wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 06:57 AM EDT

PDB ID	:	1YBT
Title	:	MYCOBACTERIUM TUBERCULOSIS ADENYLYL CYCLASE, RV1900C CHD
Authors	:	Sinha, S.C.; Wetterer, M.; Sprang, S.R.; Schultz, J.E.; Linder, J.U.
Deposited on	:	2004-12-21
Resolution	:	2.31 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:

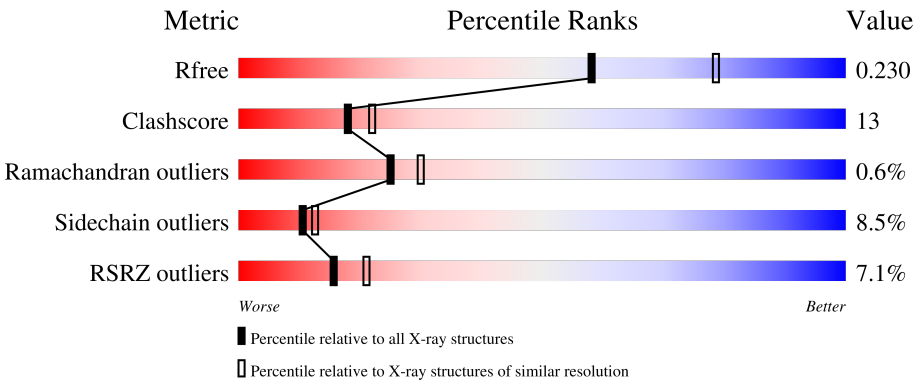
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.31 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	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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 <=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	184	2%, 66%, 21%, •, 9%
1	B	184	11%, 67%, 23%, •, 7%
1	C	184	10%, 71%, 21%, •, 7%
1	D	184	2%, 65%, 23%, •, 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5213 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 hydrolase, alpha/beta hydrolase fold family.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	167	1244	764	234	239	4	3	0	0	0
1	B	172	1286	787	245	247	4	3	0	0	0
1	C	172	1286	787	245	247	4	3	0	0	0
1	D	167	1244	764	234	239	4	3	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	279	MSE	-	CLONING ARTIFACT	UNP O07732
A	280	ARG	-	CLONING ARTIFACT	UNP O07732
A	281	GLY	-	CLONING ARTIFACT	UNP O07732
A	282	SER	-	CLONING ARTIFACT	UNP O07732
A	283	HIS	-	EXPRESSION TAG	UNP O07732
A	284	HIS	-	EXPRESSION TAG	UNP O07732
A	285	HIS	-	EXPRESSION TAG	UNP O07732
A	286	HIS	-	EXPRESSION TAG	UNP O07732
A	287	HIS	-	EXPRESSION TAG	UNP O07732
A	288	HIS	-	EXPRESSION TAG	UNP O07732
A	289	GLY	-	CLONING ARTIFACT	UNP O07732
A	290	SER	-	CLONING ARTIFACT	UNP O07732
A	294	MSE	MET	MODIFIED RESIDUE	UNP O07732
A	299	MSE	MET	MODIFIED RESIDUE	UNP O07732
A	454	MSE	MET	MODIFIED RESIDUE	UNP O07732
B	279	MSE	-	CLONING ARTIFACT	UNP O07732
B	280	ARG	-	CLONING ARTIFACT	UNP O07732
B	281	GLY	-	CLONING ARTIFACT	UNP O07732
B	282	SER	-	CLONING ARTIFACT	UNP O07732
B	283	HIS	-	EXPRESSION TAG	UNP O07732
B	284	HIS	-	EXPRESSION TAG	UNP O07732

