



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 12:27 AM BST

PDB ID : 4B3V  
Title : Crystal structure of the Rubella virus glycoprotein E1 in its post-fusion form crystallized in presence of 20mM of Calcium Acetate  
Authors : Vaney, M.C.; DuBois, R.M.; Tortorici, M.A.; Rey, F.A.  
Deposited on : 2012-07-26  
Resolution : 1.98 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](mailto: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 ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) 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

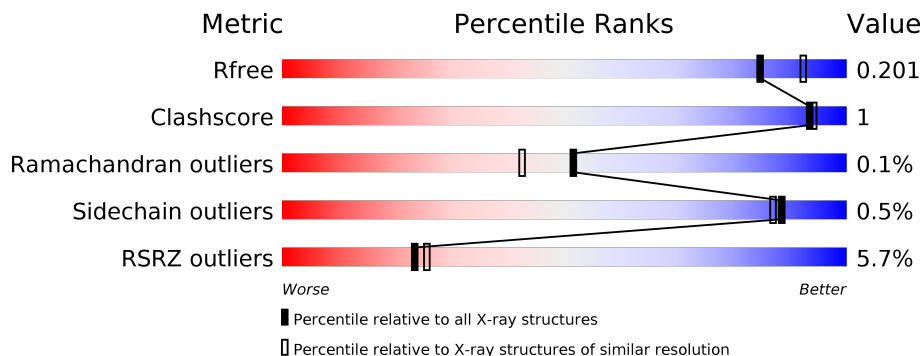
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.98 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	<div style="display: flex; align-items: center;"> <div style="width: 5%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 88%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 9%; height: 10px; background-color: grey; margin-right: 5px;"></div> </div> <p style="text-align: center;">5%      88%      9%</p>
1	B	473	<div style="display: flex; align-items: center;"> <div style="width: 5%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 87%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 9%; height: 10px; background-color: grey; margin-right: 5px;"></div> </div> <p style="text-align: center;">5%      87%      9%</p>
1	C	473	<div style="display: flex; align-items: center;"> <div style="width: 5%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 88%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 9%; height: 10px; background-color: grey; margin-right: 5px;"></div> </div> <p style="text-align: center;">5%      88%      9%</p>

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	NGA	A	1437	-	-	-	X
3	NGA	C	1438	-	-	-	X

## 2 Entry composition [i](#)

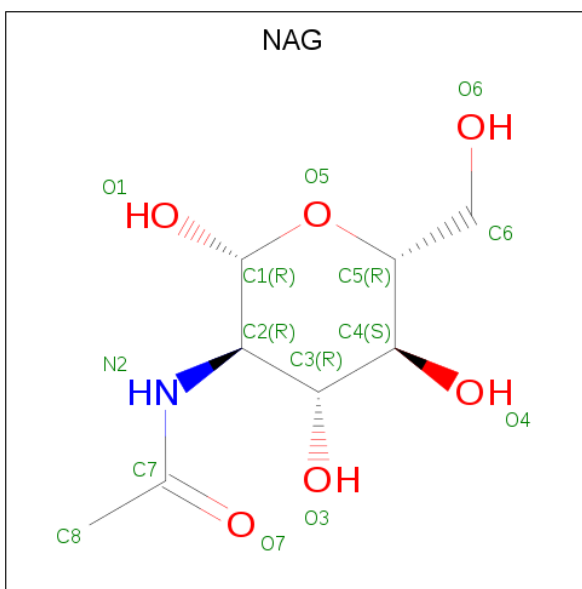
There are 10 unique types of molecules in this entry. The entry contains 11200 atoms, of which 0 are hydrogens and 0 are deuteriums.

