



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 11:52 pm BST

PDB ID : 2INY
Title : Nanoporous Crystals of Chicken Embryo Lethal Orphan (CELO) Adenovirus
Major Coat Protein, Hexon
Authors : Xu, L.; Benson, S.D.; Burnett, R.M.
Deposited on : 2006-10-09
Resolution : 3.90 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

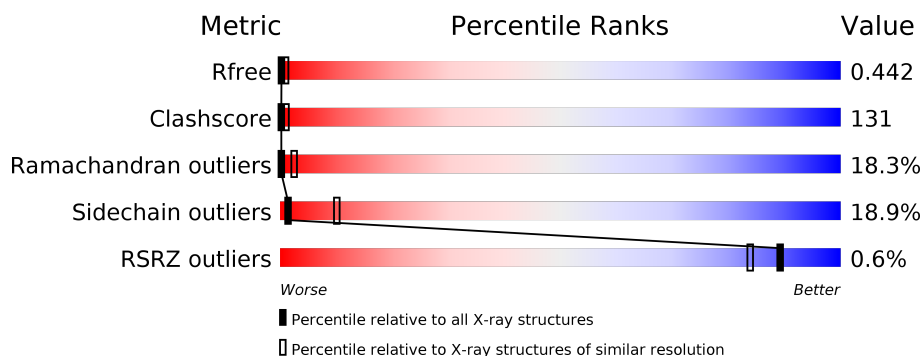
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.90 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	1002 (4.14-3.66)
Clashscore	141614	1004 (4.12-3.68)
Ramachandran outliers	138981	1021 (4.14-3.66)
Sidechain outliers	138945	1014 (4.14-3.66)
RSRZ outliers	127900	1275 (4.20-3.60)

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 ≥ 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	942	<div> <div></div> <div>27%</div> <div>47%</div> <div>22%</div> <div>.</div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 7523 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.

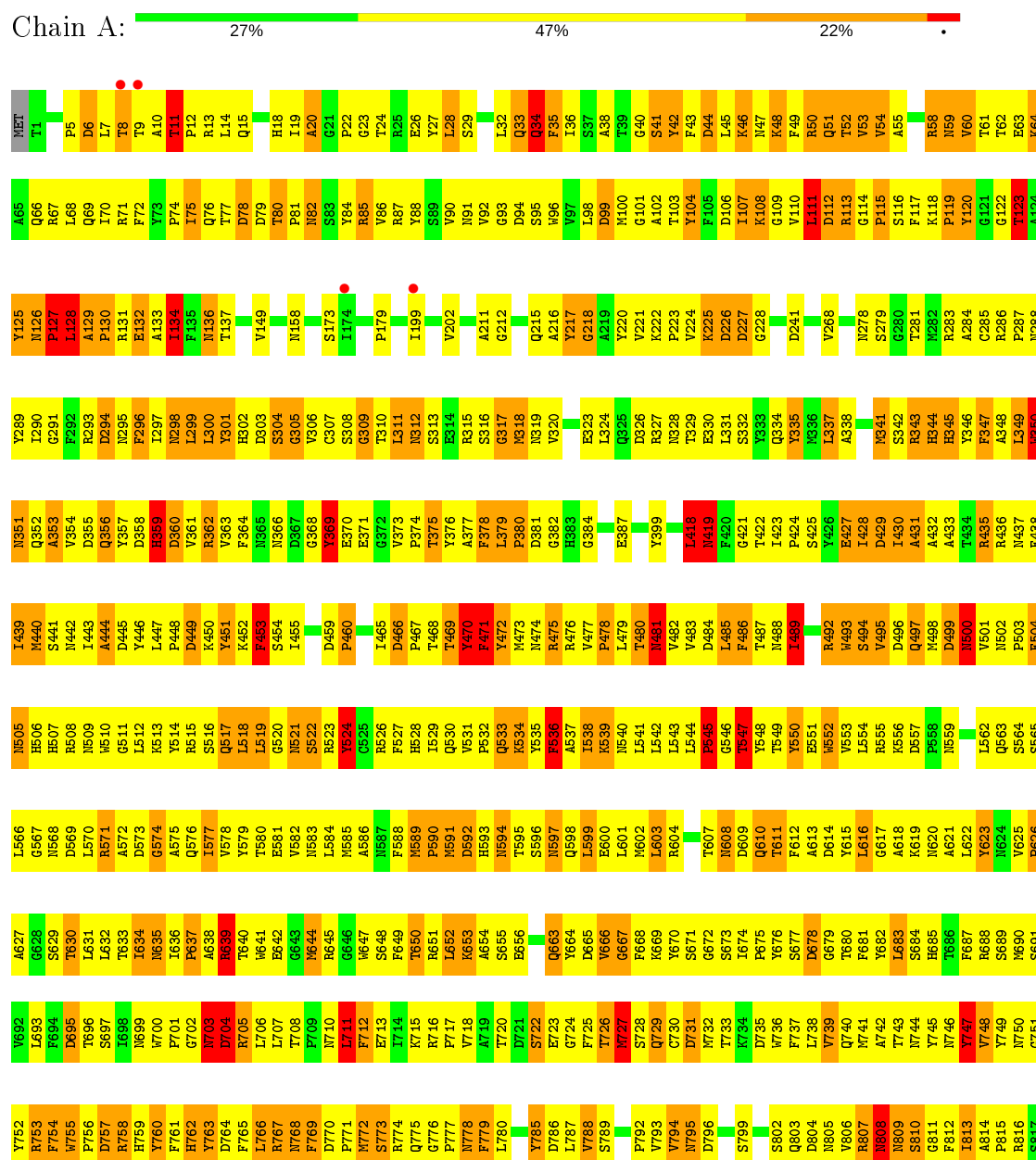
- Molecule 1 is a protein called Hexon protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	941	7523	4764	1281	1444	34	1667	0	0

