



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 07:34 pm BST

PDB ID : 1GN4  
Title : H145E mutant of Mycobacterium tuberculosis iron-superoxide dismutase.  
Authors : Bunting, K.A.; Cooper, J.B.; Badasso, M.O.; Tickle, I.J.; Newton, M.; Wood, S.P.; Zhang, Y.; Young, D.B.  
Deposited on : 2001-10-02  
Resolution : 2.50 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

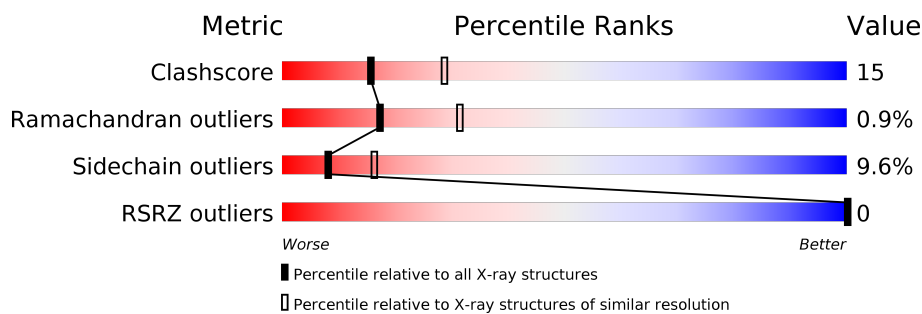
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	207	
1	B	207	
1	C	207	
1	D	207	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6560 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 SUPEROXIDE DISMUTASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	198	Total	C	N	O	S	0	0	0
			1567	1011	267	288	1			
1	B	198	Total	C	N	O	S	0	0	0
			1567	1011	267	288	1			
1	C	198	Total	C	N	O	S	0	0	0
			1567	1011	267	288	1			
1	D	198	Total	C	N	O	S	0	0	0
			1567	1011	267	288	1			

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mn	0	0
			1	1		
2	A	1	Total	Mn	0	0
			1	1		
2	D	1	Total	Mn	0	0
			1	1		
2	C	1	Total	Mn	0	0
			1	1		

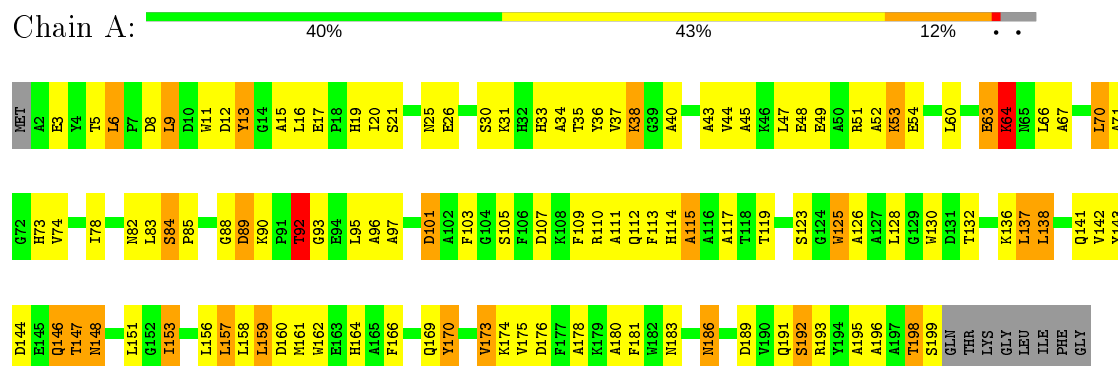
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	70	Total	O	0	0
			70	70		
3	B	81	Total	O	0	0
			81	81		
3	C	69	Total	O	0	0
			69	69		
3	D	68	Total	O	0	0
			68	68		

