



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 12:30 PM BST

PDB ID : 5VOC  
Title : Crystal structure of HCMV Pentamer in complex with neutralizing antibody 8I21 - Low resolution dataset for initial phasing by SAD  
Authors : Malito, E.; Chandramouli, S.  
Deposited on : 2017-05-02  
Resolution : 3.99 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](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:

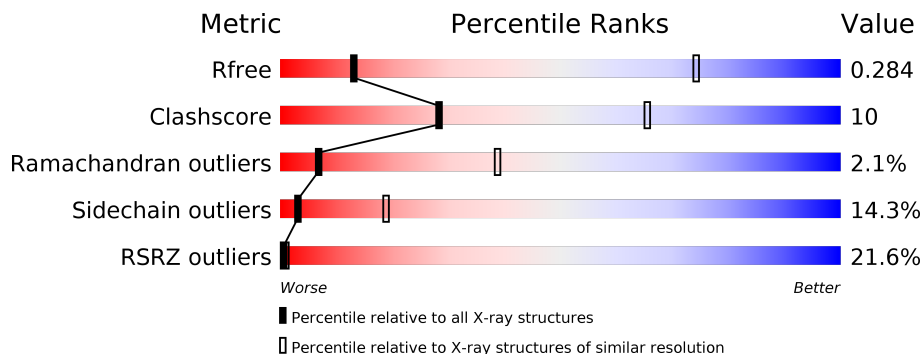
MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
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 3.99 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.



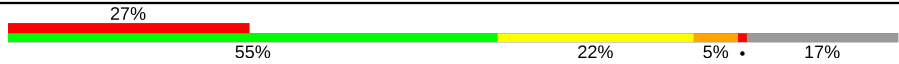

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1087 (4.30-3.70)
Clashscore	141614	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
Sidechain outliers	138945	1099 (4.30-3.70)
RSRZ outliers	127900	1028 (4.34-3.66)

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	725	
2	B	278	
3	C	171	
4	D	252	
5	E	129	
6	H	289	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
7	L	235	
8	F	3	

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
9	NAG	A	803	-	-	-	X

## 2 Entry composition i

There are 9 unique types of molecules in this entry. The entry contains 13701 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 Envelope glycoprotein H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	662	5321	3402	904	990	25	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	716	GLY	-	expression tag	UNP Q6SW67
A	717	SER	-	expression tag	UNP Q6SW67
A	718	GLY	-	expression tag	UNP Q6SW67
A	719	SER	-	expression tag	UNP Q6SW67
A	720	HIS	-	expression tag	UNP Q6SW67
A	721	HIS	-	expression tag	UNP Q6SW67
A	722	HIS	-	expression tag	UNP Q6SW67
A	723	HIS	-	expression tag	UNP Q6SW67
A	724	HIS	-	expression tag	UNP Q6SW67
A	725	HIS	-	expression tag	UNP Q6SW67

- Molecule 2 is a protein called Envelope glycoprotein L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	237	1867	1188	326	345	8	0	0	0

- Molecule 3 is a protein called Envelope glycoprotein UL128.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	135	1090	683	197	199	11	0	0	0

