



wwPDB X-ray Structure Validation Summary Report

Mar 31, 2014 – 01:46 PM BST

PDB ID : 4OBB
Title : The crystal structure of a solute-binding protein from *Anabaena variabilis* ATCC 29413 in complex with (3S)-3-methyl-2-oxopentanoic acid.
Authors : Tan, K.; Li, H.; Jedrzejczak, R.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on : 2014-01-07
Resolution : 1.53 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

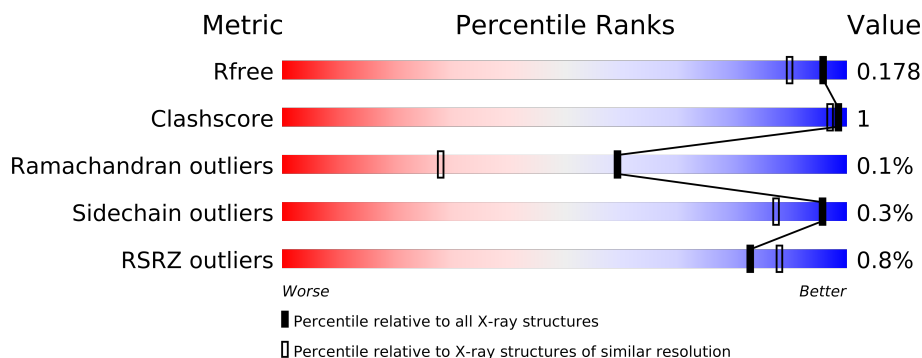
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable23004
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable23004

1 Overall quality at a glance

The reported resolution of this entry is 1.53 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}	66092	1879 (1.54-1.50)
Clashscore	79885	2184 (1.54-1.50)
Ramachandran outliers	78287	2123 (1.54-1.50)
Sidechain outliers	78261	2121 (1.54-1.50)
RSRZ outliers	66119	1880 (1.54-1.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	395	
1	B	395	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6413 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

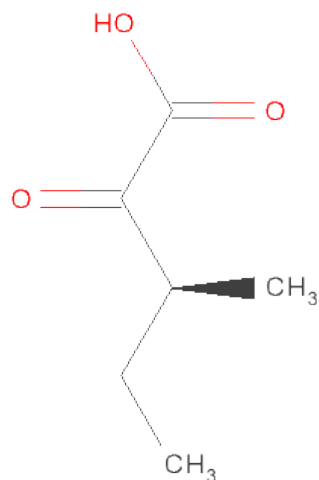
- Molecule 1 is a protein called Amino acid/amide ABC transporter substrate-binding protein, HAAT family.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	364	Total	C	N	O	S	Se	0	6	0
			2765	1764	462	536	2	1			
1	B	364	Total	C	N	O	S	Se	0	8	0
			2789	1782	465	539	2	1			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	SER	-	EXPRESSION TAG	UNP Q3MFZ5
A	23	ASN	-	EXPRESSION TAG	UNP Q3MFZ5
A	24	ALA	-	EXPRESSION TAG	UNP Q3MFZ5
B	22	SER	-	EXPRESSION TAG	UNP Q3MFZ5
B	23	ASN	-	EXPRESSION TAG	UNP Q3MFZ5
B	24	ALA	-	EXPRESSION TAG	UNP Q3MFZ5

- Molecule 2 is (3S)-3-METHYL-2-OXOPENTANOICACID (three-letter code: 1QQ) (formula: C₆H₁₀O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			9	6	3		
2	B	1	Total	C	O	0	0
			9	6	3		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	454	Total	O	0	0
			454	454		
4	B	385	Total	O	0	0
			385	385		

i

- Molecule 1: Amino acid/amide ABC transporter substrate-binding protein, HAAT family

SER
ASN
ALA
THR
ASN
THR
ASP
THR
ASN
SER
THR
ASN
ASN
SER
SER
PRO
ASN
ASN
THR
THR
ASN
THR
THR
THR
THR
ASN
VAL
THR
THR
THR
SER
ASP
LYS
N5-3
C291
T309
K326
R405
S408
F416

- Molecule 1: Amino acid/amide ABC transporter substrate-binding protein, HAAT family

SER	ASN	ALA	THR	ASN	THR	ASP	THR	SER	ASN	ASN	SER	THR	ASN	ASN	SER	PRO	ASN	ASN	THR	THR	ASN	THR	THR	THR	THR	ASN	VAL	THR	THR	THR	SER	SER	LYS	553	\$153	C291	T309	K314	Q318	D322	K405	K416
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	98.74Å 100.81Å 150.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.86 – 1.53 25.86 – 1.53	Depositor EDS
% Data completeness (in resolution range)	99.1 (25.86-1.53) 99.1 (25.86-1.53)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.72 (at 1.53Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.132 , 0.178 0.133 , 0.178	Depositor DCC
R_{free} test set	5643 reflections (5.01%)	DCC
Wilson B-factor (Å ²)	13.9	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 41.0	EDS
Estimated twinning fraction	0.021 for -k,-h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 112655 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	6413	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.11% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, 1QQ

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	0.32	0/2813	0.51	0/3816
1	B	0.32	0/2856	0.50	0/3870
All	All	0.32	0/5669	0.51	0/7686

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2765	0	0	1	0
1	B	2789	0	0	2	0
2	A	9	0	0	0	0
2	B	9	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	454	0	0	1	0
4	B	385	0	0	0	0
All	All	6413	0	0	3	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 1.

All (3) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:309:THR:O	1:B:314:LYS:NZ	2.24	0.70
1:B:318:GLN:NE2	1:B:322:ASP:OD1	2.39	0.55
1:A:326:LYS:NZ	4:A:1014:HOH:O	2.53	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	368/395 (93%)	360 (98%)	8 (2%)	0	100	100
1	B	370/395 (94%)	363 (98%)	6 (2%)	1 (0%)	50	20
All	All	738/790 (93%)	723 (98%)	14 (2%)	1 (0%)	59	28

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	153	SER

5.3.2 Protein sidechains ⓘ

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	289/325 (89%)	288 (100%)	1 (0%)	96	87
1	B	297/325 (91%)	296 (100%)	1 (0%)	96	87

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	586/650 (90%)	584 (100%)	2 (0%)	96	87

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	291	CYS
1	B	291	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	1QQ	A	501	-	8,8,8	0.76	0	10,10,10	0.87	0
2	1QQ	B	501	-	8,8,8	0.73	0	10,10,10	1.18	1 (10%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	1QQ	A	501	-	-	0/10/10/10	0/0/0/0
2	1QQ	B	501	-	-	0/10/10/10	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	1QQ	OXT-C42-C44	2.64	119.68	114.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	364/395 (92%)	-0.47	4 (1%) 77 84	8, 12, 25, 38	0
1	B	364/395 (92%)	-0.43	2 (0%) 88 92	8, 13, 26, 35	0
All	All	728/790 (92%)	-0.45	6 (0%) 83 88	8, 12, 26, 38	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	405	LYS	2.8
1	A	405	LYS	2.6
1	A	416	LYS	2.5
1	A	408	SER	2.2
1	A	309	THR	2.2

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	1QQ	A	501	9/9	0.05	-0.46	11,11,18,18	0
2	1QQ	B	501	9/9	0.05	-1.09	11,12,17,18	0
3	MG	A	502	1/1	0.03	-1.99	10,10,10,10	0
3	MG	B	502	1/1	0.02	-3.32	11,11,11,11	0

6.5 Other polymers

There are no such residues in this entry.