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	285	HIS	-	EXPRESSION TAG	UNP O07732
B	286	HIS	-	EXPRESSION TAG	UNP O07732
B	287	HIS	-	EXPRESSION TAG	UNP O07732
B	288	HIS	-	EXPRESSION TAG	UNP O07732
B	289	GLY	-	CLONING ARTIFACT	UNP O07732
B	290	SER	-	CLONING ARTIFACT	UNP O07732
B	294	MSE	MET	MODIFIED RESIDUE	UNP O07732
B	299	MSE	MET	MODIFIED RESIDUE	UNP O07732
B	454	MSE	MET	MODIFIED RESIDUE	UNP O07732
C	279	MSE	-	CLONING ARTIFACT	UNP O07732
C	280	ARG	-	CLONING ARTIFACT	UNP O07732
C	281	GLY	-	CLONING ARTIFACT	UNP O07732
C	282	SER	-	CLONING ARTIFACT	UNP O07732
C	283	HIS	-	EXPRESSION TAG	UNP O07732
C	284	HIS	-	EXPRESSION TAG	UNP O07732
C	285	HIS	-	EXPRESSION TAG	UNP O07732
C	286	HIS	-	EXPRESSION TAG	UNP O07732
C	287	HIS	-	EXPRESSION TAG	UNP O07732
C	288	HIS	-	EXPRESSION TAG	UNP O07732
C	289	GLY	-	CLONING ARTIFACT	UNP O07732
C	290	SER	-	CLONING ARTIFACT	UNP O07732
C	294	MSE	MET	CLONING ARTIFACT	UNP O07732
C	299	MSE	MET	MODIFIED RESIDUE	UNP O07732
C	454	MSE	MET	MODIFIED RESIDUE	UNP O07732
D	279	MSE	-	MODIFIED RESIDUE	UNP O07732
D	280	ARG	-	CLONING ARTIFACT	UNP O07732
D	281	GLY	-	CLONING ARTIFACT	UNP O07732
D	282	SER	-	CLONING ARTIFACT	UNP O07732
D	283	HIS	-	EXPRESSION TAG	UNP O07732
D	284	HIS	-	EXPRESSION TAG	UNP O07732
D	285	HIS	-	EXPRESSION TAG	UNP O07732
D	286	HIS	-	EXPRESSION TAG	UNP O07732
D	287	HIS	-	EXPRESSION TAG	UNP O07732
D	288	HIS	-	EXPRESSION TAG	UNP O07732
D	289	GLY	-	CLONING ARTIFACT	UNP O07732
D	290	SER	-	CLONING ARTIFACT	UNP O07732
D	294	MSE	MET	MODIFIED RESIDUE	UNP O07732
D	299	MSE	MET	MODIFIED RESIDUE	UNP O07732
D	454	MSE	MET	MODIFIED RESIDUE	UNP O07732

- Molecule 2 is water.

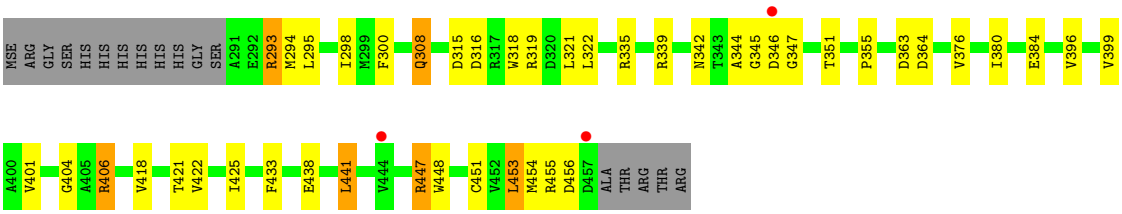
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	43	43	43	0	0
			Total	O		
2	B	35	35	35	0	0
			Total	O		
2	C	37	37	37	0	0
			Total	O		
2	D	38	38	38	0	0

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

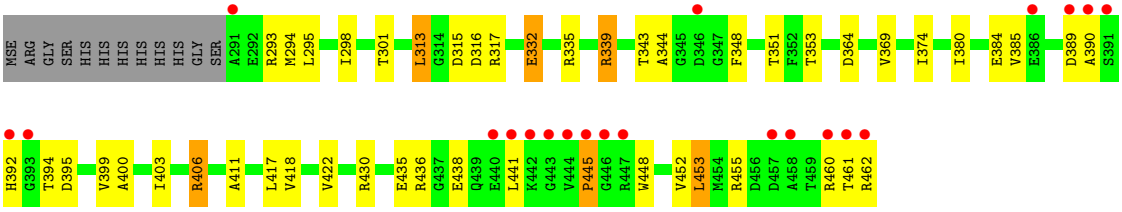
• Molecule 1: hydrolase, alpha/beta hydrolase fold family

Chain A:



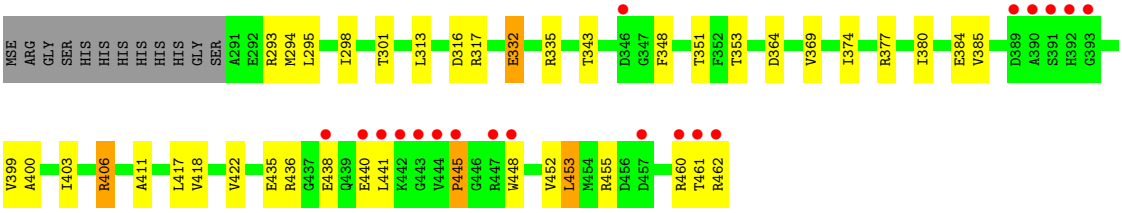
• Molecule 1: hydrolase, alpha/beta hydrolase fold family

Chain B:



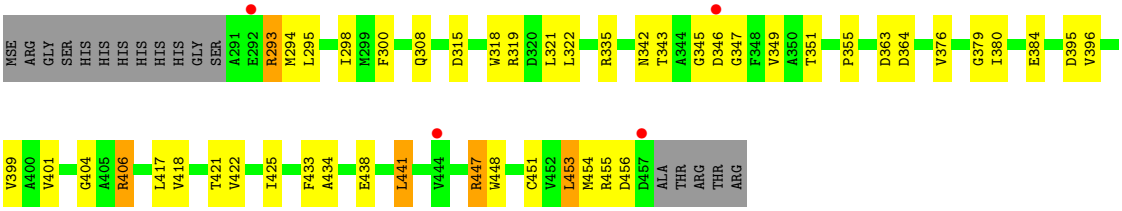
• Molecule 1: hydrolase, alpha/beta hydrolase fold family