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
			Total	C	N	O	S			
1	A	431	Total	C	N	O	S	0	10	1
			3299	2087	576	613	23			
1	B	432	Total	C	N	O	S	0	5	2
			3270	2067	575	605	23			
1	C	432	Total	C	N	O	S	0	4	2
			3252	2056	571	602	23			

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



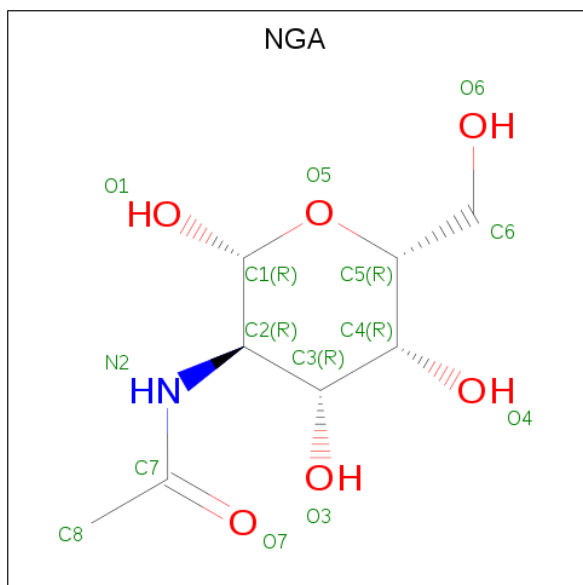
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

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

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-galactopyranose (three-letter code: NGA) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



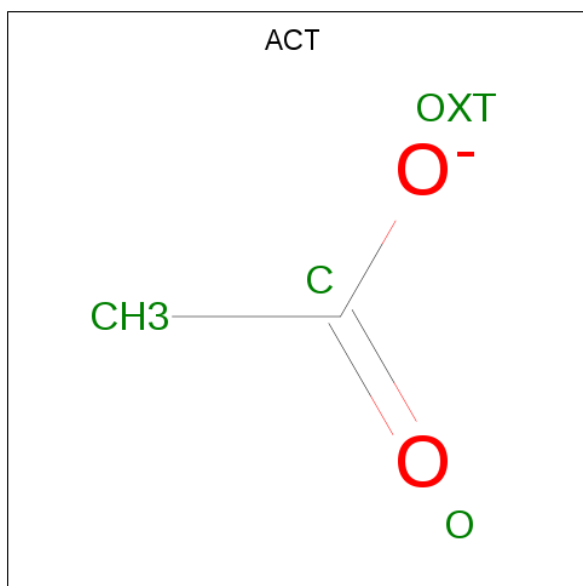
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 12 6 6	0	1
4	C	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0

*Continued on next page...*

Continued from previous page...

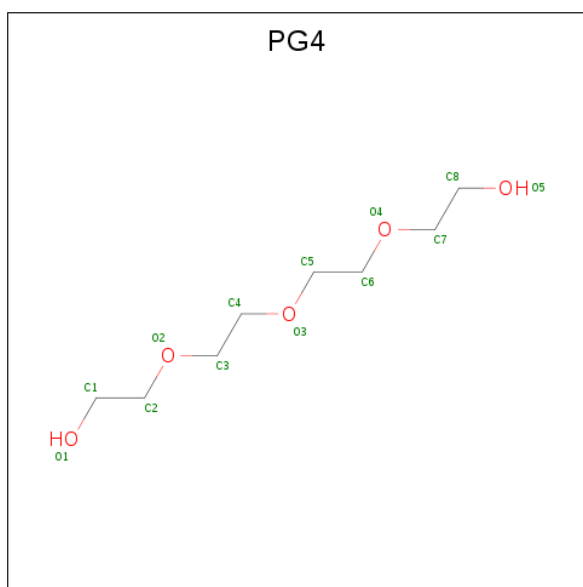
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	C	1	6	3	3	0	0

