WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 5, 2024 – 03:35 pm BST

PDB ID	:	2BI0
Title	:	RV0216, A conserved hypothetical protein from Mycobacterium tuberculosis that is essential for bacterial survival during infection, has a double hotdogfold
Authors	:	Castell, A.; Johansson, P.; Unge, T.; Jones, T.A.; Backbro, K.
Deposited on	:	2005-01-20
Resolution	:	1.90 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

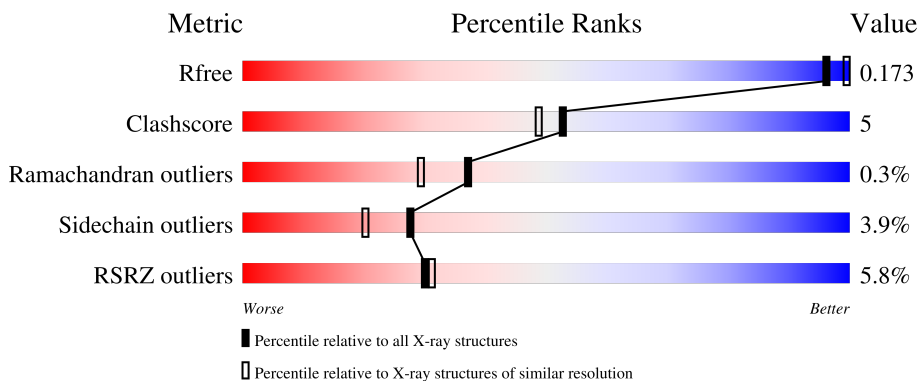|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance ⓘ

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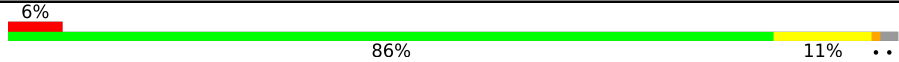*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	Percentile Ranks	Value
Rfree		0.173
Clashscore		5
Ramachandran outliers		0.3%
Sidechain outliers		3.9%
RSRZ outliers		5.8%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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 <=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	337	6% — 86% — 11% ··

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 2805 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 RV0216.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	330	Total 2482	C 1565	N 455	O 455	S 4	Se 3	0	0	0

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Cl 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	322	Total 322	O 322	0	0

3 Residue-property plots ⓘ

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: HYPOTHETICAL PROTEIN RV0216
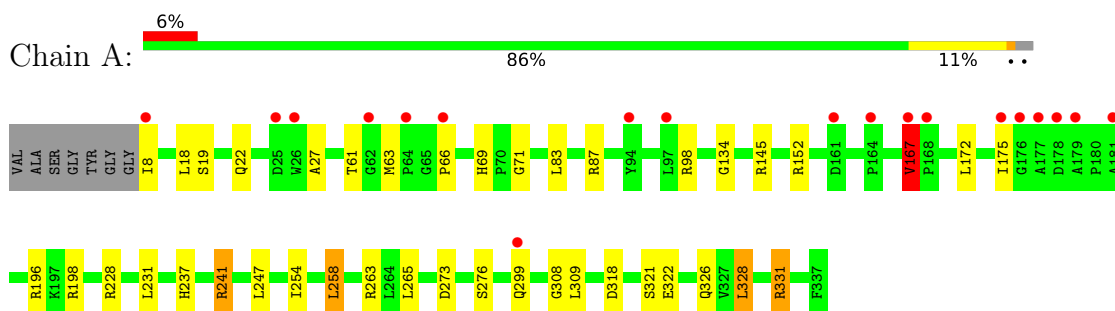
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	78.75Å 78.75Å 180.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	91.29 – 1.90 90.16 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (91.29-1.90) 100.0 (90.16-1.90)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.89 (at 1.49Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.173 , 0.217 0.175 , 0.173	Depositor DCC
R_{free} test set	1382 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å^2)	17.9	Xtriage
Anisotropy	0.016	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 43.0	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2805	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.01% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5　Model quality ⓘ

5.1　Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CL

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	0.93	1/2542 (0.0%)	1.08	16/3468 (0.5%)

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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	241	ARG	CB-CG	-6.52	1.34	1.52

