



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 07:29 am BST

PDB ID : 4PZU  
Title : Crystal structure of a putative uncharacterize protein Rv3404c and likely sugar N-formyltransferase from Mycobacterium tuberculosis  
Authors : Seattle Structural Genomics Center for Infectious Disease; Seattle Structural Genomics Center for Infectious Disease (SSGCID)  
Deposited on : 2014-03-31  
Resolution : 2.10 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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
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

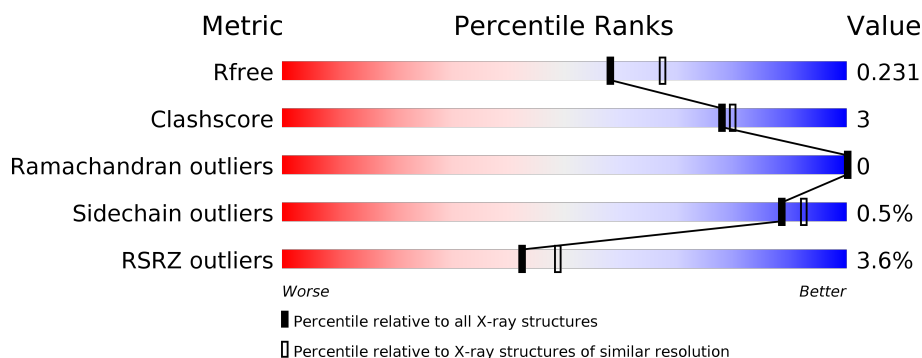
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	256	<div> <div>4%</div> <div> <div></div> <div>84%</div> <div>8%</div> <div>8%</div> </div> </div>
1	B	256	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>•</div> <div>9%</div> </div> </div>
1	C	256	<div> <div>2%</div> <div> <div></div> <div>86%</div> <div>5%</div> <div>8%</div> </div> </div>
1	D	256	<div> <div>3%</div> <div> <div></div> <div>84%</div> <div>7%</div> <div>8%</div> </div> </div>
1	E	256	<div> <div>•</div> <div> <div></div> <div>87%</div> <div>•</div> <div>9%</div> </div> </div>
1	F	256	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>7%</div> <div>8%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	256	<div><div></div><div>5%</div><div>86%</div><div>5%</div><div>9%</div></div>
1	H	256	<div><div></div><div>8%</div><div>84%</div><div>6%</div><div>10%</div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15706 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 Uncharacterized protein Rv3404c/MT3512.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	235	Total	C	N	O	S	0	0	0
			1836	1160	329	341	6			
1	B	232	Total	C	N	O	S	0	1	0
			1837	1162	333	336	6			
1	C	235	Total	C	N	O	S	0	1	0
			1853	1171	337	339	6			
1	D	235	Total	C	N	O	S	0	1	0
			1853	1171	336	340	6			
1	E	232	Total	C	N	O	S	0	1	0
			1825	1154	323	341	7			
1	F	235	Total	C	N	O	S	0	0	0
			1840	1164	332	338	6			
1	G	232	Total	C	N	O	S	0	0	0
			1791	1135	317	333	6			
1	H	231	Total	C	N	O	S	0	0	0
			1810	1144	323	337	6			

There are 176 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	EXPRESSION TAG	UNP P65073
A	-19	ALA	-	EXPRESSION TAG	UNP P65073
A	-18	HIS	-	EXPRESSION TAG	UNP P65073
A	-17	HIS	-	EXPRESSION TAG	UNP P65073
A	-16	HIS	-	EXPRESSION TAG	UNP P65073
A	-15	HIS	-	EXPRESSION TAG	UNP P65073
A	-14	HIS	-	EXPRESSION TAG	UNP P65073
A	-13	HIS	-	EXPRESSION TAG	UNP P65073
A	-12	MET	-	EXPRESSION TAG	UNP P65073
A	-11	GLY	-	EXPRESSION TAG	UNP P65073
A	-10	THR	-	EXPRESSION TAG	UNP P65073
A	-9	LEU	-	EXPRESSION TAG	UNP P65073
A	-8	GLU	-	EXPRESSION TAG	UNP P65073

