wwPDB EM Validation Summary Report (i)

Apr 21, 2024 – 11:57 am BST

PDB ID	:	1H6I
Title	:	A REFINED STRUCTURE OF HUMAN AQUAPORIN 1
Authors	:	De Groot, B.L.; Engel, A.; Grubmuller, H.
Deposited on	:	2001-06-15
Resolution	:	3.54 Å(reported)

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:

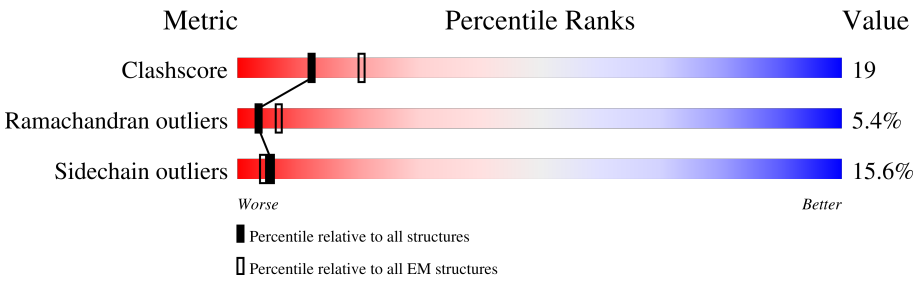
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.54 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	269	42% 36% 6% 16%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1653 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 AQUAPORIN-1.

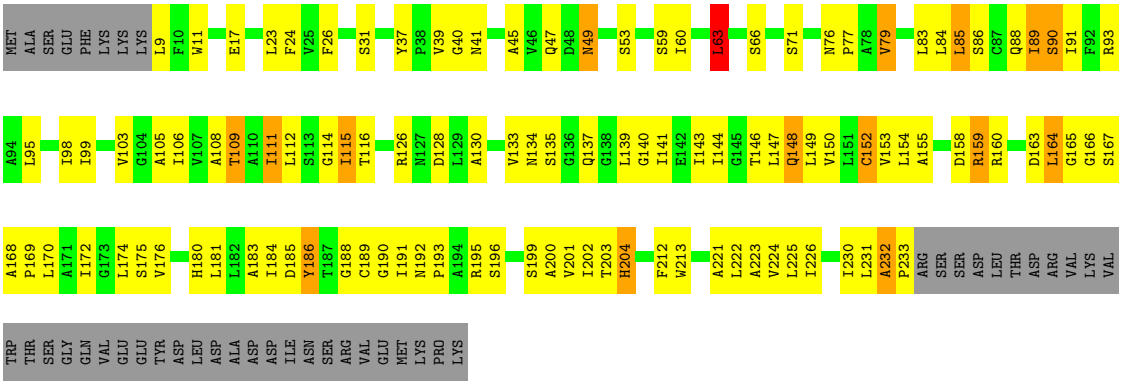
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	225	1653	1071	281	296	5	0	0

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: AQUAPORIN-1

Chain A: 42% 36% 6% 16%



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 4 21 2	Depositor
Cell constants a, b, c, α, β, γ	99.58Å 99.58Å 100.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.04 – 3.54	Depositor
% Data completeness (in resolution range)	82.8 (30.04-3.54)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.364 , 0.376	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1653	wwPDB-VP
Average B, all atoms (Å^2)	44.0	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	0.41	0/1685	0.73	3/2297 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	37	TYR	N-CA-C	6.63	128.90	111.00
1	A	63	LEU	CA-CB-CG	6.61	130.49	115.30
1	A	45	ALA	N-CA-C	-5.49	96.18	111.00

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	1653	0	1699	64	0
All	All	1653	0	1699	64	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 19.

The worst 5 of 64 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:200:ALA:O	1:A:204:HIS:HA	1.71	0.89
1:A:91:ILE:HD12	1:A:91:ILE:H	1.52	0.73
1:A:139:LEU:O	1:A:143:ILE:HG12	1.88	0.72
1:A:95:LEU:O	1:A:99:ILE:HG13	1.90	0.70
1:A:86:SER:O	1:A:88:GLN:HG2	1.92	0.69

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	223/269 (83%)	180 (81%)	31 (14%)	12 (5%)	2 19

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	39	VAL
1	A	49	ASN
1	A	90	SER
1	A	164	LEU
1	A	40	GLY

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
1	A	173/214 (81%)	146 (84%)	27 (16%)	2 17

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	135	SER
1	A	152	CYS
1	A	199	SER
1	A	148	GLN
1	A	159	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	41	ASN
1	A	88	GLN
1	A	101	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.