



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 16, 2024 – 06:42 AM EDT

PDB ID : 1YBF
Title : Crystal structure of AMP nucleosidase from *Bacteroides thetaiotaomicron* VPI-5482
Authors : Krishnamurthy, N.R.; Swaminathan, S.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2004-12-20
Resolution : 2.90 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 1.20.1
EDS : 2.37.1
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) : 2.37.1

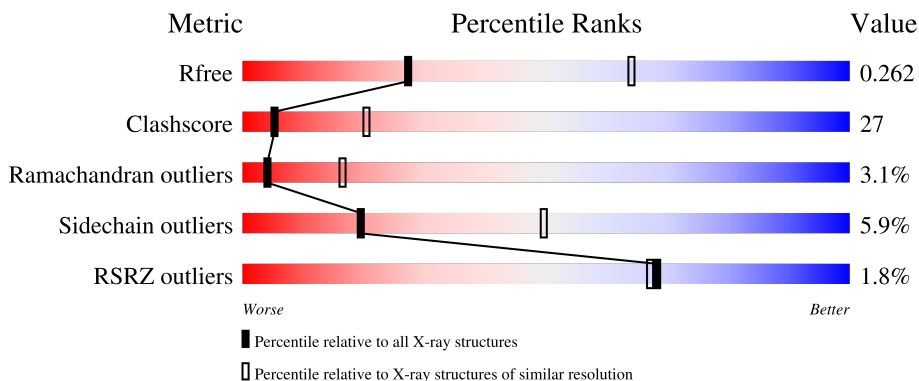
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 of 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	268	 3% 41% 44% 10%
1	B	268	 % 48% 32% 15%
1	C	268	 51% 31% 14%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5471 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 AMP nucleosidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	240	Total 1876	C 1205	N 308	O 354	S 1	Se 8	0	0	0
1	B	228	Total 1778	C 1142	N 294	O 334	S 1	Se 7	0	0	0
1	C	230	Total 1801	C 1161	N 295	O 336	S 1	Se 8	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total 3	O 3	0	0
2	B	3	Total 3	O 3	0	0
2	C	10	Total 10	O 10	0	0

A85	S182	ASN
P88	G183	TRP
V91	V184	GLU
I92	D185	GLY
K96	M186	HIS
L100	E187	HIS
LEU	T188	HIS
GLU	A189	HIS
ASN	T190	HIS
A105	L191	HIS
L106	M192	
G107	G195	
D108	I200	
I116	P201	
R117	M202	
E119	R211	
M123	P212	
E128	M213	
L133	F214	
R140	P215	
A141	E216	
I142	GLY	
I146	VAL	
Q147	LYS	
M148	THR	
K149	GLU	
G150	GLU	
Y153	SER	
W154	ASP	
T155	ASP	
R163	LEU	
E167	GLN	
E170	LEU	
K171	VAL	
F172	THR	
K173	ASP	
D174	M230	
Y175	F231	
L176	A232	
R177	E233	
S178	M237	
T179	D241	
H180	A242	
A181	L243	
	R247	
	E248	
	M249	
	K250	
	SER	
	SER	
	ILE	
	LYS	
	HIS	
	ILE	
	ARG	
	PHE	

4 Data and refinement statistics

Property	Value	Source
Space group	P 41 2 2	Depositor
Cell constants a, b, c, α , β , γ	108.12Å 108.12Å 160.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.90 46.30 – 2.70	Depositor EDS
% Data completeness (in resolution range)	95.0 (50.00-2.90) 97.1 (46.30-2.70)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.75 (at 2.69Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.228 , 0.264 0.228 , 0.262	Depositor DCC
R_{free} test set	1021 reflections (3.80%)	wwPDB-VP
Wilson B-factor (Å ²)	48.3	Xtrriage
Anisotropy	0.203	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 50.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	5471	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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	0.37	0/1913	0.62	0/2587
1	B	0.39	0/1813	0.63	0/2454
1	C	0.41	0/1837	0.64	0/2483
All	All	0.39	0/5563	0.63	0/7524

There are no bond length outliers.

There are no bond angle outliers.

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	1876	0	1836	128	0
1	B	1778	0	1742	87	0
1	C	1801	0	1763	76	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
2	C	10	0	0	0	0
All	All	5471	0	5341	287	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 27.

The worst 5 of 287 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:188:THR:HG22	1:B:192:MSE:HE2	1.24	1.18
1:A:71:MSE:HE3	1:A:164:ARG:HH21	1.29	0.94
1:C:117:ARG:HH21	1:C:123:ASN:HD22	1.13	0.94
1:C:17:TYR:HB3	1:C:56:MSE:HE2	1.54	0.90
1:A:102:LEU:HB3	1:A:208:ILE:HD13	1.56	0.88

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 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	236/268 (88%)	200 (85%)	27 (11%)	9 (4%)	3	13
1	B	224/268 (84%)	196 (88%)	21 (9%)	7 (3%)	4	16
1	C	224/268 (84%)	202 (90%)	17 (8%)	5 (2%)	6	24
All	All	684/804 (85%)	598 (87%)	65 (10%)	21 (3%)	4	16

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	52	GLU
1	C	215	PRO
1	A	120	GLY
1	B	104	ASN
1	A	71	MSE

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	200/222 (90%)	190 (95%)	10 (5%)	24	57
1	B	189/222 (85%)	178 (94%)	11 (6%)	20	50
1	C	191/222 (86%)	178 (93%)	13 (7%)	16	42
All	All	580/666 (87%)	546 (94%)	34 (6%)	19	49

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	128	GLU
1	C	155	THR
1	C	243	LEU
1	B	50	VAL
1	B	33	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	39	HIS
1	B	148	ASN
1	C	162	ASN
1	C	123	ASN
1	C	148	ASN

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.

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	232/268 (86%)	0.36	9 (3%) 39 35	28, 55, 70, 78	0
1	B	221/268 (82%)	-0.14	3 (1%) 75 75	20, 37, 57, 75	0
1	C	222/268 (82%)	-0.25	0 100 100	15, 33, 51, 67	0
All	All	675/804 (83%)	-0.01	12 (1%) 68 67	15, 40, 67, 78	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	102	LEU	4.4
1	A	102	LEU	3.7
1	B	104	ASN	3.1
1	A	15	PRO	2.8
1	A	142	ILE	2.7

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 monosaccharides 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.