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	ALA	-	EXPRESSION TAG	UNP P65073
A	-6	GLN	-	EXPRESSION TAG	UNP P65073
A	-5	THR	-	EXPRESSION TAG	UNP P65073
A	-4	GLN	-	EXPRESSION TAG	UNP P65073
A	-3	GLY	-	EXPRESSION TAG	UNP P65073
A	-2	PRO	-	EXPRESSION TAG	UNP P65073
A	-1	GLY	-	EXPRESSION TAG	UNP P65073
A	0	SER	-	EXPRESSION TAG	UNP P65073
A	2	VAL	-	EXPRESSION TAG	UNP P65073
B	-20	MET	-	EXPRESSION TAG	UNP P65073
B	-19	ALA	-	EXPRESSION TAG	UNP P65073
B	-18	HIS	-	EXPRESSION TAG	UNP P65073
B	-17	HIS	-	EXPRESSION TAG	UNP P65073
B	-16	HIS	-	EXPRESSION TAG	UNP P65073
B	-15	HIS	-	EXPRESSION TAG	UNP P65073
B	-14	HIS	-	EXPRESSION TAG	UNP P65073
B	-13	HIS	-	EXPRESSION TAG	UNP P65073
B	-12	MET	-	EXPRESSION TAG	UNP P65073
B	-11	GLY	-	EXPRESSION TAG	UNP P65073
B	-10	THR	-	EXPRESSION TAG	UNP P65073
B	-9	LEU	-	EXPRESSION TAG	UNP P65073
B	-8	GLU	-	EXPRESSION TAG	UNP P65073
B	-7	ALA	-	EXPRESSION TAG	UNP P65073
B	-6	GLN	-	EXPRESSION TAG	UNP P65073
B	-5	THR	-	EXPRESSION TAG	UNP P65073
B	-4	GLN	-	EXPRESSION TAG	UNP P65073
B	-3	GLY	-	EXPRESSION TAG	UNP P65073
B	-2	PRO	-	EXPRESSION TAG	UNP P65073
B	-1	GLY	-	EXPRESSION TAG	UNP P65073
B	0	SER	-	EXPRESSION TAG	UNP P65073
B	2	VAL	-	EXPRESSION TAG	UNP P65073
C	-20	MET	-	EXPRESSION TAG	UNP P65073
C	-19	ALA	-	EXPRESSION TAG	UNP P65073
C	-18	HIS	-	EXPRESSION TAG	UNP P65073
C	-17	HIS	-	EXPRESSION TAG	UNP P65073
C	-16	HIS	-	EXPRESSION TAG	UNP P65073
C	-15	HIS	-	EXPRESSION TAG	UNP P65073
C	-14	HIS	-	EXPRESSION TAG	UNP P65073
C	-13	HIS	-	EXPRESSION TAG	UNP P65073
C	-12	MET	-	EXPRESSION TAG	UNP P65073
C	-11	GLY	-	EXPRESSION TAG	UNP P65073
C	-10	THR	-	EXPRESSION TAG	UNP P65073