- Molecule 4 is a protein called Envelope glycoprotein UL130.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	170	1388	887	243	250	8	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	215	GLY	-	expression tag	UNP F5HCP3
D	216	SER	-	expression tag	UNP F5HCP3
D	217	GLU	-	expression tag	UNP F5HCP3
D	218	ASN	-	expression tag	UNP F5HCP3
D	219	LEU	-	expression tag	UNP F5HCP3
D	220	TYR	-	expression tag	UNP F5HCP3
D	221	PHE	-	expression tag	UNP F5HCP3
D	222	GLN	-	expression tag	UNP F5HCP3
D	223	ALA	-	expression tag	UNP F5HCP3
D	224	GLY	-	expression tag	UNP F5HCP3
D	225	TRP	-	expression tag	UNP F5HCP3
D	226	SER	-	expression tag	UNP F5HCP3
D	227	HIS	-	expression tag	UNP F5HCP3
D	228	PRO	-	expression tag	UNP F5HCP3
D	229	GLN	-	expression tag	UNP F5HCP3
D	230	PHE	-	expression tag	UNP F5HCP3
D	231	GLU	-	expression tag	UNP F5HCP3
D	232	LYS	-	expression tag	UNP F5HCP3
D	233	GLY	-	expression tag	UNP F5HCP3
D	234	GLY	-	expression tag	UNP F5HCP3
D	235	GLY	-	expression tag	UNP F5HCP3
D	236	SER	-	expression tag	UNP F5HCP3
D	237	GLY	-	expression tag	UNP F5HCP3
D	238	GLY	-	expression tag	UNP F5HCP3
D	239	GLY	-	expression tag	UNP F5HCP3
D	240	SER	-	expression tag	UNP F5HCP3
D	241	GLY	-	expression tag	UNP F5HCP3
D	242	GLY	-	expression tag	UNP F5HCP3
D	243	GLY	-	expression tag	UNP F5HCP3
D	244	SER	-	expression tag	UNP F5HCP3
D	245	TRP	-	expression tag	UNP F5HCP3
D	246	SER	-	expression tag	UNP F5HCP3
D	247	HIS	-	expression tag	UNP F5HCP3
D	248	PRO	-	expression tag	UNP F5HCP3
D	249	GLN	-	expression tag	UNP F5HCP3
D	250	PHE	-	expression tag	UNP F5HCP3
D	251	GLU	-	expression tag	UNP F5HCP3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	252	LYS	-	expression tag	UNP F5HCP3

- Molecule 5 is a protein called Envelope glycoprotein UL131A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	108	893	557	168	166	2	0	0	0

- Molecule 6 is a protein called Fab 8I21 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	H	200	1540	978	265	290	7	0	0	0

There are 43 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	228	SER	-	expression tag	UNP S6B291
H	229	SER	-	expression tag	UNP S6B291
H	230	GLY	-	expression tag	UNP S6B291
H	231	LEU	-	expression tag	UNP S6B291
H	232	GLU	-	expression tag	UNP S6B291
H	233	VAL	-	expression tag	UNP S6B291
H	234	LEU	-	expression tag	UNP S6B291
H	235	PHE	-	expression tag	UNP S6B291
H	236	GLN	-	expression tag	UNP S6B291
H	237	GLY	-	expression tag	UNP S6B291
H	238	PRO	-	expression tag	UNP S6B291
H	239	LEU	-	expression tag	UNP S6B291
H	240	GLY	-	expression tag	UNP S6B291
H	241	SER	-	expression tag	UNP S6B291
H	242	ALA	-	expression tag	UNP S6B291
H	243	TRP	-	expression tag	UNP S6B291
H	244	SER	-	expression tag	UNP S6B291
H	245	HIS	-	expression tag	UNP S6B291
H	246	PRO	-	expression tag	UNP S6B291
H	247	GLN	-	expression tag	UNP S6B291
H	248	PHE	-	expression tag	UNP S6B291
H	249	GLU	-	expression tag	UNP S6B291
H	250	LYS	-	expression tag	UNP S6B291
H	251	GLY	-	expression tag	UNP S6B291
H	252	GLY	-	expression tag	UNP S6B291

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	253	GLY	-	expression tag	UNP S6B291
H	254	SER	-	expression tag	UNP S6B291
H	255	GLY	-	expression tag	UNP S6B291
H	256	GLY	-	expression tag	UNP S6B291
H	257	GLY	-	expression tag	UNP S6B291
H	258	SER	-	expression tag	UNP S6B291
H	259	GLY	-	expression tag	UNP S6B291
H	260	GLY	-	expression tag	UNP S6B291
H	261	GLY	-	expression tag	UNP S6B291
H	262	SER	-	expression tag	UNP S6B291
H	263	TRP	-	expression tag	UNP S6B291
H	264	SER	-	expression tag	UNP S6B291
H	265	HIS	-	expression tag	UNP S6B291
H	266	PRO	-	expression tag	UNP S6B291
H	267	GLN	-	expression tag	UNP S6B291
H	268	PHE	-	expression tag	UNP S6B291
H	269	GLU	-	expression tag	UNP S6B291
H	270	LYS	-	expression tag	UNP S6B291

