wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 11:40 PM GMT

PDB ID	:	1Y8T
Title	:	Crystal Structure of RV0983 from Mycobacterium tuberculosis- Proteolytically active form
Authors	:	Palaninathan, S.K.; MohamedMohaideen, N.N.; Sacchettini, J.C.; TB Structural Genomics Consortium (TBSGC)
Deposited on	:	2004-12-13
Resolution	:	2.00 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

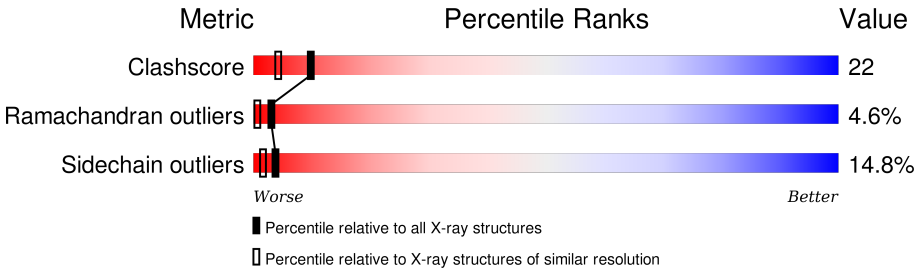
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.00 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	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	324	53% 24% 6% • 13%
1	B	324	52% 20% 8% • 16%
1	C	324	49% 22% 5% • 21%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5749 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 hypothetical protein Rv0983.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	283	Total	C	N	O	Se	0	0	0
			1911	1186	331	392	2			
1	B	272	Total	C	N	O	Se	0	0	0
			1817	1131	314	370	2			
1	C	255	Total	C	N	O	Se	0	0	0
			1703	1059	295	347	2			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	21	MSE	MET	MODIFIED RESIDUE	UNP O53896
A	176	MSE	MET	MODIFIED RESIDUE	UNP O53896
A	317	LEU	-	EXPRESSION TAG	UNP O53896
A	318	GLU	-	EXPRESSION TAG	UNP O53896
A	319	HIS	-	EXPRESSION TAG	UNP O53896
A	320	HIS	-	EXPRESSION TAG	UNP O53896
A	321	HIS	-	EXPRESSION TAG	UNP O53896
A	322	HIS	-	EXPRESSION TAG	UNP O53896
A	323	HIS	-	EXPRESSION TAG	UNP O53896
A	324	HIS	-	EXPRESSION TAG	UNP O53896
B	21	MSE	MET	MODIFIED RESIDUE	UNP O53896
B	176	MSE	MET	MODIFIED RESIDUE	UNP O53896
B	317	LEU	-	EXPRESSION TAG	UNP O53896
B	318	GLU	-	EXPRESSION TAG	UNP O53896
B	319	HIS	-	EXPRESSION TAG	UNP O53896
B	320	HIS	-	EXPRESSION TAG	UNP O53896
B	321	HIS	-	EXPRESSION TAG	UNP O53896
B	322	HIS	-	EXPRESSION TAG	UNP O53896
B	323	HIS	-	EXPRESSION TAG	UNP O53896
B	324	HIS	-	EXPRESSION TAG	UNP O53896
C	21	MSE	MET	MODIFIED RESIDUE	UNP O53896
C	176	MSE	MET	MODIFIED RESIDUE	UNP O53896
C	317	LEU	-	EXPRESSION TAG	UNP O53896

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	318	GLU	-	EXPRESSION TAG	UNP O53896
C	319	HIS	-	EXPRESSION TAG	UNP O53896
C	320	HIS	-	EXPRESSION TAG	UNP O53896
C	321	HIS	-	EXPRESSION TAG	UNP O53896
C	322	HIS	-	EXPRESSION TAG	UNP O53896
C	323	HIS	-	EXPRESSION TAG	UNP O53896
C	324	HIS	-	EXPRESSION TAG	UNP O53896

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	116	Total	O	0	0
			116	116		
2	B	119	Total	O	0	0
			119	119		
2	C	83	Total	O	0	0
			83	83		