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-9	LEU	-	EXPRESSION TAG	UNP P65073
C	-8	GLU	-	EXPRESSION TAG	UNP P65073
C	-7	ALA	-	EXPRESSION TAG	UNP P65073
C	-6	GLN	-	EXPRESSION TAG	UNP P65073
C	-5	THR	-	EXPRESSION TAG	UNP P65073
C	-4	GLN	-	EXPRESSION TAG	UNP P65073
C	-3	GLY	-	EXPRESSION TAG	UNP P65073
C	-2	PRO	-	EXPRESSION TAG	UNP P65073
C	-1	GLY	-	EXPRESSION TAG	UNP P65073
C	0	SER	-	EXPRESSION TAG	UNP P65073
C	2	VAL	-	EXPRESSION TAG	UNP P65073
D	-20	MET	-	EXPRESSION TAG	UNP P65073
D	-19	ALA	-	EXPRESSION TAG	UNP P65073
D	-18	HIS	-	EXPRESSION TAG	UNP P65073
D	-17	HIS	-	EXPRESSION TAG	UNP P65073
D	-16	HIS	-	EXPRESSION TAG	UNP P65073
D	-15	HIS	-	EXPRESSION TAG	UNP P65073
D	-14	HIS	-	EXPRESSION TAG	UNP P65073
D	-13	HIS	-	EXPRESSION TAG	UNP P65073
D	-12	MET	-	EXPRESSION TAG	UNP P65073
D	-11	GLY	-	EXPRESSION TAG	UNP P65073
D	-10	THR	-	EXPRESSION TAG	UNP P65073
D	-9	LEU	-	EXPRESSION TAG	UNP P65073
D	-8	GLU	-	EXPRESSION TAG	UNP P65073
D	-7	ALA	-	EXPRESSION TAG	UNP P65073
D	-6	GLN	-	EXPRESSION TAG	UNP P65073
D	-5	THR	-	EXPRESSION TAG	UNP P65073
D	-4	GLN	-	EXPRESSION TAG	UNP P65073
D	-3	GLY	-	EXPRESSION TAG	UNP P65073
D	-2	PRO	-	EXPRESSION TAG	UNP P65073
D	-1	GLY	-	EXPRESSION TAG	UNP P65073
D	0	SER	-	EXPRESSION TAG	UNP P65073
D	2	VAL	-	EXPRESSION TAG	UNP P65073
E	-20	MET	-	EXPRESSION TAG	UNP P65073
E	-19	ALA	-	EXPRESSION TAG	UNP P65073
E	-18	HIS	-	EXPRESSION TAG	UNP P65073
E	-17	HIS	-	EXPRESSION TAG	UNP P65073
E	-16	HIS	-	EXPRESSION TAG	UNP P65073
E	-15	HIS	-	EXPRESSION TAG	UNP P65073
E	-14	HIS	-	EXPRESSION TAG	UNP P65073
E	-13	HIS	-	EXPRESSION TAG	UNP P65073
E	-12	MET	-	EXPRESSION TAG	UNP P65073

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-11	GLY	-	EXPRESSION TAG	UNP P65073
E	-10	THR	-	EXPRESSION TAG	UNP P65073
E	-9	LEU	-	EXPRESSION TAG	UNP P65073
E	-8	GLU	-	EXPRESSION TAG	UNP P65073
E	-7	ALA	-	EXPRESSION TAG	UNP P65073
E	-6	GLN	-	EXPRESSION TAG	UNP P65073
E	-5	THR	-	EXPRESSION TAG	UNP P65073
E	-4	GLN	-	EXPRESSION TAG	UNP P65073
E	-3	GLY	-	EXPRESSION TAG	UNP P65073
E	-2	PRO	-	EXPRESSION TAG	UNP P65073
E	-1	GLY	-	EXPRESSION TAG	UNP P65073
E	0	SER	-	EXPRESSION TAG	UNP P65073
E	2	VAL	-	EXPRESSION TAG	UNP P65073
F	-20	MET	-	EXPRESSION TAG	UNP P65073
F	-19	ALA	-	EXPRESSION TAG	UNP P65073
F	-18	HIS	-	EXPRESSION TAG	UNP P65073
F	-17	HIS	-	EXPRESSION TAG	UNP P65073
F	-16	HIS	-	EXPRESSION TAG	UNP P65073
F	-15	HIS	-	EXPRESSION TAG	UNP P65073
F	-14	HIS	-	EXPRESSION TAG	UNP P65073
F	-13	HIS	-	EXPRESSION TAG	UNP P65073
F	-12	MET	-	EXPRESSION TAG	UNP P65073
F	-11	GLY	-	EXPRESSION TAG	UNP P65073
F	-10	THR	-	EXPRESSION TAG	UNP P65073
F	-9	LEU	-	EXPRESSION TAG	UNP P65073
F	-8	GLU	-	EXPRESSION TAG	UNP P65073
F	-7	ALA	-	EXPRESSION TAG	UNP P65073
F	-6	GLN	-	EXPRESSION TAG	UNP P65073
F	-5	THR	-	EXPRESSION TAG	UNP P65073
F	-4	GLN	-	EXPRESSION TAG	UNP P65073
F	-3	GLY	-	EXPRESSION TAG	UNP P65073
F	-2	PRO	-	EXPRESSION TAG	UNP P65073
F	-1	GLY	-	EXPRESSION TAG	UNP P65073
F	0	SER	-	EXPRESSION TAG	UNP P65073
F	2	VAL	-	EXPRESSION TAG	UNP P65073
G	-20	MET	-	EXPRESSION TAG	UNP P65073
G	-19	ALA	-	EXPRESSION TAG	UNP P65073
G	-18	HIS	-	EXPRESSION TAG	UNP P65073
G	-17	HIS	-	EXPRESSION TAG	UNP P65073
G	-16	HIS	-	EXPRESSION TAG	UNP P65073
G	-15	HIS	-	EXPRESSION TAG	UNP P65073
G	-14	HIS	-	EXPRESSION TAG	UNP P65073

