wwPDB X-ray Structure Validation Summary Report (i)

Dec 18, 2023 – 09:06 am GMT

PDB ID	:	4ADG
Title	:	Crystal structure of the Rubella virus envelope Glycoprotein E1 in post-fusion form (crystal form II)
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	:	2.18 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

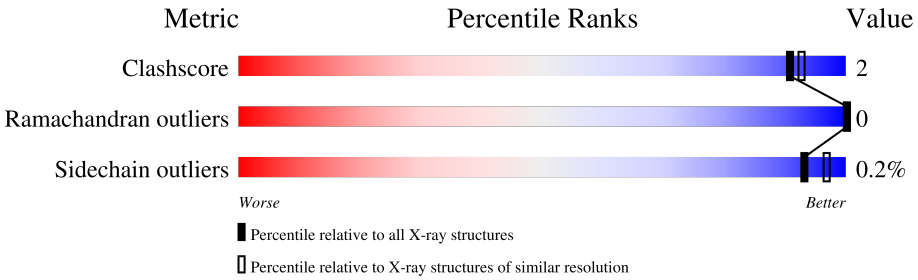
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.18 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))
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)

2 Entry composition

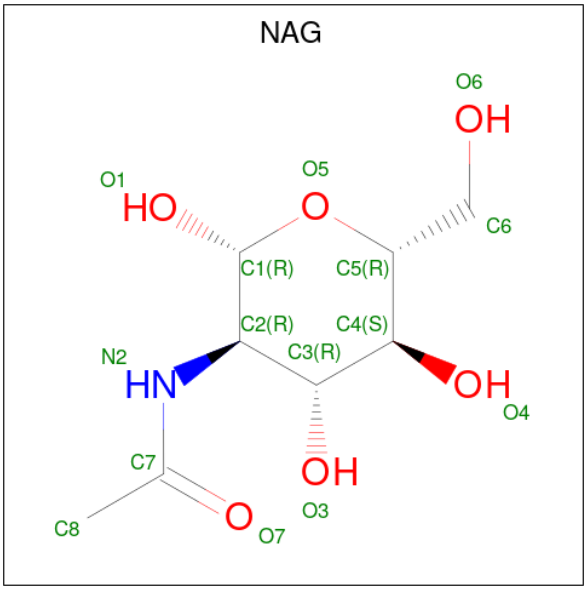
There are 9 unique types of molecules in this entry. The entry contains 11144 atoms, of which 0 are hydrogens and 0 are deuteriums.

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	432	Total	C	N	O	S	0	6	1
			3278	2073	574	608	23			
1	B	433	Total	C	N	O	S	0	7	2
			3282	2078	572	609	23			
1	C	433	Total	C	N	O	S	0	4	1
			3268	2066	573	606	23			

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



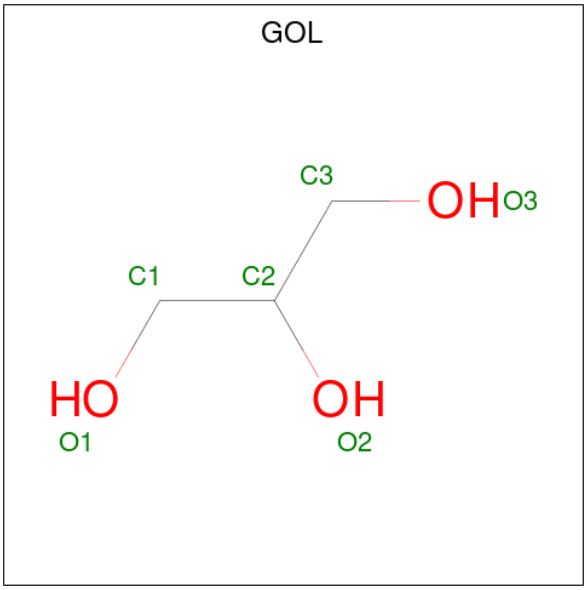
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		
2	C	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total C N O 14 8 1 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 C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0

Continued on next page...