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



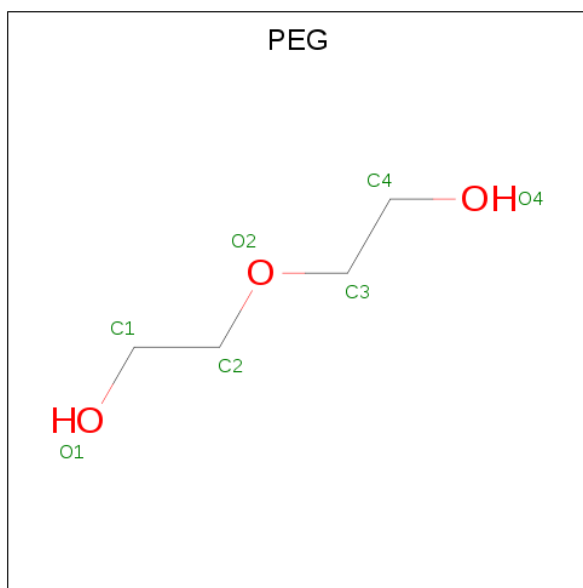
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	A	1	4	2	2	0	0
5	B	1	4	2	2	0	0
5	B	1	4	2	2	0	0
5	C	1	4	2	2	0	0
5	C	1	4	2	2	0	0

- Molecule 6 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			8	5	3		

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			7	4	3		
7	B	1	Total	C	O	0	0
			7	4	3		
7	C	1	Total	C	O	0	0
			7	4	3		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	C	1	Total C O 7 4 3	0	0

- Molecule 8 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total Na 1 1	0	0

- Molecule 9 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	B	1	Total Ca 1 1	0	0
9	A	1	Total Ca 1 1	0	0
9	C	1	Total Ca 1 1	0	0

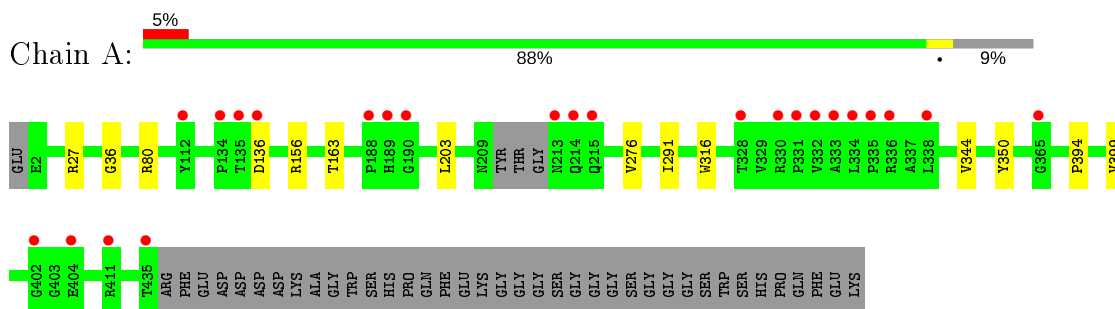
- Molecule 10 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	419	Total O 419 419	0	0
10	B	378	Total O 378 378	0	0
10	C	334	Total O 334 334	0	0

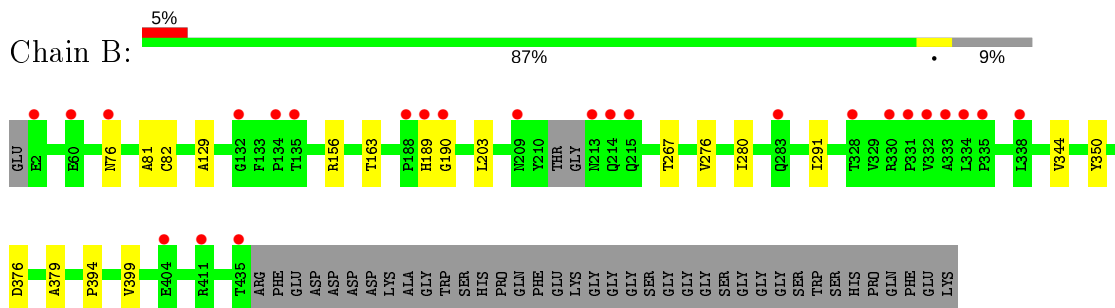
### 3 Residue-property plots [i](#)

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.

