



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Mar 2, 2017 – 11:30 am GMT

PDB ID : 3J4P  
EMDB ID: : EMD-5681  
Title : Electron Microscopy Analysis of a Disaccharide Analog complex Reveals Receptor Interactions of Adeno-Associated Virus  
Authors : Xie, Q.; Chapman, M.S.  
Deposited on : 2013-09-10  
Resolution : 4.80 Å(reported)  
Based on PDB ID : 3J1Q

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc29047

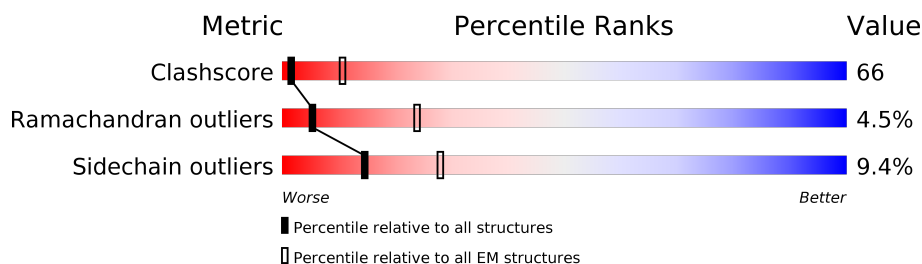
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.80 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)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	517	

## 2 Entry composition [i](#)

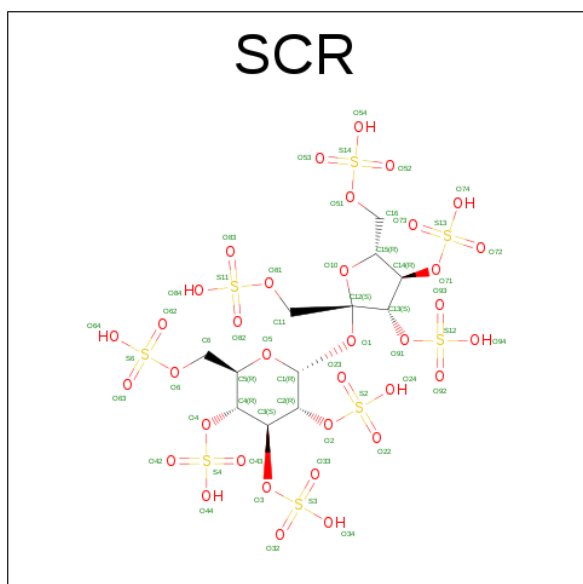
There are 5 unique types of molecules in this entry. The entry contains 4243 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Capsid protein VP1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	517	Total	C	N	O	S	6	0
			4185	2635	728	808	14		

- Molecule 2 is SUGAR (SUCROSE OCTASULFATE) (three-letter code: SCR) (formula:  $C_{12}H_{22}O_{35}S_8$ ).



Mol	Chain	Residues	Atoms				AltConf
2	A	1	Total	C	O	S	0
			55	12	35	8	

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		AltConf
3	A	1	Total	Na	0
			1	1	