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	-13	HIS	-	EXPRESSION TAG	UNP P65073
G	-12	MET	-	EXPRESSION TAG	UNP P65073
G	-11	GLY	-	EXPRESSION TAG	UNP P65073
G	-10	THR	-	EXPRESSION TAG	UNP P65073
G	-9	LEU	-	EXPRESSION TAG	UNP P65073
G	-8	GLU	-	EXPRESSION TAG	UNP P65073
G	-7	ALA	-	EXPRESSION TAG	UNP P65073
G	-6	GLN	-	EXPRESSION TAG	UNP P65073
G	-5	THR	-	EXPRESSION TAG	UNP P65073
G	-4	GLN	-	EXPRESSION TAG	UNP P65073
G	-3	GLY	-	EXPRESSION TAG	UNP P65073
G	-2	PRO	-	EXPRESSION TAG	UNP P65073
G	-1	GLY	-	EXPRESSION TAG	UNP P65073
G	0	SER	-	EXPRESSION TAG	UNP P65073
G	2	VAL	-	EXPRESSION TAG	UNP P65073
H	-20	MET	-	EXPRESSION TAG	UNP P65073
H	-19	ALA	-	EXPRESSION TAG	UNP P65073
H	-18	HIS	-	EXPRESSION TAG	UNP P65073
H	-17	HIS	-	EXPRESSION TAG	UNP P65073
H	-16	HIS	-	EXPRESSION TAG	UNP P65073
H	-15	HIS	-	EXPRESSION TAG	UNP P65073
H	-14	HIS	-	EXPRESSION TAG	UNP P65073
H	-13	HIS	-	EXPRESSION TAG	UNP P65073
H	-12	MET	-	EXPRESSION TAG	UNP P65073
H	-11	GLY	-	EXPRESSION TAG	UNP P65073
H	-10	THR	-	EXPRESSION TAG	UNP P65073
H	-9	LEU	-	EXPRESSION TAG	UNP P65073
H	-8	GLU	-	EXPRESSION TAG	UNP P65073
H	-7	ALA	-	EXPRESSION TAG	UNP P65073
H	-6	GLN	-	EXPRESSION TAG	UNP P65073
H	-5	THR	-	EXPRESSION TAG	UNP P65073
H	-4	GLN	-	EXPRESSION TAG	UNP P65073
H	-3	GLY	-	EXPRESSION TAG	UNP P65073
H	-2	PRO	-	EXPRESSION TAG	UNP P65073
H	-1	GLY	-	EXPRESSION TAG	UNP P65073
H	0	SER	-	EXPRESSION TAG	UNP P65073
H	2	VAL	-	EXPRESSION TAG	UNP P65073

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	A	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	B	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		
2	E	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	E	1	Total C O 4 2 2	0	0
2	E	1	Total C O 4 2 2	0	0
2	F	1	Total C O 4 2 2	0	0
2	F	1	Total C O 4 2 2	0	0
2	F	1	Total C O 4 2 2	0	0
2	G	1	Total C O 4 2 2	0	0
2	G	1	Total C O 4 2 2	0	0
2	H	1	Total C O 4 2 2	0	0

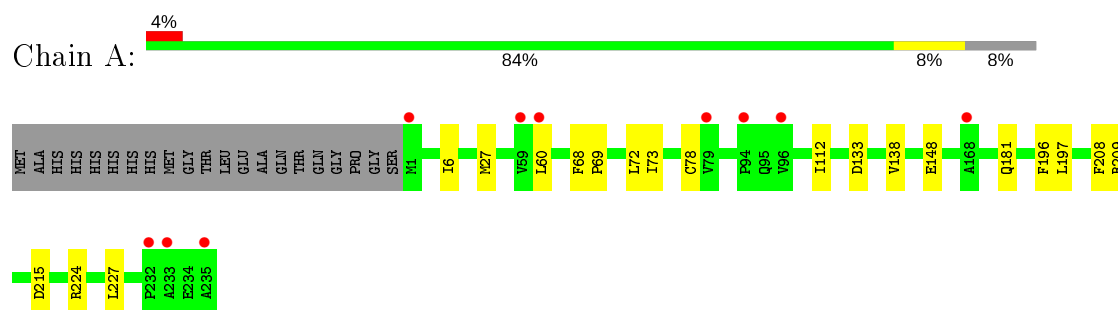
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	153	Total O 153 153	0	0
3	B	136	Total O 136 136	0	0
3	C	158	Total O 158 158	0	0
3	D	139	Total O 139 139	0	0
3	E	156	Total O 156 156	0	0
3	F	92	Total O 92 92	0	0
3	G	90	Total O 90 90	0	0
3	H	49	Total O 49 49	0	0