- Molecule 7 is a protein called Fab 8I21 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	L	194	1493	939	254	295	5	0	0	0

There are 19 discrepancies between the modelled and reference sequences:

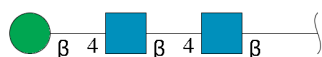
Chain	Residue	Modelled	Actual	Comment	Reference
L	-17	THR	-	conflict	PDB ?
L	-14	GLU	-	conflict	PDB ?
L	2	THR	-	conflict	PDB ?
L	17	GLY	-	conflict	PDB ?
L	30	GLY	-	conflict	PDB ?
L	31	ILE	-	conflict	PDB ?
L	56	SER	-	conflict	PDB ?
L	58	PHE	-	conflict	PDB ?
L	76	THR	-	conflict	PDB ?
L	93	ASP	-	conflict	PDB ?
L	96	PRO	-	linker	PDB ?
L	97	TRP	-	linker	PDB ?
L	98	THR	-	linker	PDB ?
L	99	PHE	-	linker	PDB ?

*Continued on next page...*

Continued from previous page...

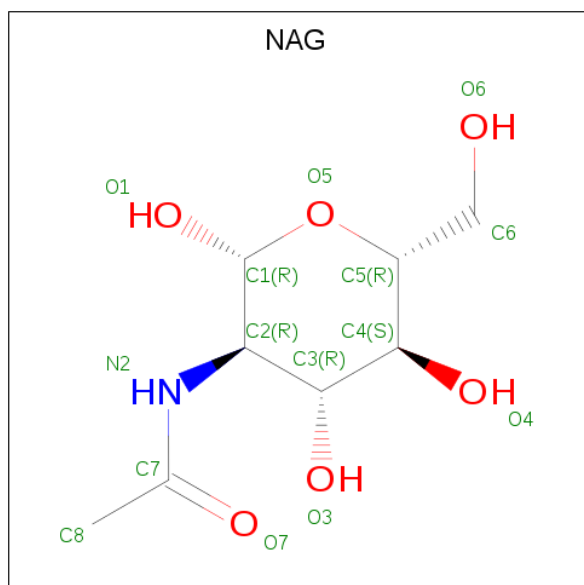
Chain	Residue	Modelled	Actual	Comment	Reference
L	100	GLY	-	linker	PDB ?
L	101	GLN	-	linker	PDB ?
L	102	GLY	-	linker	PDB ?
L	103	THR	-	linker	PDB ?
L	104	LYS	-	linker	PDB ?

- Molecule 8 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
8	F	3	39	22	2	15	0	0	0

- Molecule 9 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

Continued on next page...