Chain C:



• Molecule 1: hydrolase, alpha/beta hydrolase fold family

Chain D:


MSE ARG GLY SER HIS HIS HIS HIS HIS HIS GLY SER A291 E292 R293 M294 L295 I298 M299 F300 Q308 D315 W318 R319 D320 L321 L322 R335 N342 T343 A344 G345 D346 G347 F348 V349 A350 T351 P355 D363 D364 V376 G379 I380 E384 D395 V396
V399 A400 V401 G404 A405 R406 L417 V418 T421 V422 I425 F433 A434 E438 L441 V444 R447 W448 C451 V452 L453 M454 R455 D456 D457 ALA THR ARG THR ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	90.70Å 44.39Å 80.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.94 – 2.31 19.93 – 2.32	Depositor EDS
% Data completeness (in resolution range)	99.1 (19.94-2.31) 99.1 (19.93-2.32)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.56 (at 2.33Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.194 , 0.230 0.195 , 0.230	Depositor DCC
R_{free} test set	1401 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	39.3	Xtriage
Anisotropy	0.448	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 40.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.479 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5213	wwPDB-VP
Average B, all atoms (Å^2)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.14% 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.

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	0.55	0/1259	0.78	1/1702 (0.1%)
1	B	0.53	0/1301	0.82	2/1757 (0.1%)
1	C	0.54	0/1301	0.80	2/1757 (0.1%)
1	D	0.55	0/1259	0.78	1/1702 (0.1%)
All	All	0.54	0/5120	0.79	6/6918 (0.1%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	445	PRO	CA-N-CD	-8.89	99.06	111.50
1	B	445	PRO	CA-N-CD	-7.92	100.42	111.50
1	B	453	LEU	CA-CB-CG	6.29	129.77	115.30
1	C	453	LEU	CA-CB-CG	5.82	128.68	115.30
1	D	453	LEU	CA-CB-CG	5.39	127.69	115.30

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	1244	0	1216	39	0
1	B	1286	0	1261	35	0
1	C	1286	0	1261	29	0
1	D	1244	0	1216	37	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	43	0	0	1	0
2	B	35	0	0	2	0
2	C	37	0	0	1	0
2	D	38	0	0	0	0
All	All	5213	0	4954	135	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 13.

The worst 5 of 135 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:344:ALA:HB2	2:B:109:HOH:O	1.76	0.85
1:C:377:ARG:NH1	1:C:411:ALA:O	2.22	0.72
1:C:436:ARG:HH21	1:C:460:ARG:HB3	1.56	0.70
1:A:438:GLU:HG2	1:A:447:ARG:NH1	2.06	0.70
1:D:345:GLY:O	1:D:347:GLY:N	2.25	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 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	165/184 (90%)	160 (97%)	4 (2%)	1 (1%)	25	30
1	B	170/184 (92%)	161 (95%)	8 (5%)	1 (1%)	25	30
1	C	170/184 (92%)	161 (95%)	8 (5%)	1 (1%)	25	30
1	D	165/184 (90%)	159 (96%)	5 (3%)	1 (1%)	25	30
All	All	670/736 (91%)	641 (96%)	25 (4%)	4 (1%)	25	30

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	346	ASP
1	D	346	ASP
1	B	343	THR
1	C	343	THR

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	127/137 (93%)	116 (91%)	11 (9%)	10 12
1	B	131/137 (96%)	118 (90%)	13 (10%)	8 9
1	C	131/137 (96%)	121 (92%)	10 (8%)	13 16
1	D	127/137 (93%)	117 (92%)	10 (8%)	12 15
All	All	516/548 (94%)	472 (92%)	44 (8%)	10 12

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	317	ARG
1	D	295	LEU
1	C	332	GLU
1	C	406	ARG
1	D	376	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	308	GLN
1	C	324	ASN
1	C	439	GLN
1	B	324	ASN
1	A	308	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.

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	164/184 (89%)	-0.13	3 (1%)	68	75	23, 42, 99, 131	0
1	B	169/184 (91%)	0.23	21 (12%)	4	6	26, 44, 125, 143	1 (0%)
1	C	169/184 (91%)	0.29	19 (11%)	5	7	23, 44, 123, 145	1 (0%)
1	D	164/184 (89%)	-0.16	4 (2%)	59	66	23, 41, 97, 127	0
All	All	666/736 (90%)	0.06	47 (7%)	16	21	23, 43, 113, 145	2 (0%)

The worst 5 of 47 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	391	SER	14.1
1	C	390	ALA	13.0
1	A	457	ASP	7.0
1	B	392	HIS	5.9
1	C	460	ARG	5.6

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.