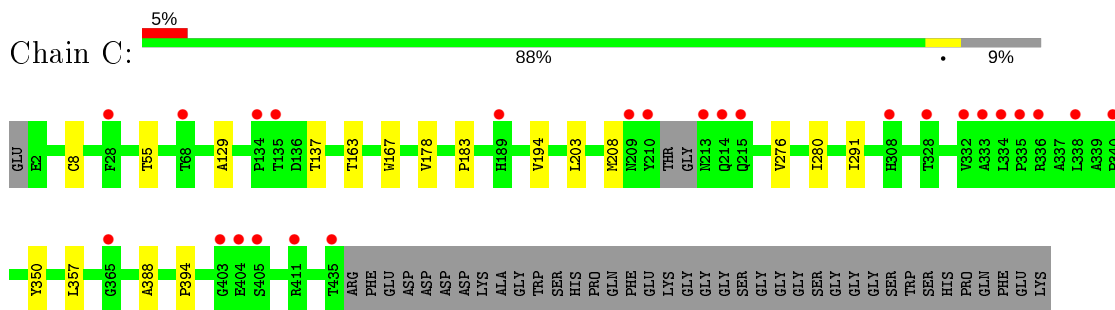
- Molecule 1: E1 ENVELOPE GLYCOPROTEIN



- Molecule 1: E1 ENVELOPE GLYCOPROTEIN



- Molecule 1: E1 ENVELOPE GLYCOPROTEIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	121.27Å 126.16Å 129.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.99 – 1.98 28.78 – 1.98	Depositor EDS
% Data completeness (in resolution range)	97.3 (28.99-1.98) 97.3 (28.78-1.98)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.89 (at 1.98Å)	Xtrriage
Refinement program	BUSTER 2.11.2	Depositor
R, $R_{free}$	0.185 , 0.198 0.188 , 0.201	Depositor DCC
$R_{free}$ test set	6752 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	28.1	Xtrriage
Anisotropy	0.051	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 47.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.016 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11200	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: GOL, NAG, NA, CA, PG4, 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	0.35	0/3421	0.58	0/4703
1	B	0.36	0/3391	0.59	0/4659
1	C	0.35	0/3372	0.58	0/4635
All	All	0.35	0/10184	0.58	0/13997

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	3299	0	3136	7	0
1	B	3270	0	3109	10	0
1	C	3252	0	3096	12	0
2	A	14	0	13	0	0
2	B	14	0	13	0	0
2	C	28	0	26	0	0
3	A	14	0	13	0	0
3	C	14	0	13	0	0
4	A	24	0	32	0	0
4	B	42	0	56	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	30	0	40	1	0
5	A	12	0	9	0	0
5	B	8	0	6	0	0
5	C	8	0	6	1	0
6	A	8	0	9	2	0
7	A	7	0	10	0	0
7	B	7	0	10	0	0
7	C	14	0	20	1	0
8	A	1	0	0	0	0
9	A	1	0	0	0	0
9	B	1	0	0	0	0
9	C	1	0	0	0	0
10	A	419	0	0	0	0
10	B	378	0	0	0	0
10	C	334	0	0	0	0
All	All	11200	0	9617	27	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 1.

