wwPDB EM Validation Summary Report (i)

Mar 6, 2024 – 12:07 PM JST

PDB ID	:	1UCU
Title	:	R-type straight flagellar filament made of full-length flagellin
Authors	:	Yonekura, K.; Maki-Yonekura, S.; Namba, K.
Deposited on	:	2003-04-22
Resolution	:	4.00 Å(reported)
Based on initial model	:	1IO1

This is a wwPDB EM 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
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
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:

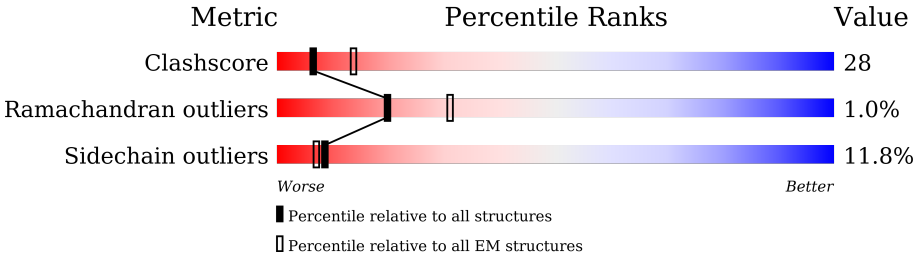
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 4.00 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	494	56% 36% 6% •

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3617 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 phase 1 Flagellin.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	494	3617	2191	640	784	2	0	0

There is a discrepancy between the modelled and reference sequences:

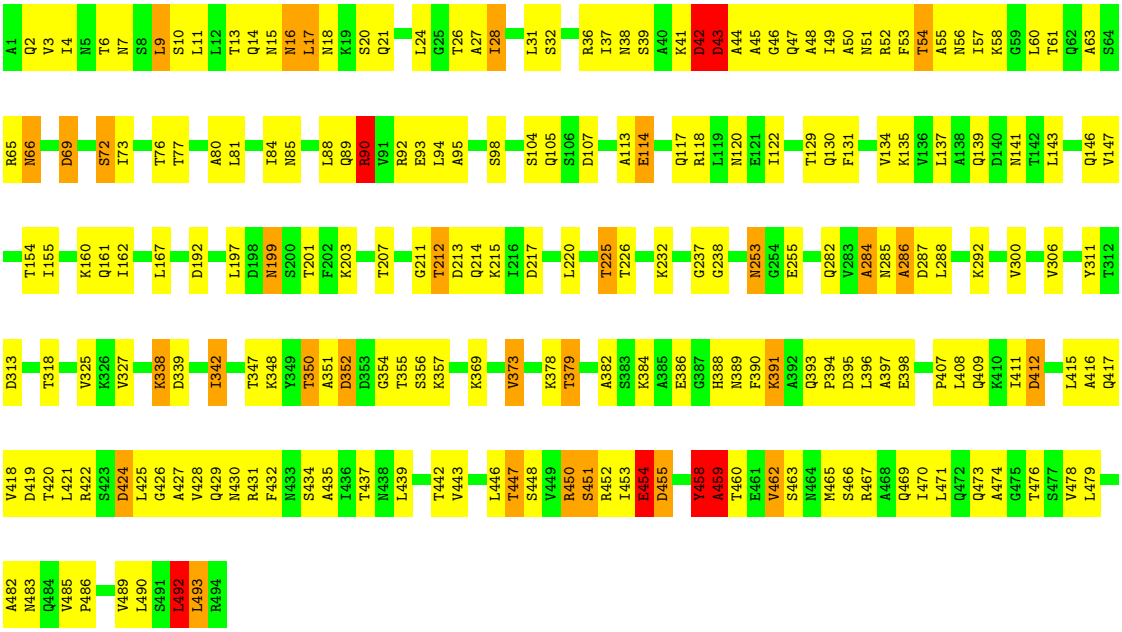
Chain	Residue	Modelled	Actual	Comment	Reference
A	449	VAL	ALA	SEE REMARK 999	UNP P06179

3 Residue-property plots (i)

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. 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: phase 1 Flagellin

Chain A: 56% 36% 6% •



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	1.00Å 1.00Å 1.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 4.00	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-4.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	FEX-PLOR, FX-PLOR	Depositor
R, R_{free}	0.300 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3617	wwPDB-VP
Average B, all atoms (Å^2)	10.0	wwPDB-VP

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	0.67	4/3640 (0.1%)	0.95	20/4941 (0.4%)

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	1	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	212	THR	CB-CG2	9.16	1.82	1.52
1	A	238	GLY	N-CA	5.50	1.54	1.46
1	A	452	ARG	CB-CG	-5.49	1.37	1.52
1	A	237	GLY	CA-C	-5.36	1.43	1.51

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	90	ARG	NE-CZ-NH2	-13.09	113.75	120.30
1	A	66	ASN	CB-CG-OD1	-8.66	104.28	121.60
1	A	118	ARG	NE-CZ-NH2	-8.20	116.20	120.30
1	A	458	TYR	CB-CA-C	7.52	125.45	110.40
1	A	211	GLY	N-CA-C	-6.99	95.62	113.10

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	54	THR	CB

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	66	ASN	Sidechain
1	A	90	ARG	Sidechain

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	3617	0	3580	203	0
All	All	3617	0	3580	203	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 28.

The worst 5 of 203 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:212:THR:CG2	1:A:212:THR:CB	1.82	1.54
1:A:450:ARG:HD2	1:A:454:GLU:HG2	1.21	1.11
1:A:424:ASP:O	1:A:428:VAL:HG23	1.53	1.07
1:A:53:PHE:HA	1:A:56:ASN:HD22	1.17	1.05
1:A:427:ALA:O	1:A:431:ARG:HG3	1.63	0.99

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	492/494 (100%)	460 (94%)	27 (6%)	5 (1%)	15 53

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	286	ALA
1	A	352	ASP
1	A	395	ASP
1	A	459	ALA
1	A	394	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 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	391/391 (100%)	345 (88%)	46 (12%)	5 24

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	386	GLU
1	A	437	THR
1	A	391	LYS
1	A	419	ASP
1	A	450	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	141	ASN
1	A	282	GLN
1	A	469	GLN
1	A	388	HIS
1	A	89	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 monosaccharides 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.