



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Mar 2, 2017 – 11:24 am GMT

PDB ID : 3J1Q
EMDB ID: : EMD-5415
Title : Structure of AAV-DJ, a Retargeted Gene Therapy Vector: Cryo-Electron Microscopy at 4.5Å resolution
Authors : Lerch, T.F.; O'Donnell, J.K.; Meyer, N.L.; Xie, Q.; Taylor, K.A.; Stagg, S.M.; Chapman, M.S.
Deposited on : 2012-04-30
Resolution : 4.50 Å (reported)
Based on PDB ID : 1LP3

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
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
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

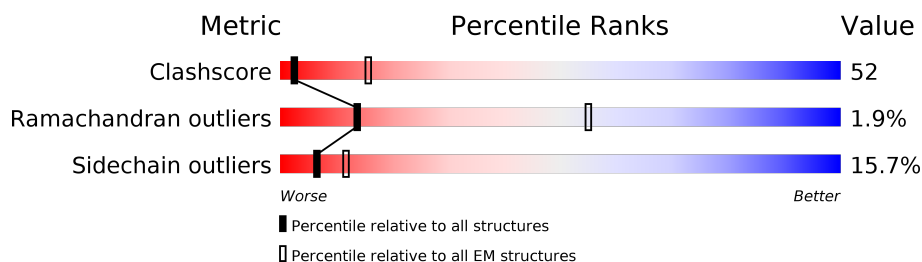
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.50 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 ≥ 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	737	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 4136 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.

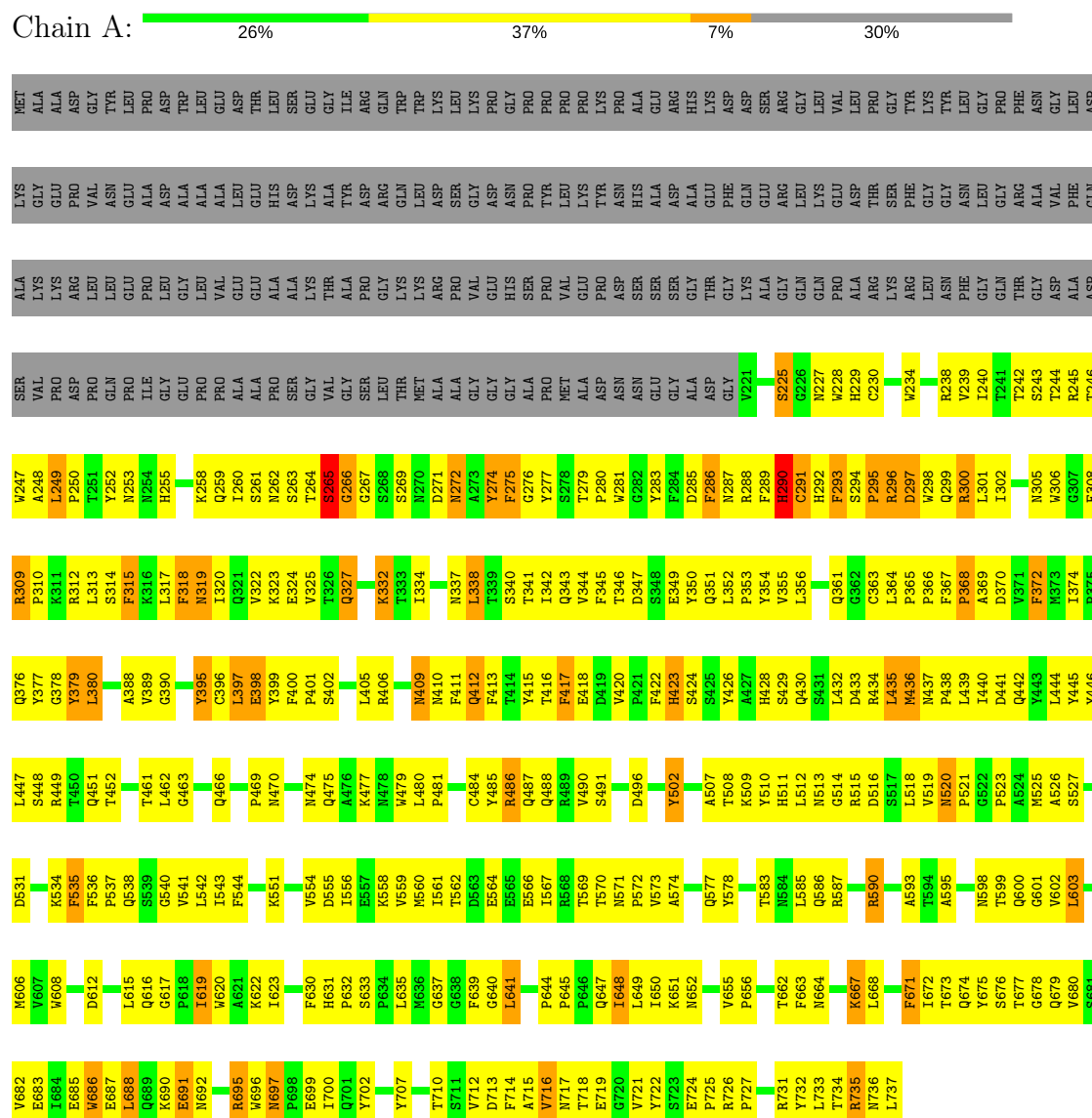
- Molecule 1 is a protein called Adeno-associated virus DJ.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	517	Total	C	N	O	S	0	0
			4136	2608	716	798	14		

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: Adeno-associated virus DJ



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	27312	Depositor
Resolution determination method	FSC at 0.143 cut-off	Depositor
CTF correction method	CTF was estimated using Appion software	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.49	0/4259	0.73	3/5806 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	266	GLY	N-CA-C	-10.12	87.80	113.10
1	A	265	SER	N-CA-C	8.57	134.13	111.00
1	A	267	GLY	N-CA-C	6.23	128.68	113.10

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	4136	0	3902	421	0
All	All	4136	0	3902	421	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 52.

The worst 5 of 421 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:648:ILE:HD13	1:A:648:ILE:H	1.24	1.00
1:A:398:GLU:H	1:A:398:GLU:CD	1.63	0.99
1:A:440:ILE:HG22	1:A:441:ASP:H	1.25	0.98
1:A:265:SER:HB2	1:A:266:GLY:O	1.65	0.95
1:A:312:ARG:HB2	1:A:685:GLU:HB2	1.51	0.93

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	515/737 (70%)	463 (90%)	42 (8%)	10 (2%)	9 49

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	265	SER
1	A	290	HIS
1	A	417	PHE
1	A	632	PRO
1	A	262	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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
-----	-------	----------	-----------	----------	-------------

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	458/626 (73%)	386 (84%)	72 (16%)	3 20

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	412	GLN
1	A	485	TYR
1	A	695	ARG
1	A	413	PHE
1	A	435	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	387	GLN
1	A	428	HIS
1	A	689	GLN
1	A	409	ASN
1	A	475	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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.