All (27) 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:55:THR:H	7:C:1448:PEG:H21	1.61	0.65
1:B:189:HIS:N	1:B:190:GLY:HA2	2.19	0.57
1:C:178[A]:VAL:HG13	1:C:183:PRO:HA	1.87	0.56
6:A:1447:PG4:H11	1:C:388:ALA:HB3	1.89	0.54
1:B:276:VAL:HG12	1:B:291:ILE:HG12	1.93	0.50
1:A:80:ARG:HH22	1:B:267:THR:HG21	1.76	0.50
1:A:316:TRP:HA	6:A:1447:PG4:H12	1.95	0.49
1:C:167:TRP:HB2	1:C:178[B]:VAL:HG23	1.95	0.48
1:B:344:VAL:HG11	1:B:399:VAL:HG11	1.95	0.48
1:B:81:ALA:HA	4:B:1442[A]:GOL:H2	1.95	0.47
1:A:36:GLY:HA2	1:A:156:ARG:HH21	1.78	0.47
1:C:137:THR:H	5:C:1444:ACT:H3	1.81	0.46
1:C:178[A]:VAL:HG11	1:C:194:VAL:HB	1.98	0.46
1:C:276:VAL:HG12	1:C:291:ILE:HG12	1.96	0.46
1:B:350:TYR:CZ	1:B:394:PRO:HB3	2.51	0.45
1:A:350:TYR:CZ	1:A:394:PRO:HB3	2.51	0.45
1:B:129:ALA:HA	1:B:280:ILE:HA	1.97	0.45
1:A:344:VAL:HG11	1:A:399:VAL:HG11	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:276:VAL:HG12	1:A:291:ILE:HG12	1.99	0.45
1:C:350:TYR:CZ	1:C:394:PRO:HB3	2.51	0.45
1:C:163[B]:THR:HG21	1:C:203:LEU:HG	2.01	0.43
1:C:129:ALA:HA	1:C:280:ILE:HA	2.01	0.42
1:B:379:ALA:HB3	1:C:8:CYS:HB3	2.01	0.42
1:A:163[B]:THR:HG21	1:A:203:LEU:HG	2.01	0.42
1:B:163[B]:THR:HG21	1:B:203:LEU:HG	2.01	0.42
1:C:357:LEU:HB2	4:C:1441:GOL:H12	2.03	0.41
1:B:82:CYS:H	4:B:1442[B]:GOL:H32	1.86	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	437/473 (92%)	429 (98%)	6 (1%)	2 (0%)	29	16
1	B	433/473 (92%)	425 (98%)	8 (2%)	0	100	100
1	C	432/473 (91%)	425 (98%)	7 (2%)	0	100	100
All	All	1302/1419 (92%)	1279 (98%)	21 (2%)	2 (0%)	51	38

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	136[A]	ASP
1	A	136[B]	ASP

### 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	355/376 (94%)	354 (100%)	1 (0%)	92	92
1	B	350/376 (93%)	347 (99%)	3 (1%)	78	77
1	C	349/376 (93%)	348 (100%)	1 (0%)	92	92
All	All	1054/1128 (93%)	1049 (100%)	5 (0%)	88	87

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	ARG
1	B	76	ASN
1	B	156	ARG
1	B	376	ASP
1	C	208	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	398	GLN
1	B	111	GLN
1	B	398	GLN
1	C	111	GLN
1	C	398	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