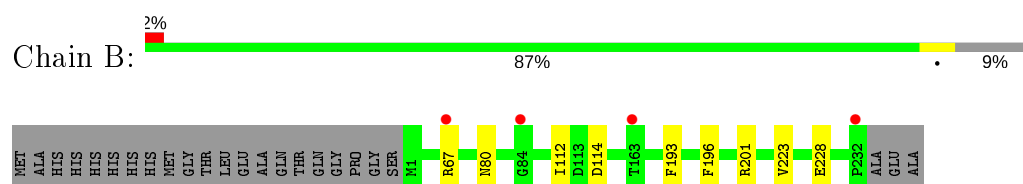
### 3 Residue-property plots [i](#)

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.

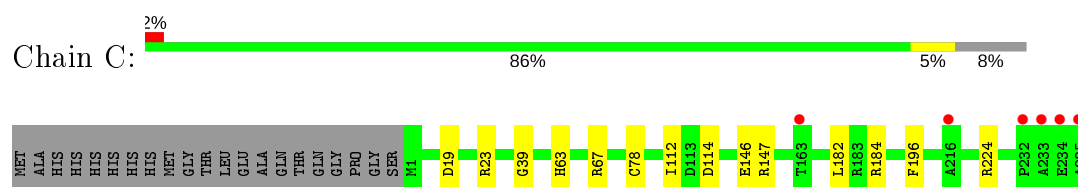
- Molecule 1: Uncharacterized protein Rv3404c/MT3512



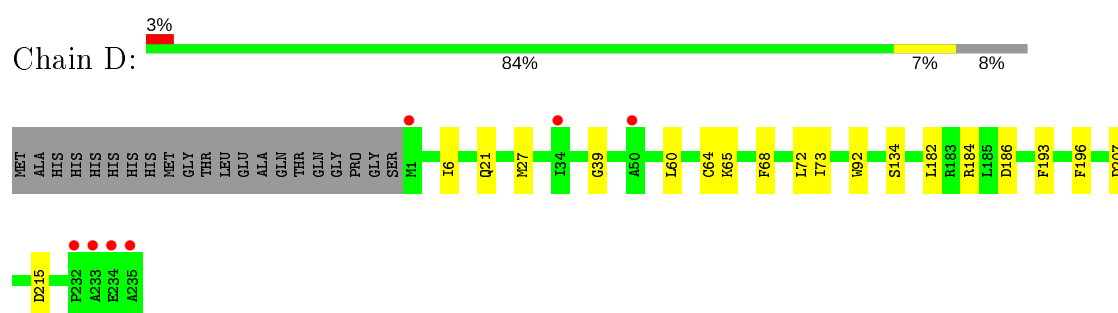
- Molecule 1: Uncharacterized protein Rv3404c/MT3512



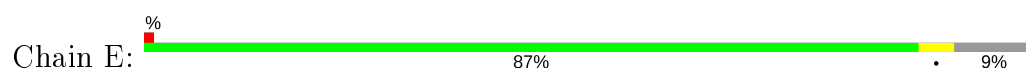
- Molecule 1: Uncharacterized protein Rv3404c/MT3512



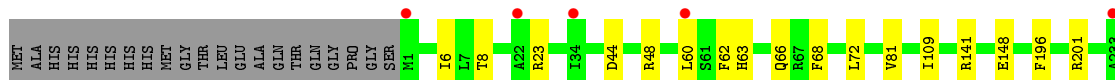
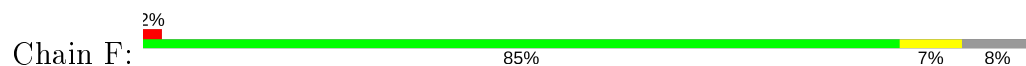
- Molecule 1: Uncharacterized protein Rv3404c/MT3512



