	GOL	A	1438	6/6	0.72	0.25	63,65,66,67	0
3	GOL	C	1442	6/6	0.72	0.22	72,73,73,73	0
3	GOL	C	1441	6/6	0.74	0.20	69,69,70,70	0
3	GOL	B	1438	6/6	0.75	0.22	57,59,60,60	0
3	GOL	A	1441	6/6	0.77	0.15	75,76,77,77	0
3	GOL	C	1440	6/6	0.77	0.28	43,46,47,48	0
6	PEG	C	1445	7/7	0.78	0.21	27,36,42,43	0
5	ACT	B	1441	4/4	0.79	0.20	53,54,55,55	0
3	GOL	A	1437	6/6	0.79	0.26	39,46,47,49	0
5	ACT	A	1443	4/4	0.79	0.15	59,62,62,62	0
3	GOL	A	1436	6/6	0.80	0.31	37,41,42,45	0
6	PEG	B	1444	7/7	0.81	0.19	63,64,66,67	0
2	NAG	A	2001	14/15	0.85	0.14	39,42,47,48	0
3	GOL	B	1436	6/6	0.85	0.33	42,44,45,45	0
2	NAG	B	2001	14/15	0.85	0.16	40,43,47,49	0
2	NAG	C	2001	14/15	0.87	0.12	37,40,44,45	0
3	GOL	C	1437	6/6	0.88	0.11	25,28,29,30	0
5	ACT	B	1443	4/4	0.88	0.16	72,72,72,73	0
3	GOL	C	1439	6/6	0.89	0.18	41,41,42,43	0
3	GOL	C	1438	6/6	0.92	0.12	30,31,31,31	0
4	NA	B	1439	1/1	0.97	0.08	29,29,29,29	0
4	NA	C	1443	1/1	0.98	0.11	32,32,32,32	0
4	NA	A	1442	1/1	0.99	0.14	31,31,31,31	0

6.5 Other polymers (i)

There are no such residues in this entry.