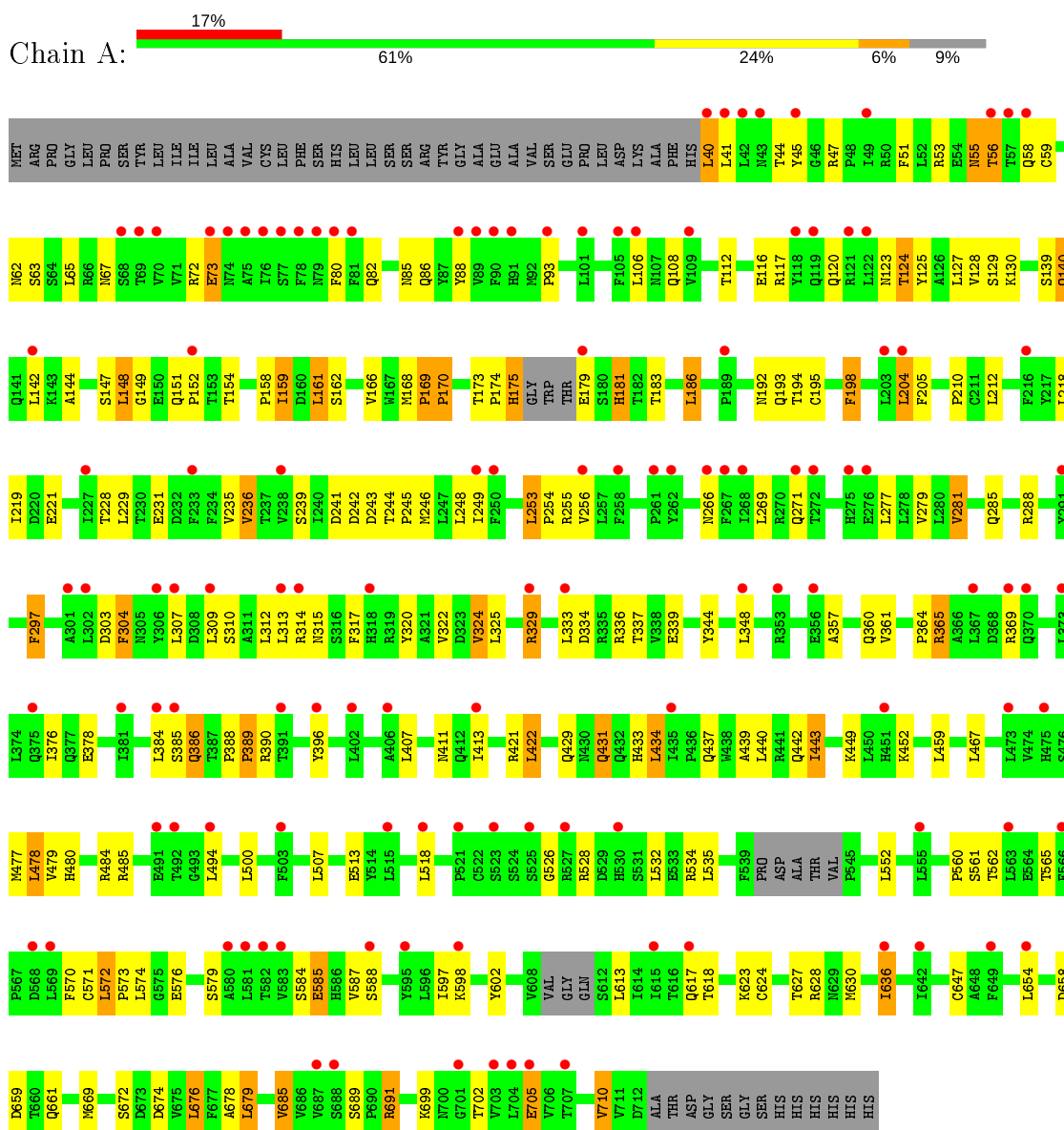
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>				<b>ZeroOcc</b>	<b>AltConf</b>
9	A	1	Total	C	N	O	0	0
			14	8	1	5		
9	D	1	Total	C	N	O	0	0
			14	8	1	5		
9	D	1	Total	C	N	O	0	0
			14	8	1	5		