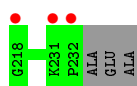
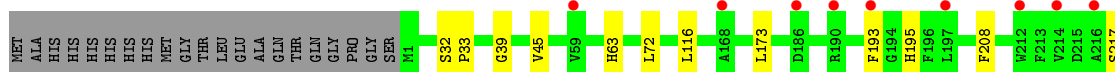
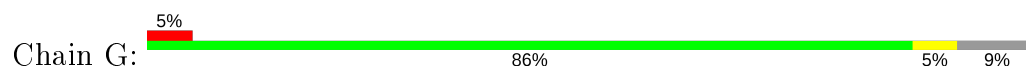
- Molecule 1: Uncharacterized protein Rv3404c/MT3512



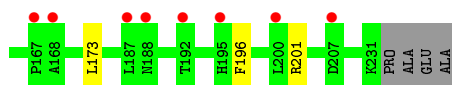
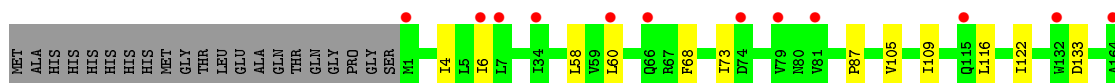
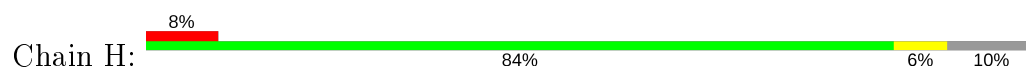
- Molecule 1: Uncharacterized protein Rv3404c/MT3512



- Molecule 1: Uncharacterized protein Rv3404c/MT3512



- Molecule 1: Uncharacterized protein Rv3404c/MT3512



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.08 Å 136.45 Å 99.97 Å 90.00° 103.74° 90.00°	Depositor
Resolution (Å)	20.00 – 2.10 20.00 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-2.10) 99.9 (20.00-2.10)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.35 (at 2.09 Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, $R_{free}$	0.187 , 0.231 0.187 , 0.231	Depositor DCC
$R_{free}$ test set	6535 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.6	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 42.3	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	0.96	EDS
Total number of atoms	15706	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% 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](#)

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

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	0.61	0/1879	0.73	0/2558
1	B	0.60	0/1883	0.74	1/2560 (0.0%)
1	C	0.65	0/1899	0.75	1/2581 (0.0%)
1	D	0.62	0/1899	0.72	1/2581 (0.0%)
1	E	0.64	0/1871	0.72	0/2548
1	F	0.57	0/1883	0.70	2/2561 (0.1%)
1	G	0.55	0/1834	0.68	0/2502
1	H	0.53	0/1852	0.67	0/2522
All	All	0.60	0/15000	0.71	5/20413 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	182	LEU	CB-CG-CD1	-6.44	100.06	111.00
1	F	141	ARG	NE-CZ-NH1	-6.14	117.23	120.30
1	B	201	ARG	NE-CZ-NH2	-5.78	117.41	120.30
1	C	182	LEU	CB-CG-CD1	-5.48	101.69	111.00
1	F	201	ARG	NE-CZ-NH1	5.01	122.80	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	1836	0	1750	19	0
1	B	1837	0	1776	10	0
1	C	1853	0	1790	19	0
1	D	1853	0	1788	18	0
1	E	1825	0	1740	11	0
1	F	1840	0	1770	14	0
1	G	1791	0	1684	14	0
1	H	1810	0	1724	13	0
2	A	12	0	18	1	0
2	B	12	0	18	0	0
2	C	12	0	18	0	0
2	D	16	0	24	3	0
2	E	12	0	18	0	0
2	F	12	0	18	0	0
2	G	8	0	12	0	0
2	H	4	0	6	0	0
3	A	153	0	0	3	0
3	B	136	0	0	0	0
3	C	158	0	0	5	0
3	D	139	0	0	1	0
3	E	156	0	0	2	0
3	F	92	0	0	0	0
3	G	90	0	0	1	0
3	H	49	0	0	2	0
All	All	15706	0	14154	85	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 3.