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	196	ARG	NE-CZ-NH2	-21.52	109.54	120.30
1	A	196	ARG	NE-CZ-NH1	15.21	127.91	120.30
1	A	145	ARG	NE-CZ-NH2	-11.61	114.49	120.30
1	A	145	ARG	NE-CZ-NH1	8.68	124.64	120.30
1	A	241	ARG	CG-CD-NE	-7.32	96.44	111.80
1	A	196	ARG	CD-NE-CZ	6.80	133.12	123.60
1	A	263	ARG	NE-CZ-NH1	6.60	123.60	120.30
1	A	231	LEU	CB-CG-CD2	6.58	122.18	111.00
1	A	196	ARG	CG-CD-NE	-6.50	98.16	111.80
1	A	228	ARG	NE-CZ-NH1	6.28	123.44	120.30
1	A	258	LEU	CB-CG-CD2	6.20	121.54	111.00
1	A	145	ARG	CB-CG-CD	-5.97	96.06	111.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	309	LEU	CA-CB-CG	-5.91	101.70	115.30
1	A	152	ARG	NE-CZ-NH2	-5.48	117.56	120.30
1	A	87	ARG	NE-CZ-NH1	5.25	122.93	120.30
1	A	198	ARG	NE-CZ-NH2	-5.08	117.76	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	167	VAL	Peptide

5.2 Too-close contacts ⓘ

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	2482	0	2449	24	0
2	A	1	0	0	0	0
3	A	322	0	0	12	2
All	All	2805	0	2449	24	2

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 5.

All (24) 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:254:ILE:HD11	3:A:2267:HOH:O	1.79	0.83
1:A:167:VAL:HG13	1:A:167:VAL:O	1.86	0.75
1:A:167:VAL:O	1:A:167:VAL:CG1	2.38	0.71
1:A:69:HIS:CD2	1:A:71:GLY:H	2.16	0.63
1:A:237:HIS:HD2	1:A:247:LEU:O	1.81	0.63
1:A:276:SER:CB	3:A:2267:HOH:O	2.49	0.61
1:A:276:SER:HB3	3:A:2267:HOH:O	2.02	0.59
1:A:18:LEU:HA	1:A:22:GLN:HE22	1.68	0.59
1:A:237:HIS:HE1	3:A:2225:HOH:O	1.87	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:299:GLN:HG3	3:A:2289:HOH:O	2.08	0.53
1:A:8:ILE:HD11	3:A:2014:HOH:O	2.11	0.50
1:A:322:GLU:OE1	3:A:2310:HOH:O	2.19	0.49
1:A:27:ALA:HB3	3:A:2126:HOH:O	2.14	0.45
1:A:321:SER:O	1:A:321:SER:OG	2.29	0.45
1:A:331:ARG:HD3	3:A:2319:HOH:O	2.17	0.43
1:A:318:ASP:HB2	3:A:2304:HOH:O	2.18	0.43
1:A:326:GLN:NE2	3:A:2317:HOH:O	2.52	0.43
1:A:19:SER:H	1:A:22:GLN:HE21	1.67	0.42
1:A:254:ILE:HG13	1:A:328:LEU:HD11	2.00	0.42
1:A:66:PRO:CB	3:A:2079:HOH:O	2.67	0.42
1:A:61:THR:OG1	1:A:63:MSE:HB2	2.19	0.41
1:A:134:GLY:HA3	1:A:175:ILE:HD13	2.03	0.41
1:A:308:GLY:CA	1:A:331:ARG:HH22	2.33	0.41
1:A:18:LEU:HA	1:A:22:GLN:NE2	2.35	0.40

All (2) 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)
3:A:2160:HOH:O	3:A:2160:HOH:O[12_555]	1.20	1.00
3:A:2015:HOH:O	3:A:2031:HOH:O[3_665]	2.17	0.03

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	328/337 (97%)	323 (98%)	4 (1%)	1 (0%)	37	29

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	167	VAL

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	254/254 (100%)	244 (96%)	10 (4%)	27	20

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	83	LEU
1	A	98	ARG
1	A	167	VAL
1	A	172	LEU
1	A	241	ARG
1	A	258	LEU
1	A	265	LEU
1	A	273	ASP
1	A	328	LEU
1	A	331	ARG

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

Mol	Chain	Res	Type
1	A	22	GLN
1	A	69	HIS
1	A	237	HIS
1	A	279	HIS
1	A	326	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

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 ⓘ

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(Å²)	Q<0.9
1	A	327/337 (97%)	-0.15	19 (5%)	30	31	8, 16, 35, 55	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	167	VAL	5.1
1	A	161	ASP	4.0
1	A	26	TRP	3.9
1	A	177	ALA	3.4
1	A	178	ASP	3.1
1	A	25	ASP	3.0
1	A	168	PRO	2.8
1	A	179	ALA	2.7
1	A	181	ALA	2.7
1	A	175	ILE	2.7
1	A	164	PRO	2.5
1	A	64	PRO	2.5
1	A	62	GLY	2.5
1	A	97	LEU	2.4
1	A	8	ILE	2.4
1	A	176	GLY	2.3
1	A	94	TYR	2.3
1	A	299	GLN	2.3
1	A	66	PRO	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CL	A	1338	1/1	0.96	0.18	27,27,27,27	1

6.5 Other polymers ⓘ

There are no such residues in this entry.