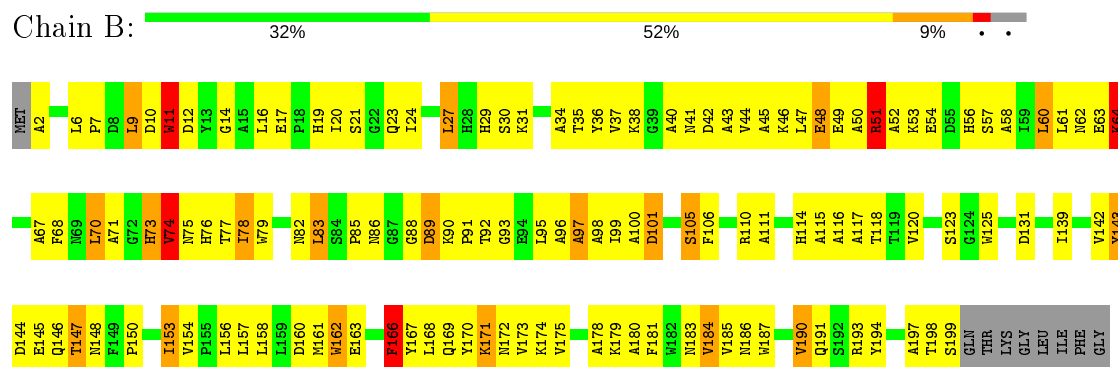
### 3 Residue-property plots

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.

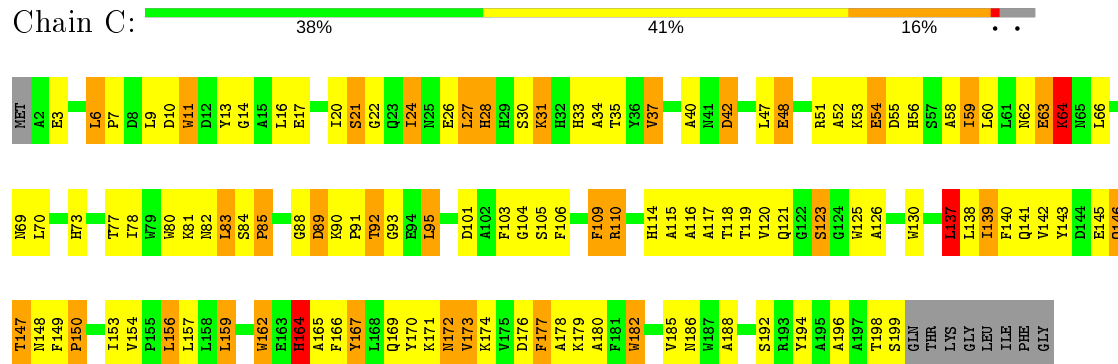
#### • Molecule 1: SUPEROXIDE DISMUTASE



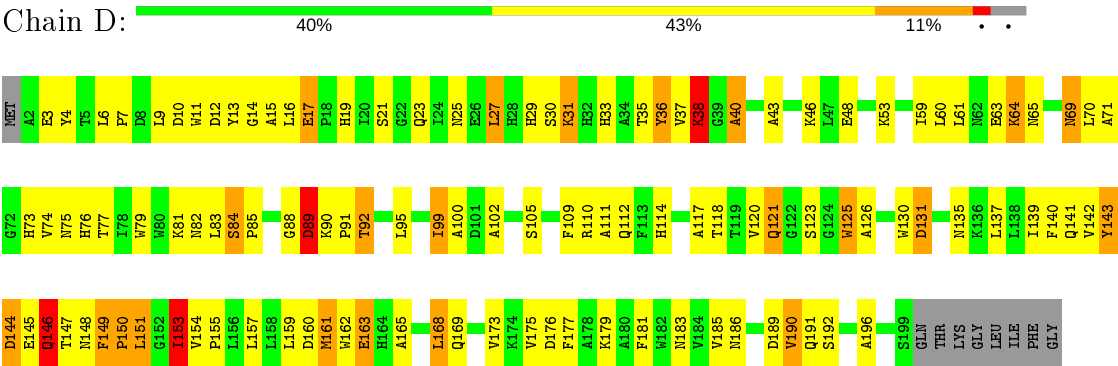
#### • Molecule 1: SUPEROXIDE DISMUTASE



#### • Molecule 1: SUPEROXIDE DISMUTASE



● Molecule 1: SUPEROXIDE DISMUTASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.66Å 102.86Å 73.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 12.54 – 2.50	Depositor EDS
% Data completeness (in resolution range)	94.4 (20.00-2.50) 95.0 (12.54-2.50)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.44 (at 2.50Å)	Xtriage
Refinement program	CCP4	Depositor
R, $R_{free}$	0.172 , 0.225 0.153 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.4	Xtriage
Anisotropy	0.434	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 86.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.024 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6560	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% 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: MN

