



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 04:13 pm BST

PDB ID : 1UE7  
Title : Crystal structure of the single-stranded dna-binding protein from mycobacterium tuberculosis  
Authors : Saikrishnan, K.; Jeyakanthan, J.; Venkatesh, J.; Acharya, N.; Sekar, K.; Varshney, U.; Vijayan, M.; TB Structural Genomics Consortium (TBSGC)  
Deposited on : 2003-05-09  
Resolution : 3.20 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

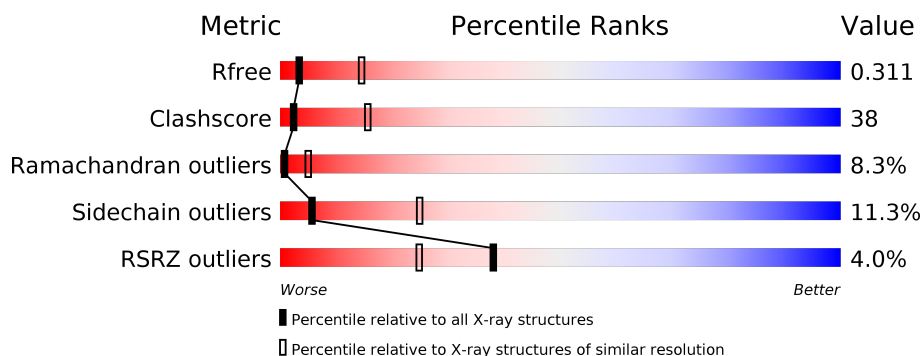
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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  $\geq 3$ , 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	164	<div> <div>4%</div> <div> <div></div> <div>37%</div> <div>22%</div> <div>5%</div> <div>35%</div> </div> </div>
1	B	164	<div> <div>2%</div> <div> <div></div> <div>29%</div> <div>24%</div> <div>7%</div> <div>40%</div> </div> </div>
1	C	164	<div> <div>2%</div> <div> <div></div> <div>23%</div> <div>30%</div> <div>5%</div> <div>40%</div> </div> </div>
1	D	164	<div> <div>2%</div> <div> <div></div> <div>25%</div> <div>28%</div> <div>• •</div> <div>42%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3036 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 Single-strand binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	106	Total	C	N	O	S	0	0	0
			758	468	141	148	1			
1	B	98	Total	C	N	O	S	0	0	0
			705	440	125	139	1			
1	C	98	Total	C	N	O	S	0	0	0
			712	445	126	140	1			
1	D	95	Total	C	N	O	S	0	0	0
			678	425	117	135	1			

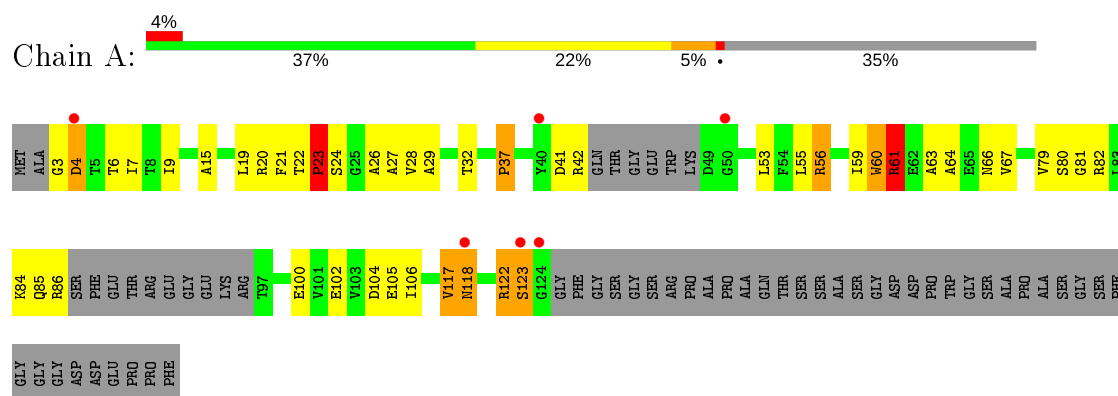
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	51	Total	O	0	0
			51	51		
2	B	18	Total	O	0	0
			18	18		
2	C	46	Total	O	0	0
			46	46		
2	D	68	Total	O	0	0
			68	68		