### 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.

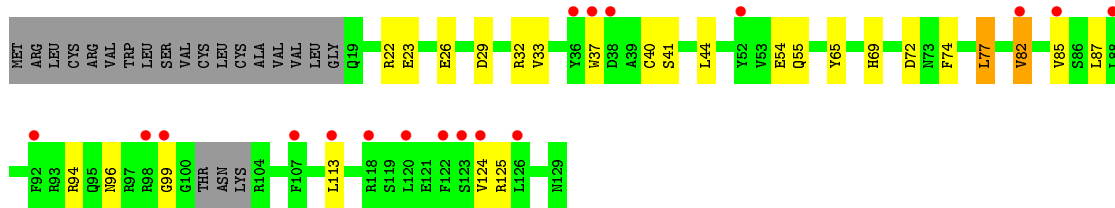
- Molecule 1: Envelope glycoprotein H



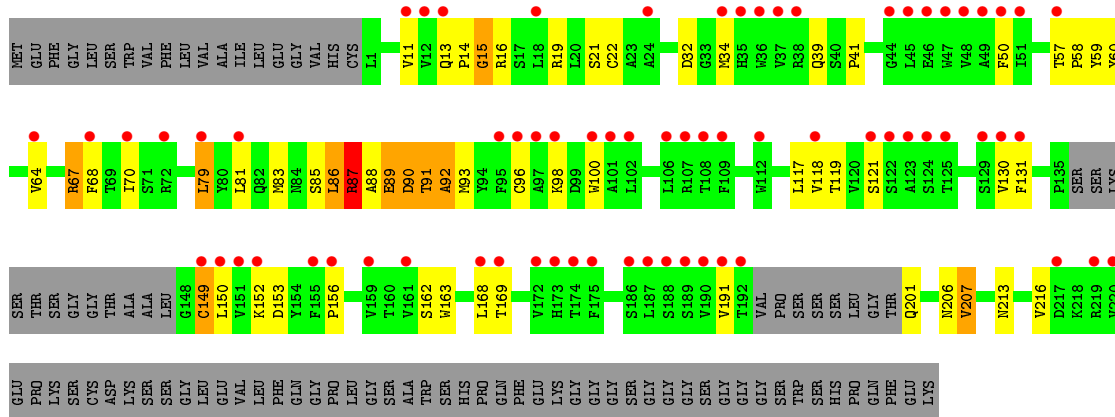
- Molecule 2: Envelope glycoprotein L



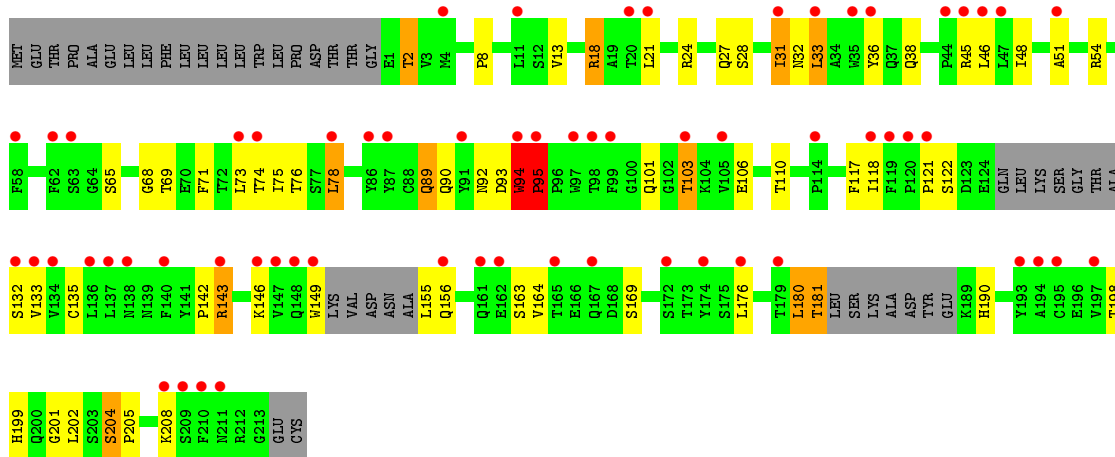




• Molecule 6: Fab 8I21 heavy chain



• Molecule 7: Fab 8I21 light chain



• Molecule 8: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	124.42Å 146.53Å 190.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.06 – 3.99 116.25 – 3.99	Depositor EDS
% Data completeness (in resolution range)	99.9 (35.06-3.99) 100.0 (116.25-3.99)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.53 (at 4.01Å)	Xtrriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.259 , 0.276 0.268 , 0.284	Depositor DCC
$R_{free}$ test set	1525 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	153.8	Xtrriage
Anisotropy	0.575	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 127.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	13701	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	121.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.78% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, 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	0.55	0/5444	0.76	1/7415 (0.0%)
2	B	0.56	0/1913	0.76	0/2612
3	C	0.49	0/1112	0.78	0/1503
4	D	0.51	0/1427	0.79	1/1939 (0.1%)
5	E	0.47	0/912	0.65	0/1233
6	H	0.39	0/1580	0.73	1/2148 (0.0%)
7	L	0.43	0/1528	0.76	1/2076 (0.0%)
All	All	0.51	0/13916	0.76	4/18926 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
4	D	136	ASN	C-N-CA	6.22	137.26	121.70
1	A	174	PRO	C-N-CA	5.58	135.65	121.70
7	L	95	PRO	CA-C-N	5.22	131.72	117.10
6	H	15	GLY	C-N-CA	5.19	134.68	121.70