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	1.75	44/1615 (2.7%)	2.24	87/2203 (3.9%)
1	B	1.76	53/1615 (3.3%)	2.24	95/2203 (4.3%)
1	C	1.75	50/1615 (3.1%)	2.25	98/2203 (4.4%)
1	D	1.62	39/1615 (2.4%)	2.19	84/2203 (3.8%)
All	All	1.72	186/6460 (2.9%)	2.23	364/8812 (4.1%)

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	B	1	0
1	C	1	0
1	D	1	0
All	All	3	0

The worst 5 of 186 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	117	ALA	CA-CB	10.24	1.74	1.52
1	C	177	PHE	N-CA	10.22	1.66	1.46
1	A	175	VAL	N-CA	9.98	1.66	1.46
1	C	150	PRO	CA-CB	9.91	1.73	1.53
1	A	114	HIS	N-CA	9.87	1.66	1.46

The worst 5 of 364 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	40	ALA	N-CA-CB	-12.73	92.28	110.10

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	92	THR	CA-CB-CG2	-11.59	96.18	112.40
1	C	153	ILE	CA-CB-CG2	10.55	131.99	110.90
1	D	117	ALA	N-CA-CB	-10.32	95.66	110.10
1	B	67	ALA	N-CA-CB	-10.23	95.78	110.10

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	20	ILE	CB
1	C	153	ILE	CB
1	D	161	MET	CA

There are no planarity outliers.

## 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	1567	0	1488	44	0
1	B	1567	0	1488	44	0
1	C	1567	0	1488	53	0
1	D	1567	0	1488	46	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	70	0	0	1	0
3	B	81	0	0	0	0
3	C	69	0	0	5	0
3	D	68	0	0	3	0
All	All	6560	0	5952	180	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 15.

The worst 5 of 180 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:88:GLY:O	1:C:89:ASP:HB2	1.39	1.07
1:B:95:LEU:HD21	1:B:190:VAL:HG12	1.41	1.02
1:A:88:GLY:O	1:A:89:ASP:HB2	1.66	0.92
1:C:88:GLY:O	1:C:89:ASP:CB	2.22	0.86
1:B:95:LEU:CD2	1:B:190:VAL:HG12	2.06	0.86

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	196/207 (95%)	177 (90%)	17 (9%)	2 (1%)	15	28
1	B	196/207 (95%)	184 (94%)	10 (5%)	2 (1%)	15	28
1	C	196/207 (95%)	175 (89%)	20 (10%)	1 (0%)	29	48
1	D	196/207 (95%)	182 (93%)	12 (6%)	2 (1%)	15	28
All	All	784/828 (95%)	718 (92%)	59 (8%)	7 (1%)	17	31

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	89	ASP
1	A	89	ASP
1	B	148	ASN
1	D	148	ASN
1	B	166	PHE

### 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	156/163 (96%)	140 (90%)	16 (10%)	7	14
1	B	156/163 (96%)	141 (90%)	15 (10%)	8	16
1	C	156/163 (96%)	142 (91%)	14 (9%)	9	19
1	D	156/163 (96%)	141 (90%)	15 (10%)	8	16
All	All	624/652 (96%)	564 (90%)	60 (10%)	8	16

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	105	SER
1	C	30	SER
1	D	146	GLN
1	B	199	SER
1	C	64	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	114	HIS
1	C	62	ASN
1	D	62	ASN
1	B	82	ASN
1	D	19	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

Of 4 ligands modelled in this entry, 4 are 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, 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	198/207 (95%)	-1.06	0 100 100	4, 20, 49, 89	0
1	B	198/207 (95%)	-0.91	0 100 100	7, 27, 61, 86	0
1	C	198/207 (95%)	-1.02	0 100 100	5, 20, 53, 84	0
1	D	198/207 (95%)	-0.98	0 100 100	7, 23, 57, 110	0
All	All	792/828 (95%)	-0.99	0 100 100	4, 22, 56, 110	0

There are no RSRZ outliers to report.

### 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 carbohydrates 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, 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(Å <sup>2</sup> )	Q<0.9
2	MN	B	1200	1/1	0.98	0.04	17,17,17,17	0
2	MN	D	1200	1/1	0.99	0.06	18,18,18,18	0
2	MN	A	1200	1/1	0.99	0.07	13,13,13,13	0
2	MN	C	1200	1/1	0.99	0.06	18,18,18,18	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.