Of 38 ligands modelled in this entry, 4 are monoatomic - leaving 34 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
4	GOL	C	1442	-	5,5,5	0.05	0	5,5,5	0.21	0
3	NGA	A	1437	1	14,14,15	0.25	0	17,19,21	0.67	0
2	NAG	B	1436	1	14,14,15	0.25	0	17,19,21	0.43	0
4	GOL	A	1439	-	5,5,5	0.06	0	5,5,5	0.20	0
5	ACT	A	1443	-	1,3,3	4.64	1 (100%)	0,3,3	0.00	-
5	ACT	A	1445	-	1,3,3	4.32	1 (100%)	0,3,3	0.00	-
4	GOL	C	1441	-	5,5,5	0.05	0	5,5,5	0.29	0
4	GOL	B	1441	-	5,5,5	0.06	0	5,5,5	0.21	0
7	PEG	C	1447	-	6,6,6	0.11	0	5,5,5	0.08	0
5	ACT	C	1445	-	1,3,3	4.74	1 (100%)	0,3,3	0.00	-
4	GOL	A	1440	-	5,5,5	0.05	0	5,5,5	0.15	0
4	GOL	C	1443	-	5,5,5	0.04	0	5,5,5	0.18	0
4	GOL	B	1442[A]	-	5,5,5	0.06	0	5,5,5	0.30	0
7	PEG	B	1446	-	6,6,6	0.13	0	5,5,5	0.05	0
2	NAG	C	1437	1	14,14,15	0.26	0	17,19,21	0.46	0
4	GOL	B	1442[B]	-	5,5,5	0.04	0	5,5,5	0.32	0
7	PEG	C	1448	-	6,6,6	0.12	0	5,5,5	0.03	0
4	GOL	B	1438	-	5,5,5	0.06	0	5,5,5	0.28	0
6	PG4	A	1447	-	7,7,12	0.40	0	6,6,11	0.81	0
7	PEG	A	1449	-	6,6,6	0.09	0	5,5,5	0.03	0
2	NAG	C	1436	1	14,14,15	0.26	0	17,19,21	0.42	0
5	ACT	B	1443	9	1,3,3	4.75	1 (100%)	0,3,3	0.00	-
5	ACT	B	1444	-	1,3,3	4.75	1 (100%)	0,3,3	0.00	-
4	GOL	B	1437	-	5,5,5	0.06	0	5,5,5	0.22	0
4	GOL	A	1441	-	5,5,5	0.04	0	5,5,5	0.20	0
5	ACT	C	1444	9	1,3,3	4.65	1 (100%)	0,3,3	0.00	-
4	GOL	A	1438	-	5,5,5	0.04	0	5,5,5	0.26	0
4	GOL	C	1440	-	5,5,5	0.05	0	5,5,5	0.14	0
3	NGA	C	1438	1	14,14,15	0.30	0	17,19,21	0.83	2 (11%)
4	GOL	B	1439	-	5,5,5	0.04	0	5,5,5	0.15	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GOL	C	1439	-	5,5,5	0.05	0	5,5,5	0.35	0
2	NAG	A	1436	1	14,14,15	0.27	0	17,19,21	0.48	0
5	ACT	A	1444	-	1,3,3	3.86	1 (100%)	0,3,3	0.00	-
4	GOL	B	1440	-	5,5,5	0.06	0	5,5,5	0.18	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
4	GOL	C	1442	-	-	0/4/4/4	-
3	NGA	A	1437	1	-	0/6/23/26	0/1/1/1
2	NAG	B	1436	1	-	0/6/23/26	0/1/1/1
4	GOL	A	1439	-	-	0/4/4/4	-
7	PEG	A	1449	-	-	1/4/4/4	-
4	GOL	C	1441	-	-	1/4/4/4	-
4	GOL	B	1441	-	-	0/4/4/4	-
7	PEG	C	1447	-	-	2/4/4/4	-
4	GOL	A	1440	-	-	0/4/4/4	-
4	GOL	C	1443	-	-	1/4/4/4	-
4	GOL	B	1442[A]	-	-	0/4/4/4	-
7	PEG	B	1446	-	-	1/4/4/4	-
2	NAG	C	1437	1	-	0/6/23/26	0/1/1/1
4	GOL	B	1442[B]	-	-	1/4/4/4	-
4	GOL	B	1438	-	-	2/4/4/4	-
6	PG4	A	1447	-	-	4/5/5/10	-
2	NAG	A	1436	1	-	0/6/23/26	0/1/1/1
2	NAG	C	1436	1	-	0/6/23/26	0/1/1/1
4	GOL	B	1437	-	-	0/4/4/4	-
4	GOL	A	1441	-	-	0/4/4/4	-
7	PEG	C	1448	-	-	2/4/4/4	-
4	GOL	A	1438	-	-	0/4/4/4	-
4	GOL	C	1440	-	-	0/4/4/4	-
3	NGA	C	1438	1	-	0/6/23/26	0/1/1/1
4	GOL	B	1439	-	-	0/4/4/4	-
4	GOL	C	1439	-	-	0/4/4/4	-
4	GOL	B	1440	-	-	0/4/4/4	-