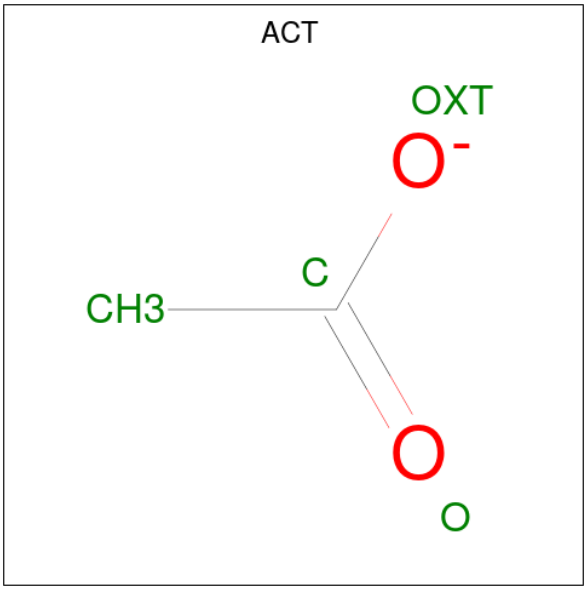
Continued from previous page...

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

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

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

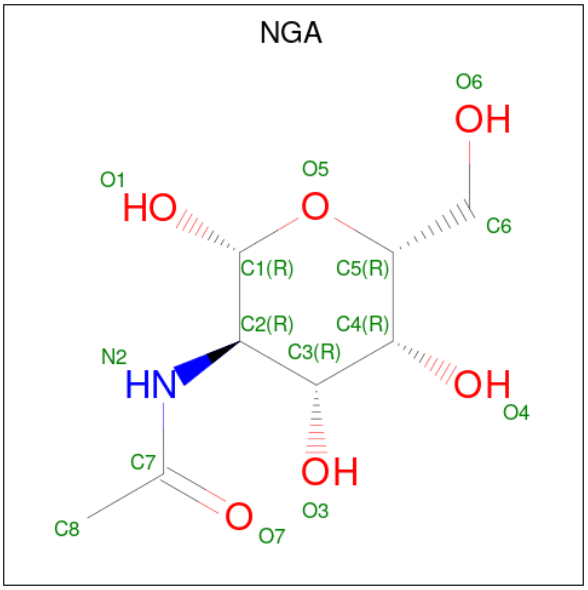


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

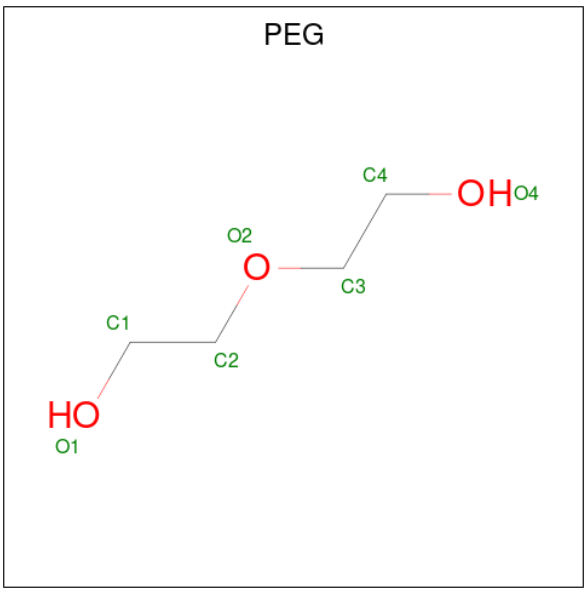
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Cl 1 1	0	0

- 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 C N O 14 8 1 5	0	0

- Molecule 8 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



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

- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	421	Total O 421 421	0	0
9	B	386	Total O 386 386	0	0
9	C	344	Total O 344 344	0	0

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	121.68Å 126.55Å 130.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.18	Depositor
% Data completeness (in resolution range)	98.9 (20.00-2.18)	Depositor
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 2.19Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, R_{free}	0.178 , 0.193	Depositor
Wilson B-factor (Å^2)	25.9	Xtriage
Anisotropy	0.061	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.016 for -h,l,k	Xtriage
Total number of atoms	11144	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.22% 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.

4 Model quality

4.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, CA, ACT, CL, GOL, 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	0.36	0/3396	0.60	0/4665
1	B	0.36	0/3402	0.58	0/4677
1	C	0.35	0/3383	0.59	0/4650
All	All	0.36	0/10181	0.59	0/13992

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

4.2 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	3278	0	3110	8	0
1	B	3282	0	3103	12	0
1	C	3268	0	3097	15	0
2	A	14	0	13	0	0
2	B	14	0	13	0	0
2	C	28	0	26	0	0
3	A	6	0	8	0	0
3	B	18	0	24	2	0
3	C	48	0	64	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	4	0	3	0	0
5	B	4	0	3	0	0
5	C	4	0	3	1	0
6	A	1	0	0	0	0
7	A	14	0	13	1	0
8	C	7	0	10	0	0
9	A	421	0	0	0	0
9	B	386	0	0	0	0
9	C	344	0	0	0	0
All	All	11144	0	9490	33	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 2.