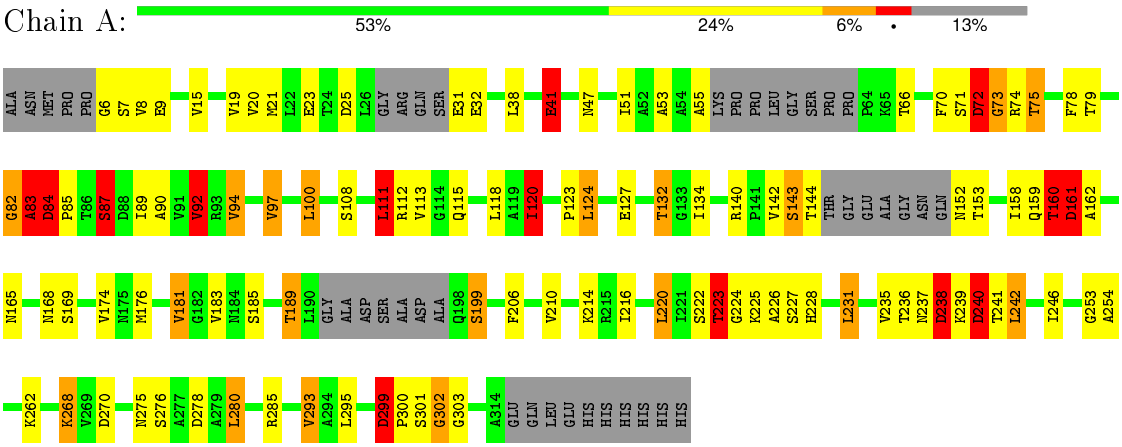
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

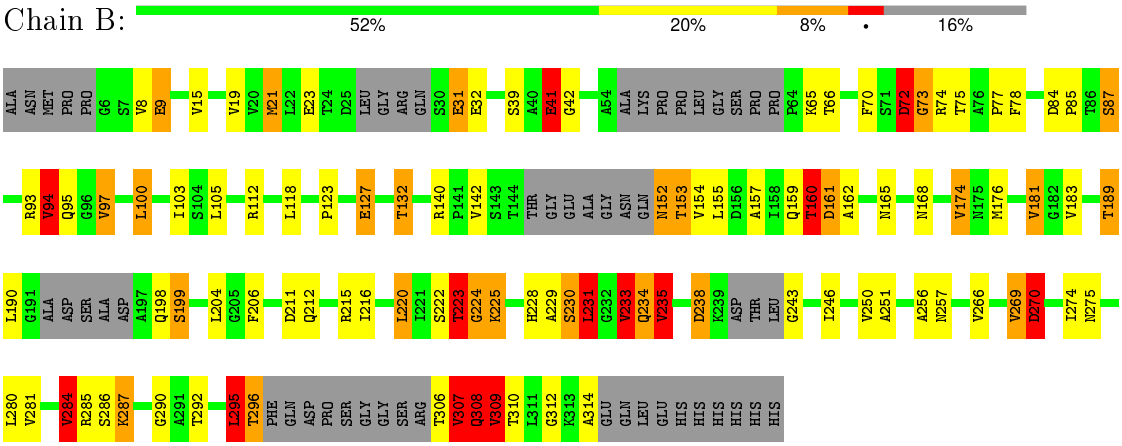
- Molecule 1: hypothetical protein Rv0983

Chain A:



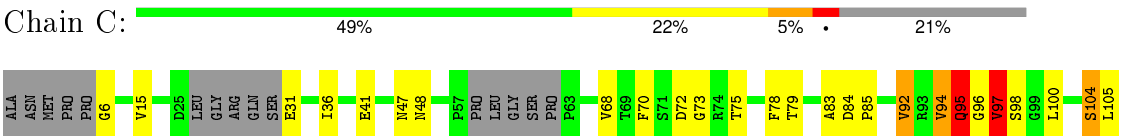
- Molecule 1: hypothetical protein Rv0983

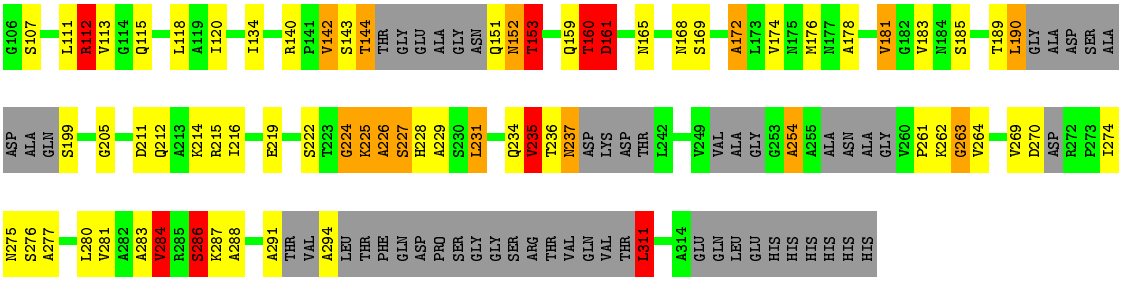
Chain B:



- Molecule 1: hypothetical protein Rv0983

Chain C:




G106 S107 L111 R112 V113 G114 Q115 L118 A119 I120 I134 R140 P141 V142 S143 T144 THR GLY GLU ALA GLY ASN Q151 N152 T153 Q159 T160 D161 N165 N168 S169 A172 L173 V174 M175 M176 M177 A178 V181 G182 V183 N184 S185 T189 L190 GLY ALA ASP SER ALA
ASP ALA GLN S199 G205 D211 Q212 A213 K214 R215 I216 E219 S222 T223 G224 K225 A226 S227 H228 A229 S230 L231 Q234 V235 T236 N237 ASP LYS ASP THR L242 V249 VAL ALA GLY G253 A254 A255 ALA ASN ALA GLY V260 P261 K262 G263 V264 V269 D270 ASP R272 P273 I274
N275 S276 A277 L280 V281 A282 A283 V284 R285 S286 K287 A288 A291 THR VAL A294 LEU THR PHE GLN ASP PRO SER GLY GLY SER ARG THR VAL GLN VAL THR L311 A314 GLU GLN LEU GLU HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	149.58Å 89.07Å 69.41Å 90.00° 97.55° 90.00°	Depositor
Resolution (Å)	76.70 – 2.00	Depositor
% Data completeness (in resolution range)	98.8 (76.70-2.00)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.227 , 0.280	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5749	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	1.57	23/1924 (1.2%)	1.52	30/2630 (1.1%)
1	B	1.86	24/1826 (1.3%)	1.84	32/2494 (1.3%)
1	C	1.39	14/1710 (0.8%)	1.48	22/2332 (0.9%)
All	All	1.62	61/5460 (1.1%)	1.62	84/7456 (1.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	9
1	B	0	8
1	C	1	6
All	All	1	23

The worst 5 of 61 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	41	GLU	CD-OE2	-30.41	0.92	1.25
1	B	41	GLU	CD-OE1	-22.51	1.00	1.25
1	B	41	GLU	CB-CG	14.68	1.80	1.52
1	A	31	GLU	CG-CD	12.29	1.70	1.51
1	A	41	GLU	CG-CD	10.71	1.68	1.51

The worst 5 of 84 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	41	GLU	OE1-CD-OE2	-49.09	64.40	123.30
1	C	161	ASP	CB-CG-OD1	-21.89	98.60	118.30
1	A	161	ASP	CB-CG-OD1	-19.09	101.11	118.30
1	B	41	GLU	CG-CD-OE1	17.86	154.02	118.30
1	C	161	ASP	CB-CG-OD2	17.66	134.19	118.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	C	235	VAL	CA

5 of 23 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	25	ASP	Peptide
1	A	71	SER	Peptide
1	A	73	GLY	Peptide
1	A	82	GLY	Peptide
1	A	83	ALA	Mainchain

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	1911	0	1922	81	0
1	B	1817	0	1831	96	0
1	C	1703	0	1686	67	0
2	A	116	0	0	4	0
2	B	119	0	0	11	0
2	C	83	0	0	1	0
All	All	5749	0	5439	241	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 22.

The worst 5 of 241 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:132:THR:CG2	1:A:132:THR:CB	1.76	1.63
1:C:112:ARG:CG	1:C:112:ARG:CD	1.74	1.61
1:B:132:THR:CG2	1:B:132:THR:CB	1.75	1.59
1:A:189:THR:CG2	1:A:189:THR:CB	1.75	1.56
1:B:41:GLU:CG	1:B:41:GLU:CB	1.80	1.54

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	273/324 (84%)	252 (92%)	11 (4%)	10 (4%)	4	1
1	B	258/324 (80%)	228 (88%)	16 (6%)	14 (5%)	2	0
1	C	234/324 (72%)	209 (89%)	14 (6%)	11 (5%)	3	0
All	All	765/972 (79%)	689 (90%)	41 (5%)	35 (5%)	3	1

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	ASP
1	A	223	THR
1	A	225	LYS
1	A	240	ASP
1	A	254	ALA

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	200/242 (83%)	171 (86%)	29 (14%)	4	2
1	B	187/242 (77%)	158 (84%)	29 (16%)	3	1
1	C	174/242 (72%)	149 (86%)	25 (14%)	4	2
All	All	561/726 (77%)	478 (85%)	83 (15%)	4	1

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	100	LEU
1	B	223	THR
1	C	222	SER
1	B	153	THR
1	B	190	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	159	GLN
1	B	168	ASN
1	C	115	GLN
1	A	237	ASN
1	B	95	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 carbohydrates 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

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.