The worst 5 of 85 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:196:PHE:HD2	1:B:196:PHE:CD2	1.44	1.32
1:E:196:PHE:HD2	1:F:196:PHE:CD2	1.47	1.31
1:E:196:PHE:CD2	1:F:196:PHE:HD2	1.56	1.23
1:A:196:PHE:CD2	1:B:196:PHE:CD2	2.31	1.17
1:A:196:PHE:CD2	1:B:196:PHE:HD2	1.61	1.17

There are no symmetry-related clashes.

## 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	233/256 (91%)	230 (99%)	3 (1%)	0	100	100
1	B	231/256 (90%)	224 (97%)	7 (3%)	0	100	100
1	C	234/256 (91%)	229 (98%)	5 (2%)	0	100	100
1	D	234/256 (91%)	231 (99%)	3 (1%)	0	100	100
1	E	231/256 (90%)	226 (98%)	5 (2%)	0	100	100
1	F	233/256 (91%)	229 (98%)	4 (2%)	0	100	100
1	G	230/256 (90%)	226 (98%)	4 (2%)	0	100	100
1	H	229/256 (90%)	227 (99%)	2 (1%)	0	100	100
All	All	1855/2048 (91%)	1822 (98%)	33 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	191/217 (88%)	190 (100%)	1 (0%)	88	92
1	B	194/217 (89%)	194 (100%)	0	100	100
1	C	194/217 (89%)	194 (100%)	0	100	100
1	D	194/217 (89%)	193 (100%)	1 (0%)	88	92
1	E	193/217 (89%)	191 (99%)	2 (1%)	76	82
1	F	192/217 (88%)	191 (100%)	1 (0%)	88	92

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	184/217 (85%)	183 (100%)	1 (0%)	88	92
1	H	190/217 (88%)	189 (100%)	1 (0%)	88	92
All	All	1532/1736 (88%)	1525 (100%)	7 (0%)	88	92

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	230	GLU
1	H	105	VAL
1	F	66	GLN
1	D	207	ASP
1	G	217	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	21	GLN
1	D	174	ASN
1	H	82	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

22 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and

the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	EDO	H	301	-	3,3,3	0.52	0	2,2,2	0.33	0
2	EDO	F	301	-	3,3,3	0.44	0	2,2,2	0.64	0
2	EDO	C	301	-	3,3,3	0.41	0	2,2,2	0.60	0
2	EDO	F	303	-	3,3,3	0.60	0	2,2,2	0.21	0
2	EDO	E	302	-	3,3,3	0.48	0	2,2,2	0.49	0
2	EDO	C	303	-	3,3,3	0.43	0	2,2,2	0.38	0
2	EDO	G	302	-	3,3,3	0.47	0	2,2,2	0.53	0
2	EDO	D	302	-	3,3,3	0.68	0	2,2,2	0.28	0
2	EDO	B	301	-	3,3,3	0.42	0	2,2,2	0.18	0
2	EDO	E	303	-	3,3,3	0.35	0	2,2,2	0.71	0
2	EDO	B	303	-	3,3,3	0.36	0	2,2,2	0.72	0
2	EDO	D	303	-	3,3,3	0.62	0	2,2,2	0.16	0
2	EDO	G	301	-	3,3,3	0.90	0	2,2,2	0.40	0
2	EDO	C	302	-	3,3,3	0.46	0	2,2,2	0.30	0
2	EDO	A	303	-	3,3,3	0.53	0	2,2,2	0.34	0
2	EDO	D	304	-	3,3,3	0.59	0	2,2,2	0.26	0
2	EDO	D	301	-	3,3,3	0.64	0	2,2,2	0.41	0
2	EDO	A	301	-	3,3,3	0.50	0	2,2,2	0.48	0
2	EDO	E	301	-	3,3,3	0.51	0	2,2,2	0.53	0
2	EDO	A	302	-	3,3,3	0.37	0	2,2,2	0.35	0
2	EDO	F	302	-	3,3,3	0.51	0	2,2,2	0.52	0
2	EDO	B	302	-	3,3,3	0.55	0	2,2,2	0.39	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	H	301	-	-	0/1/1/1	-
2	EDO	F	301	-	-	0/1/1/1	-
2	EDO	C	301	-	-	1/1/1/1	-
2	EDO	F	303	-	-	1/1/1/1	-
2	EDO	E	302	-	-	0/1/1/1	-
2	EDO	C	303	-	-	0/1/1/1	-
2	EDO	G	302	-	-	1/1/1/1	-
2	EDO	D	302	-	-	1/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EDO	B	301	-	-	1/1/1/1	-
2	EDO	E	303	-	-	0/1/1/1	-
2	EDO	B	303	-	-	0/1/1/1	-
2	EDO	D	303	-	-	1/1/1/1	-
2	EDO	G	301	-	-	0/1/1/1	-
2	EDO	C	302	-	-	0/1/1/1	-
2	EDO	A	303	-	-	0/1/1/1	-
2	EDO	D	304	-	-	1/1/1/1	-
2	EDO	D	301	-	-	0/1/1/1	-
2	EDO	A	301	-	-	1/1/1/1	-
2	EDO	E	301	-	-	0/1/1/1	-
2	EDO	A	302	-	-	1/1/1/1	-
2	EDO	F	302	-	-	0/1/1/1	-
2	EDO	B	302	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	302	EDO	O1-C1-C2-O2
2	D	304	EDO	O1-C1-C2-O2
2	D	302	EDO	O1-C1-C2-O2
2	D	303	EDO	O1-C1-C2-O2
2	A	302	EDO	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	302	EDO	1	0
2	D	303	EDO	1	0
2	D	304	EDO	1	0
2	A	302	EDO	1	0