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1443	ACT	CH3-C	4.75	1.54	1.48
5	B	1444	ACT	CH3-C	4.75	1.54	1.48
5	C	1445	ACT	CH3-C	4.74	1.54	1.48
5	C	1444	ACT	CH3-C	4.65	1.54	1.48
5	A	1443	ACT	CH3-C	4.64	1.54	1.48
5	A	1445	ACT	CH3-C	4.32	1.54	1.48
5	A	1444	ACT	CH3-C	3.86	1.53	1.48

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1438	NGA	C1-O5-C5	2.30	115.31	112.19
3	C	1438	NGA	C1-C2-N2	2.04	113.97	110.49

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	1446	PEG	O1-C1-C2-O2
4	C	1441	GOL	O1-C1-C2-C3
4	B	1438	GOL	O1-C1-C2-C3
7	C	1447	PEG	C4-C3-O2-C2
6	A	1447	PG4	C4-C3-O2-C2
7	C	1448	PEG	C4-C3-O2-C2
7	A	1449	PEG	C1-C2-O2-C3
7	C	1447	PEG	C1-C2-O2-C3
7	C	1448	PEG	O1-C1-C2-O2
4	C	1443	GOL	O1-C1-C2-C3
6	A	1447	PG4	C1-C2-O2-C3
4	B	1438	GOL	O1-C1-C2-O2
4	B	1442[B]	GOL	C1-C2-C3-O3
6	A	1447	PG4	O2-C3-C4-O3
6	A	1447	PG4	C3-C4-O3-C5

There are no ring outliers.

6 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	1441	GOL	1	0
4	B	1442[A]	GOL	1	0
4	B	1442[B]	GOL	1	0
7	C	1448	PEG	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1447	PG4	2	0
5	C	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
1	A	431/473 (91%)	0.24	24 (5%) 24 26	19, 29, 53, 81	0
1	B	432/473 (91%)	0.19	25 (5%) 23 25	19, 28, 51, 74	0
1	C	432/473 (91%)	0.11	25 (5%) 23 25	19, 28, 49, 66	1 (0%)
All	All	1295/1419 (91%)	0.18	74 (5%) 23 25	19, 28, 51, 81	1 (0%)