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5321	0	5274	114	0
2	B	1867	0	1859	47	0
3	C	1090	0	1075	28	0
4	D	1388	0	1360	31	0
5	E	893	0	851	15	0
6	H	1540	0	1495	35	0
7	L	1493	0	1437	31	0
8	F	39	0	34	1	0
9	A	42	0	39	2	0
9	D	28	0	26	2	0
All	All	13701	0	13450	266	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 10.

The worst 5 of 266 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:169:PRO:HB2	1:A:170:PRO:HD3	1.39	1.03
7:L:13:VAL:HG11	7:L:78:LEU:HD13	1.38	1.02
2:B:121:ALA:O	2:B:125:LEU:HD13	1.64	0.97
7:L:94:TRP:HB3	7:L:95:PRO:CD	1.96	0.94
4:D:117:ARG:HG2	9:D:301:NAG:H83	1.52	0.92

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	654/725 (90%)	599 (92%)	43 (7%)	12 (2%)	<b>8</b> 41
2	B	235/278 (84%)	215 (92%)	17 (7%)	3 (1%)	<b>12</b> 48

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	133/171 (78%)	108 (81%)	20 (15%)	5 (4%)	3	27
4	D	168/252 (67%)	161 (96%)	6 (4%)	1 (1%)	25	63
5	E	104/129 (81%)	98 (94%)	5 (5%)	1 (1%)	15	53
6	H	194/289 (67%)	170 (88%)	17 (9%)	7 (4%)	3	28
7	L	186/235 (79%)	169 (91%)	11 (6%)	6 (3%)	4	31
All	All	1674/2079 (80%)	1520 (91%)	119 (7%)	35 (2%)	7	39

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	56	THR
1	A	169	PRO
1	A	170	PRO
1	A	297	PHE
1	A	389	PRO

### 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	604/655 (92%)	517 (86%)	87 (14%)	3	18
2	B	204/238 (86%)	176 (86%)	28 (14%)	3	20
3	C	122/154 (79%)	101 (83%)	21 (17%)	2	13
4	D	156/222 (70%)	138 (88%)	18 (12%)	5	25
5	E	95/114 (83%)	83 (87%)	12 (13%)	4	22
6	H	171/239 (72%)	152 (89%)	19 (11%)	6	26
7	L	167/202 (83%)	135 (81%)	32 (19%)	1	9
All	All	1519/1824 (83%)	1302 (86%)	217 (14%)	3	19

5 of 217 residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
2	B	119	LEU
3	C	70	THR
7	L	94	TRP
2	B	158	LEU
2	B	235	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	177	HIS
3	C	115	ASN
7	L	156	GLN
2	B	203	ASN
4	D	123	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](#)

3 monosaccharides 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
8	NAG	F	1	8,5	14,14,15	0.54	0	17,19,21	2.26	5 (29%)
8	NAG	F	2	8	14,14,15	0.72	0	17,19,21	1.86	3 (17%)
8	BMA	F	3	8	11,11,12	1.05	0	15,15,17	1.23	3 (20%)

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
8	NAG	F	1	8,5	-	4/6/23/26	0/1/1/1
8	NAG	F	2	8	-	1/6/23/26	0/1/1/1
8	BMA	F	3	8	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	F	1	NAG	C1-O5-C5	5.54	119.69	112.19
8	F	1	NAG	C3-C4-C5	4.61	118.46	110.24
8	F	2	NAG	C2-N2-C7	4.52	129.34	122.90
8	F	1	NAG	C1-C2-N2	4.07	117.44	110.49
8	F	2	NAG	C1-O5-C5	3.77	117.29	112.19