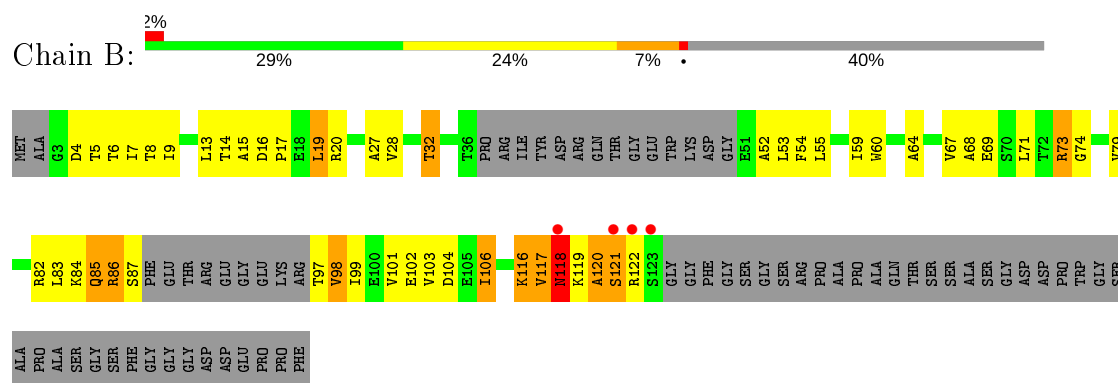
### 3 Residue-property plots [i](#)

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

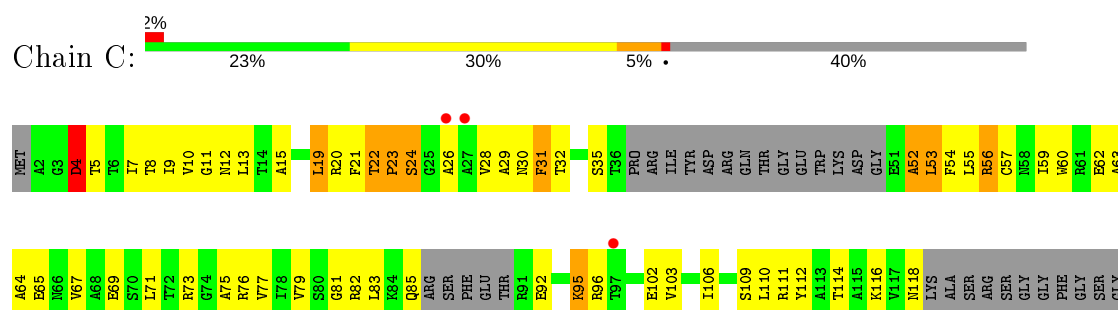
#### • Molecule 1: Single-strand binding protein



#### • Molecule 1: Single-strand binding protein

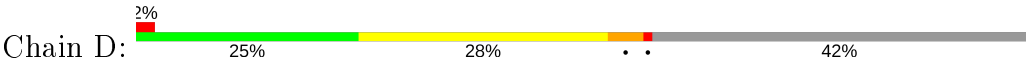


#### • Molecule 1: Single-strand binding protein



SER
ARG
PRO
ALA
PRO
ALA
GLN
THR
SER
SER
ALA
SER
GLY
ASP
ASP
TRP
GLY
SER
ALA
PRO
GLY
SER
GLY
GLY
ASP
ASP
GLU
PRO
PRO
PHE

● Molecule 1: Single-strand binding protein



MET	ALA	G3	D4	T5	T6	I7	T8	I9	N12	L13	T14	E18	L19	T22	A26	A29	N30	F31	T32	V33	A34	S35	T36	P37	R38	I1E	T1R	ASP	ARG	GLN	THR	GLY	GLU	TRP	LYS	ASP	GLY	E51	A52	L53	F54	L55	R56	C57	N58	I59	W60	R61	E62	A63	A64	V67
-----	-----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

