



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 03:19 pm BST

PDB ID : 2Q74
Title : Mycobacterium tuberculosis SuhB
Authors : Brown, A.K.; Meng, G.; Ghadbane, H.; Besra, G.S.; Futterer, K.
Deposited on : 2007-06-06
Resolution : 2.60 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 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

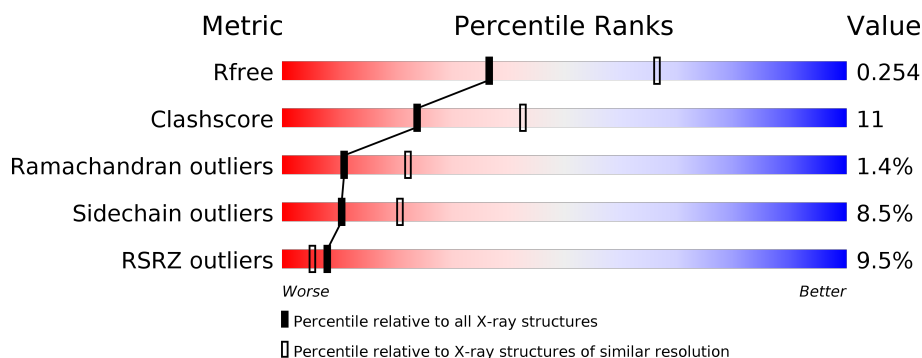
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 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	299	<div> <div>6%</div> <div> <div></div> <div>65%</div> <div>15%</div> <div>•</div> <div>18%</div> </div> </div>
1	B	299	<div> <div>9%</div> <div> <div></div> <div>66%</div> <div>12%</div> <div>•</div> <div>18%</div> </div> </div>
1	C	299	<div> <div>8%</div> <div> <div></div> <div>61%</div> <div>17%</div> <div>•</div> <div>19%</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5466 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 Inositol-1-monophosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	245	Total	C	N	O	S	0	0	0
			1784	1115	325	339	5			
1	B	245	Total	C	N	O	S	0	0	0
			1784	1115	325	338	6			
1	C	243	Total	C	N	O	S	0	0	0
			1774	1109	323	336	6			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	291	ALA	-	expression tag	UNP P65165
A	292	LEU	-	expression tag	UNP P65165
A	293	GLU	-	expression tag	UNP P65165
A	294	HIS	-	expression tag	UNP P65165
A	295	HIS	-	expression tag	UNP P65165
A	296	HIS	-	expression tag	UNP P65165
A	297	HIS	-	expression tag	UNP P65165
A	298	HIS	-	expression tag	UNP P65165
A	299	HIS	-	expression tag	UNP P65165
B	291	ALA	-	expression tag	UNP P65165
B	292	LEU	-	expression tag	UNP P65165
B	293	GLU	-	expression tag	UNP P65165
B	294	HIS	-	expression tag	UNP P65165
B	295	HIS	-	expression tag	UNP P65165
B	296	HIS	-	expression tag	UNP P65165
B	297	HIS	-	expression tag	UNP P65165
B	298	HIS	-	expression tag	UNP P65165
B	299	HIS	-	expression tag	UNP P65165
C	291	ALA	-	expression tag	UNP P65165
C	292	LEU	-	expression tag	UNP P65165
C	293	GLU	-	expression tag	UNP P65165
C	294	HIS	-	expression tag	UNP P65165
C	295	HIS	-	expression tag	UNP P65165

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	296	HIS	-	expression tag	UNP P65165
C	297	HIS	-	expression tag	UNP P65165
C	298	HIS	-	expression tag	UNP P65165
C	299	HIS	-	expression tag	UNP P65165

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	41	Total O 41 41	0	0
2	B	31	Total O 31 31	0	0
2	C	52	Total O 52 52	0	0

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 ($\text{RSRZ} > 2$). 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.

Chain A:

6% 65% 15% 18%

MET THR ARG ARG ARG PRO D5 D6 L11 R12 S13 V14 A15 F34 GLY ILE SER SER ARG ALA GLY ASP GLY ASP GLY ALA VAL ARG ALA LYS SER SER SER PRO THR D54 V58 L71 A72 Q73 I74 A75 P76 P79 I80 G82 E83 E84 G85 GLY GLY PRO ALA ASP VAL THR THR

PRO SER ASP R98 Y99 T100 W101 Y102 L103 D104 N111 I112 V127 Y132 V132 T149 L155 A158 R159 L164 V169 D170 E171 L172 A175 G176 T179 G180 F181 G182 Y183 S184 R185 R186 A189 L195 A196 H197 V198 V199 R206 R207 R222 A225 V226

Y227 D235 L241 E245 S253 THR PRO ARG ALA GLY A260 G261 L262 V263 A267 L271 R281 F282 N283 G284 L285 E286 PRO ILE PRO ASP ALA LEU GLU HIS HIS HIS HIS HIS HIS

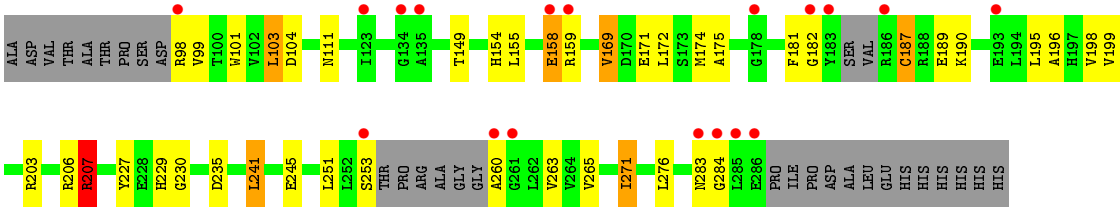
Chain B:

