wwPDB X-ray Structure Validation Summary Report (i)

May 23, 2024 – 04:24 PM EDT

PDB ID	:	4U46
Title	:	Crystal structure of an avidin mutant
Authors	:	Agrawal, N.; Lehtonen, S.; Kahkonen, N.; Riihimaki, T.; Hytonen, V.P.; Kulomaa, M.S.; Johnson, M.S.; Airenne, T.T.
Deposited on	:	2014-07-23
Resolution	:	1.95 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 (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:

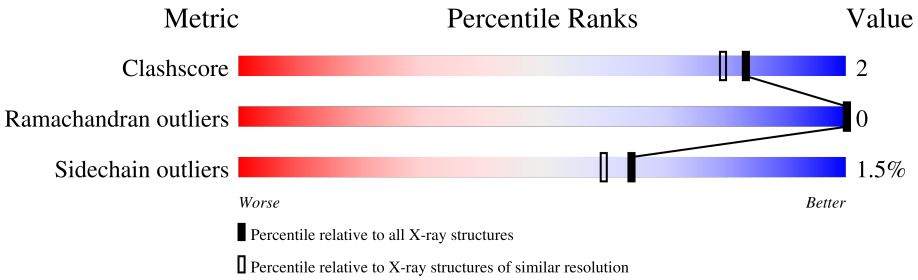
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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))
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1957 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 Avidin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	116	Total	C	N	O	S	0	2	0
			936	590	166	175	5			
1	B	116	Total	C	N	O	S	0	2	0
			937	591	166	175	5			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLN	-	expression tag	UNP P02701
A	-2	THR	-	expression tag	UNP P02701
A	-1	VAL	LEU	expression tag	UNP P02701
A	0	ASN	SER	expression tag	UNP P02701
A	13	ARG	ASP	conflict	UNP P02701
A	14	MET	LEU	conflict	UNP P02701
A	15	ASN	GLY	conflict	UNP P02701
A	16	HIS	SER	conflict	UNP P02701
A	35	ALA	THR	conflict	UNP P02701
A	36	THR	ALA	conflict	UNP P02701
A	38	ASN	THR	conflict	UNP P02701
A	117	TYR	ILE	conflict	UNP P02701
A	129	HIS	-	expression tag	UNP P02701
A	130	HIS	-	expression tag	UNP P02701
A	131	HIS	-	expression tag	UNP P02701
A	132	HIS	-	expression tag	UNP P02701
A	133	HIS	-	expression tag	UNP P02701
A	134	HIS	-	expression tag	UNP P02701
B	-3	GLN	-	expression tag	UNP P02701
B	-2	THR	-	expression tag	UNP P02701
B	-1	VAL	LEU	expression tag	UNP P02701
B	0	ASN	SER	expression tag	UNP P02701
B	13	ARG	ASP	conflict	UNP P02701
B	14	MET	LEU	conflict	UNP P02701
B	15	ASN	GLY	conflict	UNP P02701

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	16	HIS	SER	conflict	UNP P02701
B	35	ALA	THR	conflict	UNP P02701
B	36	THR	ALA	conflict	UNP P02701
B	38	ASN	THR	conflict	UNP P02701
B	117	TYR	ILE	conflict	UNP P02701
B	129	HIS	-	expression tag	UNP P02701
B	130	HIS	-	expression tag	UNP P02701
B	131	HIS	-	expression tag	UNP P02701
B	132	HIS	-	expression tag	UNP P02701
B	133	HIS	-	expression tag	UNP P02701
B	134	HIS	-	expression tag	UNP P02701

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	6	Total Cl 6 6	0	0
2	B	6	Total Cl 6 6	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	34	Total O 34 34	0	0
3	B	38	Total O 38 38	0	0

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	74.48Å 79.80Å 43.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.34 – 1.95	Depositor
% Data completeness (in resolution range)	99.5 (19.34-1.95)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.02 (at 1.94Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.182 , 0.214	Depositor
Wilson B-factor (Å^2)	24.2	Xtriage
Anisotropy	0.180	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1957	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.70% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

4 Model quality

4.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CL

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	0.79	0/961	0.84	2/1298 (0.2%)
1	B	0.78	0/962	0.86	1/1300 (0.1%)
All	All	0.78	0/1923	0.85	3/2598 (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	86	ASP	CB-CG-OD1	5.72	123.45	118.30
1	B	18	MET	CG-SD-CE	5.53	109.04	100.20
1	A	86	ASP	CB-CG-OD2	-5.34	113.49	118.30

There are no chirality outliers.

There are no planarity outliers.

4.2 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	936	0	924	2	0
1	B	937	0	926	7	0
2	A	6	0	0	0	0
2	B	6	0	0	0	0
3	A	34	0	0	0	0
3	B	38	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	1957	0	1850	9	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 2.

The worst 5 of 9 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:B:36:THR:HG22	1:B:44:ILE:CD1	2.19	0.72
1:B:94:LYS:HD2	1:B:119:ILE:HD12	1.81	0.63
1:B:36:THR:HG22	1:B:44:ILE:HD12	1.80	0.62
1:A:12:ASN:O	1:A:14:MET:HG2	2.06	0.55
1:A:46:GLU:CD	1:A:46:GLU:H	2.15	0.49

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	114/138 (83%)	113 (99%)	1 (1%)	0	100	100
1	B	114/138 (83%)	113 (99%)	1 (1%)	0	100	100
All	All	228/276 (83%)	226 (99%)	2 (1%)	0	100	100

There are no Ramachandran outliers to report.

4.3.2 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	105/123 (85%)	103 (98%)	2 (2%)	57 50
1	B	105/123 (85%)	104 (99%)	1 (1%)	76 74
All	All	210/246 (85%)	207 (99%)	3 (1%)	65 62

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	87	ARG
1	A	94	LYS
1	B	71	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	ASN
1	A	53	GLN
1	A	118	ASN
1	B	17	ASN

4.3.3 RNA (i)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

5.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

5.4 Ligands

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers

EDS failed to run properly - this section is therefore empty.