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
4	A	1	Total	Mg	0
			1	1	

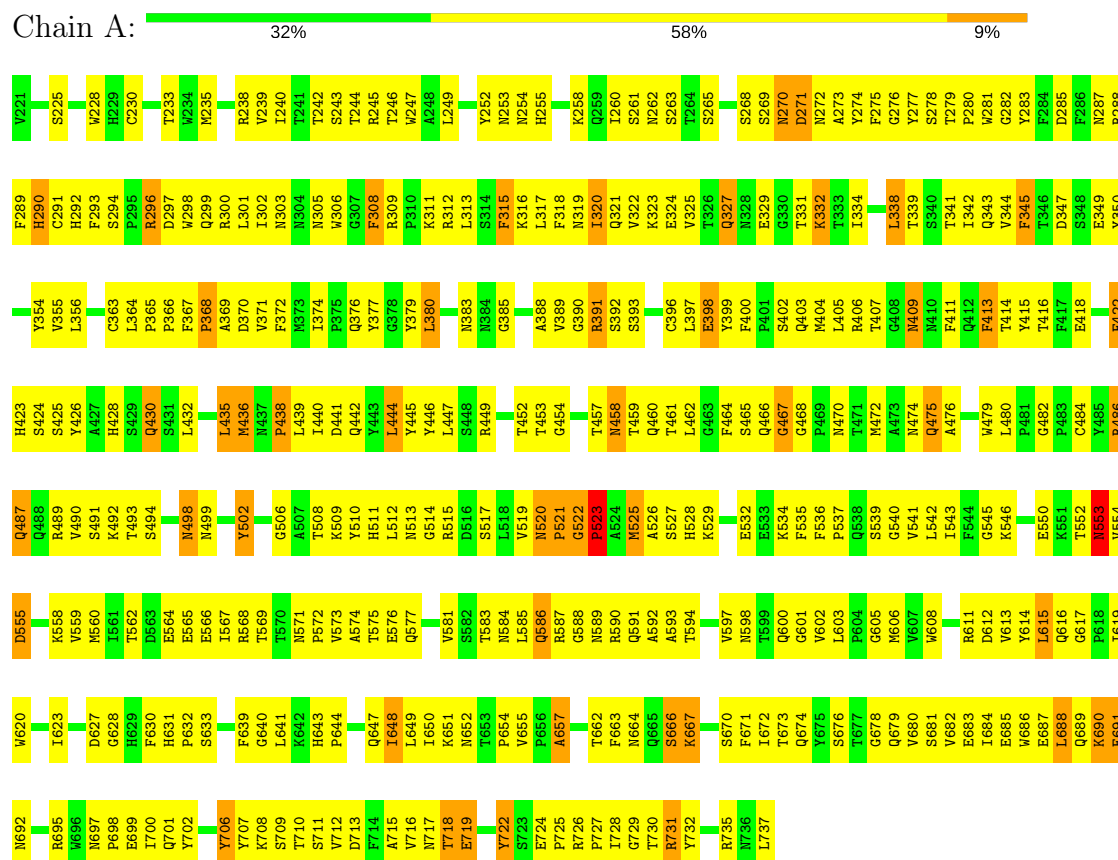
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		AltConf
5	A	1	Total	O	0
			1	1	

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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: Capsid protein VP1



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	45000	Depositor
Resolution determination method	FSC at 0.143 cut-off	Depositor
CTF correction method	each particle	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	120	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	15	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	120000	Depositor
Image detector	Gatan Ultrascan 4Kx4K CCD	Depositor

## 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: NA, MG, SCR

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	$\# Z  > 2$	RMSZ	$\# Z  > 2$
1	A	0.36	0/4308	0.63	2/5870 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	520	ASN	C-N-CD	-6.75	105.75	120.60
1	A	522	GLY	C-N-CD	-5.99	107.43	120.60

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	4185	0	3946	534	0
2	A	55	0	14	8	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	1	0	0	0	0
All	All	4243	0	3960	540	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 66.

The worst 5 of 540 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:320:ILE:HD12	1:A:320:ILE:H	1.23	1.04
1:A:262:ASN:HD21	1:A:273:ALA:HA	1.27	0.99
1:A:355:VAL:H	1:A:647:GLN:NE2	1.63	0.97
1:A:520:ASN:HA	1:A:522:GLY:H	1.26	0.96
1:A:355:VAL:N	1:A:647:GLN:HE22	1.65	0.95

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 PDB entries followed by that with respect to all EM entries.

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	521/517 (101%)	432 (83%)	65 (12%)	24 (5%)	<b>3</b> 29

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	270	ASN
1	A	458	ASN
1	A	521	PRO
1	A	523	PRO
1	A	553	ASN

### 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 PDB entries followed by that with respect to all EM entries.

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	463/458 (101%)	420 (91%)	43 (9%)	10 39

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	436	MET
1	A	487	GLN
1	A	699	GLU
1	A	444	LEU
1	A	460	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	475	GLN
1	A	487	GLN
1	A	674	GLN
1	A	409	ASN
1	A	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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	$\# Z  > 2$
2	SCR	A	801	3	55,56,56	1.94	15 (27%)	67,92,92	2.80	10 (14%)

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	SCR	A	801	3	-	1/49/88/88	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	SCR	O81-C11	-4.53	1.38	1.45
2	A	801	SCR	O91-S12	-3.69	1.46	1.56
2	A	801	SCR	O3-S3	-3.62	1.46	1.56
2	A	801	SCR	O6-S6	-3.58	1.46	1.56
2	A	801	SCR	O51-S14	-3.55	1.46	1.56

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	801	SCR	C11-C12-C13	-3.34	108.00	115.33
2	A	801	SCR	O1-C12-O10	-2.75	102.11	110.55
2	A	801	SCR	O5-C5-C6	-2.14	102.35	106.64
2	A	801	SCR	O10-C15-C14	2.45	107.55	103.45
2	A	801	SCR	O3-C3-C4	2.49	114.17	108.52

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	801	SCR	S4-O4-C4-C5

There are no ring outliers.

1 monomer is involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	801	SCR	8	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.