PRO	SER	ASP	R98	W101	V102	L103	D104	I123	L155	E158	R159	V169	D170	E171	L172	A175	G178	F181	G182	Y183	S184	V185	R186	C187	R188	E189	K190	Q191	A192	E193	L194	L195	A196	I197	V198	V199	R206	R207	L208	A225	Y226	Y227	D235	C236	A237	L244		
MET	THR	ARG	PRO	D5	K6	L11	R12	S13	V14	F34	GLY	ILE	SER	ARG	ALA	GLY	ASP	GLY	ASP	GLY	ALA	VAL	ARG	ALA	PRO	THR	D54	V58	I66	L71	A72	Q73	L74	R75	D78	P79	I80	L81	G82	G85	GLY	GLY	PRO	ALA	VAL	THR	ALA	THR

Chain C:

Sequence logo for Chain C. The y-axis represents information content in bits (0.00 to 0.10). The x-axis lists amino acids. A red dot above D5 indicates a mutation. A green bar at the top shows the distribution: 8% (red), 61% (green), 17% (yellow), and 19% (grey).

Amino Acid	Information Content (bits)
MET	0.00
THR	0.00
ARG	0.00
PRO	0.00
D5	0.08
N6	0.02
L11	0.00
R12	0.00
S13	0.00
V14	0.00
L18	0.00
R29	0.00
V33	0.00
F34	0.00
GLY	0.00
ILE	0.00
SER	0.00
ARG	0.00
ALA	0.00
GLY	0.00
ASP	0.00
ASP	0.00
GLY	0.00
ALA	0.00
VAL	0.00
ARG	0.00
ALA	0.00
LYS	0.00
SER	0.00
SER	0.00
PRO	0.00
THR	0.00
D64	0.00
P65	0.00
V66	0.00
T67	0.00
V68	0.00
L71	0.00
A72	0.00
Q73	0.00
L74	0.00
R75	0.00
P76	0.00
F79	0.00
I80	0.00
L81	0.00
G82	0.00
E83	0.00
E84	0.00
G85	0.00
GLY	0.00
PER	0.00



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	101.40Å 185.60Å 107.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.88 – 2.60 19.90 – 2.60	Depositor EDS
% Data completeness (in resolution range)	93.2 (19.88-2.60) 92.6 (19.90-2.60)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.28 (at 2.59Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.223 , 0.245 0.217 , 0.254	Depositor DCC
R_{free} test set	1469 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	68.9	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 71.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.015 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.021 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5466	wwPDB-VP
Average B, all atoms (Å ²)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% 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	0.56	0/1806	0.72	1/2457 (0.0%)
1	B	0.44	0/1806	0.66	1/2457 (0.0%)
1	C	0.58	0/1795	0.72	1/2440 (0.0%)
All	All	0.53	0/5407	0.70	3/7354 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	C	207	ARG	NE-CZ-NH1	7.12	123.86	120.30
1	B	207	ARG	NE-CZ-NH1	7.04	123.82	120.30
1	A	207	ARG	NE-CZ-NH1	6.91	123.75	120.30

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	1784	0	1769	38	0
1	B	1784	0	1769	38	1
1	C	1774	0	1764	50	0
2	A	41	0	0	5	0
2	B	31	0	0	9	1
2	C	52	0	0	16	0
All	All	5466	0	5302	118	1

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 11.

The worst 5 of 118 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:195:LEU:CD2	2:A:312:HOH:O	1.93	1.17
1:C:29:ARG:HD3	2:C:287:HOH:O	1.50	1.08
1:B:199:VAL:HG11	2:B:316:HOH:O	1.68	0.92
1:A:195:LEU:HD23	2:A:312:HOH:O	1.56	0.92
1:B:278:ALA:HB1	2:B:310:HOH:O	1.72	0.90

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:208:ILE:C	2:B:311:HOH:O[3_655]	2.17	0.03

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	237/299 (79%)	222 (94%)	11 (5%)	4 (2%)	9	18
1	B	237/299 (79%)	221 (93%)	12 (5%)	4 (2%)	9	18
1	C	233/299 (78%)	218 (94%)	13 (6%)	2 (1%)	17	35
All	All	707/897 (79%)	661 (94%)	36 (5%)	10 (1%)	11	22

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	6	ASN
1	B	6	ASN
1	C	6	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	185	VAL
1	B	188	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	172/217 (79%)	157 (91%)	15 (9%)	10	20
1	B	172/217 (79%)	159 (92%)	13 (8%)	13	26
1	C	172/217 (79%)	156 (91%)	16 (9%)	9	17
All	All	516/651 (79%)	472 (92%)	44 (8%)	10	21

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

Mol	Chain	Res	Type
1	B	103	LEU
1	B	207	ARG
1	C	207	ARG
1	B	155	LEU
1	B	169	VAL

Some 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	232	GLN
1	B	280	GLN
1	C	232	GLN
1	B	126	GLN
1	C	126	GLN

5.3.3 RNA ⓘ

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, 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	245/299 (81%)	0.31	19 (7%) 13 9	52, 69, 87, 94	0
1	B	245/299 (81%)	0.48	26 (10%) 6 4	52, 69, 87, 92	0
1	C	243/299 (81%)	0.27	25 (10%) 6 4	52, 69, 86, 93	0
All	All	733/897 (81%)	0.35	70 (9%) 8 5	52, 69, 87, 94	0

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	185	VAL	10.5
1	B	197	HIS	5.5
1	A	260	ALA	5.5
1	A	184	SER	5.4
1	A	183	TYR	5.2

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.