L71	T72	R73	G74	A75	R76	V77	I78	V79	S80	C81	R82	L83	R84	Q85	R86	S87	PHE	GLU	THR	ARG	GLY	GLY	GLY	LYS	ARG	T97	V101	E102	V103	D104	E105	I106	S109	I110	R111	Y112	A115	K116	V117	M118	LYS	ALA	SER	ARG	SER	GLY	PHE	GLY	GLY	SER	GLY	ARG	PRO	ALA	PRO
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ALA	GLN	THR	SER	SER	ALA	SER	GLY	ASP	PRO	TRP	GLY	SER	ALA	ALA	SER	GLY	SER	PHE	GLY	GLY	GLY	ASP	ASP	GLU	PRO	PRO	PHE
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.22Å 116.72Å 177.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 3.20 28.77 – 3.20	Depositor EDS
% Data completeness (in resolution range)	94.3 (15.00-3.20) 94.3 (28.77-3.20)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.62 (at 3.17Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.235 , 0.313 0.236 , 0.311	Depositor DCC
$R_{free}$ test set	1042 reflections (10.31%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	64.0	Xtriage
Anisotropy	0.357	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 101.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	3036	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.56	2/764 (0.3%)	0.83	2/1037 (0.2%)
1	B	0.43	0/711	0.76	1/968 (0.1%)
1	C	0.48	0/718	0.87	1/976 (0.1%)
1	D	0.42	0/684	0.78	1/934 (0.1%)
All	All	0.48	2/2877 (0.1%)	0.81	5/3915 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	3	GLY	N-CA	5.97	1.55	1.46
1	A	3	GLY	C-O	5.44	1.32	1.23

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	118	ASN	N-CA-C	6.62	128.88	111.00
1	A	4	ASP	N-CA-C	6.59	128.79	111.00
1	D	115	ALA	N-CA-C	6.23	127.82	111.00
1	C	4	ASP	N-CA-C	5.59	126.10	111.00
1	A	37	PRO	N-CA-CB	5.03	109.33	103.30

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	758	0	725	60	0
1	B	705	0	690	50	0
1	C	712	0	704	59	0
1	D	678	0	655	49	0
2	A	51	0	0	1	0
2	B	18	0	0	2	0
2	C	46	0	0	5	0
2	D	68	0	0	1	0
All	All	3036	0	2774	213	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 38.

The worst 5 of 213 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:C:22:THR:HB	1:C:23:PRO:HD2	1.06	1.02
1:A:117:VAL:HG12	1:A:118:ASN:H	1.27	0.99
1:C:22:THR:CB	1:C:23:PRO:HD2	1.93	0.97
1:D:117:VAL:HG13	1:D:118:ASN:H	1.29	0.94
1:A:19:LEU:HD21	1:A:27:ALA:HB1	1.53	0.91

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	100/164 (61%)	75 (75%)	16 (16%)	9 (9%)	1	4
1	B	92/164 (56%)	74 (80%)	11 (12%)	7 (8%)	1	7
1	C	92/164 (56%)	72 (78%)	11 (12%)	9 (10%)	0	3
1	D	89/164 (54%)	75 (84%)	8 (9%)	6 (7%)	1	9

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	373/656 (57%)	296 (79%)	46 (12%)	31 (8%)	<b>1</b> <b>5</b>

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	ASP
1	A	23	PRO
1	A	24	SER
1	A	118	ASN
1	A	122	ARG

### 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	73/126 (58%)	65 (89%)	8 (11%)	<b>6</b> <b>26</b>
1	B	71/126 (56%)	63 (89%)	8 (11%)	<b>6</b> <b>25</b>
1	C	72/126 (57%)	63 (88%)	9 (12%)	<b>4</b> <b>21</b>
1	D	68/126 (54%)	61 (90%)	7 (10%)	<b>7</b> <b>29</b>
All	All	284/504 (56%)	252 (89%)	32 (11%)	<b>6</b> <b>25</b>

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	116	LYS
1	C	19	LEU
1	D	58	ASN
1	C	4	ASP
1	C	32	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	ASN
1	A	85	GLN
1	B	66	ASN
1	C	66	ASN
1	D	58	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 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	106/164 (64%)	-0.04	6 (5%)	23 13	17, 57, 102, 102	0
1	B	98/164 (59%)	-0.23	4 (4%)	37 24	14, 57, 97, 102	0
1	C	98/164 (59%)	-0.09	3 (3%)	49 32	17, 60, 102, 102	0
1	D	95/164 (57%)	-0.20	3 (3%)	47 31	22, 50, 97, 101	0
All	All	397/656 (60%)	-0.14	16 (4%)	38 25	14, 57, 101, 102	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	121	SER	3.6
1	A	40	TYR	3.4
1	D	37	PRO	3.3
1	A	124	GLY	3.2
1	B	118	ASN	3.1

### 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 carbohydrates 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.