The worst 5 of 33 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:357:LEU:HB2	3:C:1439:GOL:H12	1.75	0.67
1:B:357:LEU:HB2	3:B:1436:GOL:H31	1.80	0.62
1:C:178[A]:VAL:HG13	1:C:183:PRO:HA	1.84	0.59
1:A:432:VAL:HG22	7:A:4001:NGA:H62	1.93	0.51
1:B:276:VAL:HG12	1:B:291:ILE:HG12	1.93	0.51

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	434/473 (92%)	424 (98%)	10 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	436/473 (92%)	428 (98%)	8 (2%)	0	100	100
1	C	433/473 (92%)	425 (98%)	8 (2%)	0	100	100
All	All	1303/1419 (92%)	1277 (98%)	26 (2%)	0	100	100

There are no Ramachandran outliers to report.

4.3.2 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	350/376 (93%)	348 (99%)	2 (1%)	86	92
1	B	350/376 (93%)	350 (100%)	0	100	100
1	C	349/376 (93%)	349 (100%)	0	100	100
All	All	1049/1128 (93%)	1047 (100%)	2 (0%)	93	97

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	GLU
1	A	430	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	428	HIS
1	C	398	GLN
1	C	189	HIS
1	B	428	HIS
1	C	369	HIS

4.3.3 RNA (i)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

Of 25 ligands modelled in this entry, 4 are monoatomic - leaving 21 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	$\#\|Z\| > 2$
5	ACT	A	1438	4	3,3,3	1.09	0	3,3,3	0.92	0
2	NAG	C	1001	1	14,14,15	0.25	0	17,19,21	0.48	0
5	ACT	B	1440	4	3,3,3	1.10	0	3,3,3	0.91	0
3	GOL	A	1436	-	5,5,5	0.04	0	5,5,5	0.17	0
3	GOL	B	1438	-	5,5,5	0.04	0	5,5,5	0.14	0
3	GOL	C	1442	-	5,5,5	0.05	0	5,5,5	0.13	0
3	GOL	C	1443	-	5,5,5	0.03	0	5,5,5	0.12	0
3	GOL	C	1438	-	5,5,5	0.04	0	5,5,5	0.21	0
2	NAG	A	1001	1	14,14,15	0.27	0	17,19,21	0.48	0
7	NGA	A	4001	1	14,14,15	0.31	0	17,19,21	1.08	1 (5%)
3	GOL	C	1441	-	5,5,5	0.04	0	5,5,5	0.13	0
8	PEG	C	1445	-	6,6,6	0.09	0	5,5,5	0.03	0
3	GOL	B	1437	-	5,5,5	0.05	0	5,5,5	0.17	0
2	NAG	B	1001	1	14,14,15	0.27	0	17,19,21	0.42	0
3	GOL	C	1439	-	5,5,5	0.05	0	5,5,5	0.19	0
3	GOL	B	1436	-	5,5,5	0.05	0	5,5,5	0.19	0
3	GOL	C	1437	-	5,5,5	0.05	0	5,5,5	0.13	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
2	NAG	C	1001	1	-	0/6/23/26	0/1/1/1
3	GOL	A	1436	-	-	0/4/4/4	-
3	GOL	B	1438	-	-	0/4/4/4	-
3	GOL	C	1442	-	-	0/4/4/4	-
3	GOL	C	1443	-	-	0/4/4/4	-
3	GOL	C	1438	-	-	0/4/4/4	-
2	NAG	A	1001	1	-	0/6/23/26	0/1/1/1
7	NGA	A	4001	1	-	4/6/23/26	0/1/1/1
3	GOL	C	1441	-	-	0/4/4/4	-
8	PEG	C	1445	-	-	2/4/4/4	-
3	GOL	B	1437	-	-	0/4/4/4	-
2	NAG	B	1001	1	-	0/6/23/26	0/1/1/1
3	GOL	C	1439	-	-	0/4/4/4	-
3	GOL	B	1436	-	-	2/4/4/4	-
3	GOL	C	1437	-	-	0/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	4001	NGA	C1-C2-N2	4.03	117.37	110.49

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	4001	NGA	C8-C7-N2-C2
7	A	4001	NGA	O7-C7-N2-C2
7	A	4001	NGA	O5-C5-C6-O6
7	A	4001	NGA	C4-C5-C6-O6
8	C	1445	PEG	C1-C2-O2-C3

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	4001	NGA	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1437	GOL	1	0
3	C	1439	GOL	1	0
3	B	1436	GOL	1	0

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

5.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

5.4 Ligands

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers

EDS failed to run properly - this section is therefore empty.