All (74) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	333	ALA	7.1
1	A	189	HIS	7.0
1	A	335	PRO	6.6
1	B	333	ALA	6.6
1	C	333	ALA	6.4
1	B	335	PRO	5.8
1	A	135[A]	THR	5.5
1	B	330	ARG	5.0
1	C	335	PRO	4.9
1	C	210	TYR	4.9
1	B	189	HIS	4.9
1	C	213	ASN	4.9
1	A	188	PRO	4.7
1	A	134[A]	PRO	4.5
1	B	213	ASN	4.4
1	A	334	LEU	4.3
1	B	135	THR	4.2
1	C	134	PRO	4.2
1	B	332	VAL	4.1
1	C	209	ASN	3.9
1	B	134	PRO	3.9
1	A	331	PRO	3.8
1	C	135	THR	3.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	332	VAL	3.6
1	C	334	LEU	3.5
1	A	213	ASN	3.5
1	A	365	GLY	3.5
1	B	331	PRO	3.4
1	B	328	THR	3.3
1	C	338	LEU	3.2
1	C	189	HIS	3.1
1	A	336	ARG	3.1
1	A	330	ARG	3.1
1	C	214	GLN	3.0
1	B	190	GLY	2.9
1	B	215	GLN	2.9
1	A	328	THR	2.9
1	C	336	ARG	2.9
1	A	411	ARG	2.9
1	A	435	THR	2.9
1	C	404	GLU	2.8
1	C	332	VAL	2.8
1	B	334	LEU	2.8
1	C	308	HIS	2.7
1	C	28	PHE	2.7
1	A	112	TYR	2.6
1	A	404	GLU	2.6
1	C	435	THR	2.5
1	A	214	GLN	2.5
1	C	411	ARG	2.5
1	A	338	LEU	2.5
1	B	209	ASN	2.4
1	A	215	GLN	2.4
1	C	365	GLY	2.4
1	B	435	THR	2.4
1	A	190	GLY	2.4
1	A	402	GLY	2.4
1	B	404	GLU	2.3
1	B	188	PRO	2.3
1	B	411	ARG	2.3
1	B	283	GLN	2.3
1	B	60	GLU	2.3
1	C	403	GLY	2.3
1	B	214	GLN	2.3
1	C	328	THR	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	340	PRO	2.2
1	B	2	GLU	2.2
1	B	338	LEU	2.1
1	C	405	SER	2.1
1	B	132	GLY	2.1
1	A	136[A]	ASP	2.1
1	B	76	ASN	2.1
1	C	215	GLN	2.0
1	C	68	THR	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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	NGA	A	1437	14/15	0.29	0.53	76,78,80,80	0
5	ACT	B	1444	4/4	0.46	0.32	77,77,77,78	0
5	ACT	C	1445	4/4	0.51	0.28	80,80,80,80	0
2	NAG	C	1437	14/15	0.58	0.37	73,74,76,76	0
6	PG4	A	1447	8/13	0.59	0.28	65,66,67,67	0
5	ACT	C	1444	4/4	0.60	0.29	71,71,71,71	0
5	ACT	B	1443	4/4	0.62	0.36	65,65,65,65	0
4	GOL	C	1442	6/6	0.67	0.34	69,70,70,71	0
5	ACT	A	1443	4/4	0.67	0.30	47,48,48,48	4
3	NGA	C	1438	14/15	0.67	0.49	74,77,80,80	0
7	PEG	C	1447	7/7	0.70	0.26	82,83,83,83	0
7	PEG	B	1446	7/7	0.70	0.25	45,49,52,52	0
4	GOL	A	1440	6/6	0.72	0.29	66,67,67,67	0
4	GOL	B	1442[B]	6/6	0.74	0.30	26,29,29,29	6

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	GOL	B	1442[A]	6/6	0.74	0.30	42,45,45,46	6
4	GOL	B	1441	6/6	0.75	0.17	69,70,70,70	0
7	PEG	A	1449	7/7	0.75	0.23	62,63,63,63	0
4	GOL	C	1443	6/6	0.76	0.21	74,74,75,75	0
4	GOL	B	1438	6/6	0.76	0.29	50,50,52,52	0
4	GOL	C	1441	6/6	0.77	0.29	57,58,58,59	0
7	PEG	C	1448	7/7	0.78	0.23	41,45,48,48	0
4	GOL	B	1439	6/6	0.78	0.35	51,54,55,56	0
5	ACT	A	1444	4/4	0.78	0.22	62,62,62,62	0
4	GOL	B	1440	6/6	0.78	0.23	60,61,61,61	0
5	ACT	A	1445	4/4	0.81	0.35	84,84,84,84	0
4	GOL	A	1438	6/6	0.82	0.22	49,51,51,51	0
4	GOL	B	1437	6/6	0.84	0.19	33,34,35,35	0
4	GOL	A	1441	6/6	0.84	0.24	70,71,71,71	0
4	GOL	A	1439	6/6	0.85	0.16	59,61,61,61	0
4	GOL	C	1440	6/6	0.85	0.14	39,39,39,40	0
2	NAG	B	1436	14/15	0.86	0.22	45,47,50,51	0
2	NAG	C	1436	14/15	0.88	0.19	43,47,50,50	0
2	NAG	A	1436	14/15	0.93	0.21	42,45,47,48	0
4	GOL	C	1439	6/6	0.93	0.15	33,36,37,39	0
9	CA	A	1451	1/1	0.97	0.05	33,33,33,33	1
9	CA	C	1446	1/1	0.98	0.08	36,36,36,36	0
8	NA	A	1450	1/1	0.99	0.04	41,41,41,41	1
9	CA	B	1445	1/1	0.99	0.08	36,36,36,36	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.