## 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 [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	235/256 (91%)	0.07	10 (4%) 35 41	22, 33, 53, 64	0
1	B	232/256 (90%)	-0.11	4 (1%) 70 74	24, 32, 48, 68	0
1	C	235/256 (91%)	-0.03	6 (2%) 56 61	23, 33, 53, 90	0
1	D	235/256 (91%)	-0.02	7 (2%) 50 56	23, 35, 58, 72	0
1	E	232/256 (90%)	-0.10	3 (1%) 77 80	23, 32, 49, 71	0
1	F	235/256 (91%)	0.15	6 (2%) 56 61	24, 43, 67, 78	0
1	G	232/256 (90%)	0.20	12 (5%) 27 32	23, 41, 80, 113	0
1	H	231/256 (90%)	0.56	20 (8%) 10 13	34, 54, 80, 97	0
All	All	1867/2048 (91%)	0.09	68 (3%) 42 49	22, 37, 66, 113	0

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	232	PRO	5.3
1	A	233	ALA	5.2
1	C	233	ALA	5.0
1	C	235	ALA	4.9
1	D	233	ALA	4.1

### 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](#)

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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	EDO	F	303	4/4	0.80	0.22	57,57,58,60	0
2	EDO	G	301	4/4	0.81	0.22	40,49,49,53	0
2	EDO	D	303	4/4	0.86	0.14	48,49,54,58	0
2	EDO	A	303	4/4	0.90	0.15	52,54,55,59	0
2	EDO	H	301	4/4	0.91	0.17	58,60,60,64	0
2	EDO	D	304	4/4	0.92	0.15	48,49,50,52	0
2	EDO	G	302	4/4	0.93	0.21	55,58,58,60	0
2	EDO	B	302	4/4	0.93	0.14	37,48,49,54	0
2	EDO	B	303	4/4	0.94	0.14	45,47,47,49	0
2	EDO	C	303	4/4	0.94	0.11	51,53,54,56	0
2	EDO	F	302	4/4	0.94	0.14	46,48,49,51	0
2	EDO	D	302	4/4	0.94	0.10	38,43,43,47	0
2	EDO	C	301	4/4	0.95	0.15	32,33,34,35	0
2	EDO	D	301	4/4	0.95	0.12	32,32,33,38	0
2	EDO	E	301	4/4	0.95	0.09	32,35,36,37	0
2	EDO	A	302	4/4	0.95	0.14	38,46,47,51	0
2	EDO	C	302	4/4	0.95	0.14	38,47,48,48	0
2	EDO	E	303	4/4	0.95	0.15	42,45,50,51	0
2	EDO	F	301	4/4	0.96	0.15	35,39,42,42	0
2	EDO	A	301	4/4	0.97	0.12	31,32,33,33	0
2	EDO	B	301	4/4	0.98	0.13	31,32,33,35	0
2	EDO	E	302	4/4	0.98	0.11	31,35,35,37	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.