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	F	2	NAG	C1-C2-N2-C7
8	F	1	NAG	O5-C5-C6-O6
8	F	1	NAG	C4-C5-C6-O6
8	F	1	NAG	C1-C2-N2-C7
8	F	1	NAG	C3-C2-N2-C7

There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	F	2	NAG	1	0
8	F	1	NAG	1	0

## 5.6 Ligand geometry [i](#)

5 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
9	NAG	A	803	1	14,14,15	0.48	0	17,19,21	1.03	2 (11%)
9	NAG	A	802	1	14,14,15	0.70	0	17,19,21	1.74	4 (23%)
9	NAG	D	301	4	14,14,15	0.61	0	17,19,21	1.35	3 (17%)
9	NAG	D	302	4	14,14,15	0.45	0	17,19,21	0.75	1 (5%)
9	NAG	A	801	1	14,14,15	0.96	0	17,19,21	1.98	5 (29%)

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
9	NAG	A	803	1	-	2/6/23/26	0/1/1/1
9	NAG	A	802	1	-	3/6/23/26	0/1/1/1
9	NAG	D	301	4	-	3/6/23/26	0/1/1/1
9	NAG	D	302	4	-	0/6/23/26	0/1/1/1
9	NAG	A	801	1	-	5/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	802	NAG	C3-C4-C5	4.72	118.67	110.24
9	A	801	NAG	C1-C2-N2	4.72	118.55	110.49
9	D	301	NAG	O5-C1-C2	3.71	117.15	111.29
9	A	801	NAG	C2-N2-C7	3.51	127.90	122.90
9	A	801	NAG	C1-O5-C5	3.49	116.92	112.19

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	A	801	NAG	O7-C7-N2-C2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
9	A	801	NAG	C8-C7-N2-C2
9	A	801	NAG	O5-C5-C6-O6
9	A	801	NAG	C1-C2-N2-C7
9	A	802	NAG	C8-C7-N2-C2

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	803	NAG	1	0
9	D	301	NAG	2	0
9	A	801	NAG	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	662/725 (91%)	1.10	124 (18%) <b>1</b> <b>1</b>	50, 93, 133, 300	0
2	B	237/278 (85%)	0.90	30 (12%) <b>3</b> <b>4</b>	63, 97, 148, 166	0
3	C	135/171 (78%)	0.88	19 (14%) <b>2</b> <b>3</b>	64, 101, 153, 192	2 (1%)
4	D	170/252 (67%)	1.33	45 (26%) <b>0</b> <b>0</b>	55, 76, 119, 182	0
5	E	108/129 (83%)	1.17	18 (16%) <b>1</b> <b>2</b>	55, 89, 141, 155	0
6	H	200/289 (69%)	1.99	70 (35%) <b>0</b> <b>0</b>	143, 203, 258, 285	0
7	L	194/235 (82%)	1.58	63 (32%) <b>0</b> <b>0</b>	131, 172, 278, 300	0
All	All	1706/2079 (82%)	1.24	369 (21%) <b>0</b> <b>1</b>	50, 101, 248, 300	2 (0%)

The worst 5 of 369 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	H	188	SER	15.3
6	H	124	SER	12.0
6	H	189	SER	11.4
6	H	123	ALA	10.7
6	H	187	LEU	9.8

### 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](#)

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( $\text{\AA}^2$ )	Q<0.9
8	BMA	F	3	11/12	0.70	0.29	50,50,50,50	0
8	NAG	F	2	14/15	0.85	0.18	50,50,50,50	0
8	NAG	F	1	14/15	0.91	0.24	50,50,50,50	0

## 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( $\text{\AA}^2$ )	Q<0.9
9	NAG	A	803	14/15	0.78	0.46	50,50,50,50	0
9	NAG	A	802	14/15	0.78	0.31	50,50,50,50	0
9	NAG	A	801	14/15	0.84	0.22	50,50,50,50	0
9	NAG	D	301	14/15	0.85	0.27	50,50,50,50	0
9	NAG	D	302	14/15	0.88	0.22	50,50,50,50	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.