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 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: Hexon protein



W818	W819	W820	W821	W822	W823	W824	W825	W826	W827	W828		P834	Y835	P836	L837	I838	G839	W840	D841	A842	I843	S844	S845	W846	Q847		W850	K853	F854	L855	C856	D857	W858	Y859	L860	W861	T862	W863	P864	F865	S866	S867	D868	F869		W872	G873	E874	L875	T876	D877	L878	G879	Q880	W881	P882	W883	Y884
T885	W886	W887	S888	H889	S890	W891	V892	L893	W894	F895	E896	L897	D898	P899	W900		T904	Y905	V906	Y907	W908	L909	Y910	G911	V912	F913	D914	T915	V916	R917	Y918	W919	Q920	P921	E922	R923	W924	V925	L926	A927	W928	A929	Y930	F931	R932	T933	P934	F935	A936		V941							

4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α , β , γ	157.77Å 157.77Å 114.20Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.06 – 3.90 47.05 – 3.90	Depositor EDS
% Data completeness (in resolution range)	92.3 (47.06-3.90) 92.3 (47.05-3.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.20	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.86 (at 3.88Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.372 , 0.416 0.350 , 0.442	Depositor DCC
R_{free} test set	731 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	64.5	Xtriage
Anisotropy	0.113	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.23 , 48.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.046 for -h,-k,l	Xtriage
F_o, F_c correlation	0.67	EDS
Total number of atoms	7523	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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	0.85	7/7728 (0.1%)	0.82	5/10545 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	847	GLN	CB-CG	34.13	2.44	1.52
1	A	524	TYR	CD2-CE2	27.73	1.80	1.39
1	A	524	TYR	CD1-CE1	23.54	1.74	1.39
1	A	524	TYR	CE2-CZ	18.88	1.63	1.38
1	A	524	TYR	CE1-CZ	17.19	1.60	1.38

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	847	GLN	CA-CB-CG	9.87	135.12	113.40
1	A	847	GLN	CB-CG-CD	5.74	126.52	111.60
1	A	500	ASN	N-CA-C	5.67	126.32	111.00
1	A	876	THR	N-CA-C	5.52	125.90	111.00
1	A	94	ASP	N-CA-C	-5.34	96.59	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	301	TYR	Sidechain
1	A	747	TYR	Sidechain

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	7523	0	7147	1492	12
All	All	7523	0	7147	1492	12

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 131.

The worst 5 of 1492 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:524:TYR:CE2	1:A:524:TYR:CD2	1.80	1.65
1:A:524:TYR:CD1	1:A:524:TYR:CE1	1.74	1.64
1:A:376:TYR:HE2	1:A:427:GLU:HG2	1.08	1.18
1:A:433:ALA:HA	1:A:436:ARG:NH1	1.58	1.17
1:A:433:ALA:CA	1:A:436:ARG:HH12	1.57	1.16

The worst 5 of 12 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:635:ASN:ND2	1:A:635:ASN:ND2[6_555]	1.53	0.67
1:A:524:TYR:CE2	1:A:847:GLN:CG[3_455]	1.85	0.35
1:A:524:TYR:CZ	1:A:847:GLN:CG[3_455]	1.89	0.31
1:A:524:TYR:CD1	1:A:847:GLN:CB[3_455]	2.00	0.20
1:A:524:TYR:CD2	1:A:847:GLN:CB[3_455]	2.01	0.19

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	939/942 (100%)	508 (54%)	259 (28%)	172 (18%)	0 2

5 of 172 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	6	ASP
1	A	11	THR
1	A	34	GLN
1	A	42	TYR
1	A	48	LYS

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	821/822 (100%)	666 (81%)	155 (19%)	1 10

5 of 155 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	489	ILE
1	A	577	ILE
1	A	858	ASN
1	A	493	TRP
1	A	533	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	505	ASN
1	A	540	ASN
1	A	759	HIS
1	A	517	GLN
1	A	215	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.

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, 95th 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 ²)	Q<0.9
1	A	719/942 (76%)	-0.12	4 (0%) 89 84	30, 30, 30, 30	4 (0%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	199	ILE	8.7
1	A	8	THR	3.1
1	A	9	THR	2.6
1	A	174	ILE	2.1

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.