



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 20, 2023 – 04:07 PM JST

PDB ID : 7D8R  
Title : MITF HLHLZ structure  
Authors : Guo, M.; Fang, P.; Wang, J.  
Deposited on : 2020-10-09  
Resolution : 3.00 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 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  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

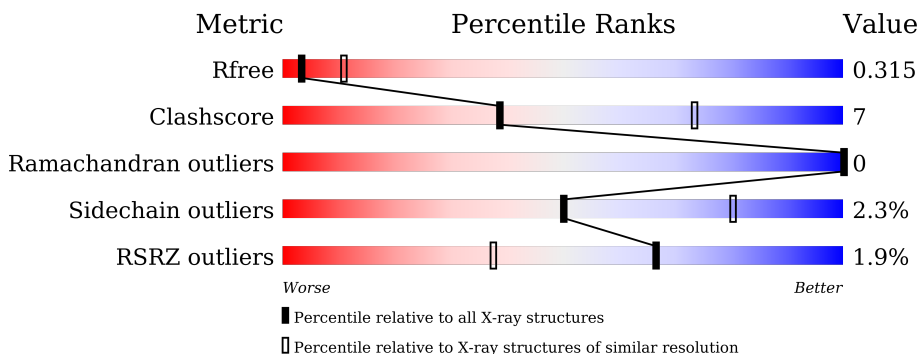
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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  $\geq 3$ , 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	182	 2% 82% 13% . .
1	B	182	 4% 77% 15% 8%
1	C	182	 % 78% 14% 8%
1	D	182	 % 76% 18% . .

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 4798 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 Microphthalmia-associated transcription factor, Methionyl-tRNA synthetase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	174	Total	C	N	O	S	0	0	0
			1218	771	208	238	1			
1	B	168	Total	C	N	O	S	0	0	0
			1171	742	205	223	1			
1	C	168	Total	C	N	O	S	0	0	0
			1154	737	197	219	1			
1	D	174	Total	C	N	O	S	0	0	0
			1217	771	205	240	1			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	322	SER	-	expression tag	UNP O75030
A	323	LEU	-	expression tag	UNP O75030
A	500	ALA	-	expression tag	UNP O66738
A	501	ALA	-	expression tag	UNP O66738
A	502	LEU	-	expression tag	UNP O66738
A	503	GLU	-	expression tag	UNP O66738
B	322	SER	-	expression tag	UNP O75030
B	323	LEU	-	expression tag	UNP O75030
B	500	ALA	-	expression tag	UNP O66738
B	501	ALA	-	expression tag	UNP O66738
B	502	LEU	-	expression tag	UNP O66738
B	503	GLU	-	expression tag	UNP O66738
C	322	SER	-	expression tag	UNP O75030
C	323	LEU	-	expression tag	UNP O75030
C	500	ALA	-	expression tag	UNP O66738
C	501	ALA	-	expression tag	UNP O66738
C	502	LEU	-	expression tag	UNP O66738
C	503	GLU	-	expression tag	UNP O66738
D	322	SER	-	expression tag	UNP O75030
D	323	LEU	-	expression tag	UNP O75030

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	500	ALA	-	expression tag	UNP O66738
D	501	ALA	-	expression tag	UNP O66738
D	502	LEU	-	expression tag	UNP O66738
D	503	GLU	-	expression tag	UNP O66738

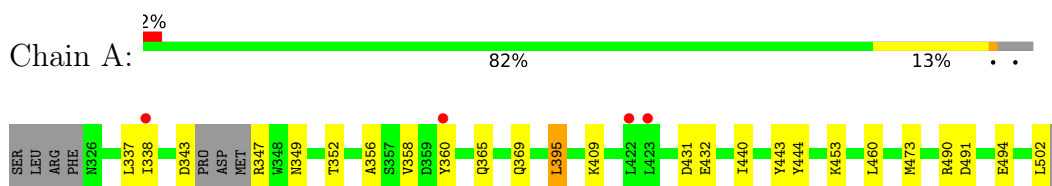
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	8	Total O 8 8	0	0
2	B	12	Total O 12 12	0	0
2	C	11	Total O 11 11	0	0
2	D	7	Total O 7 7	0	0

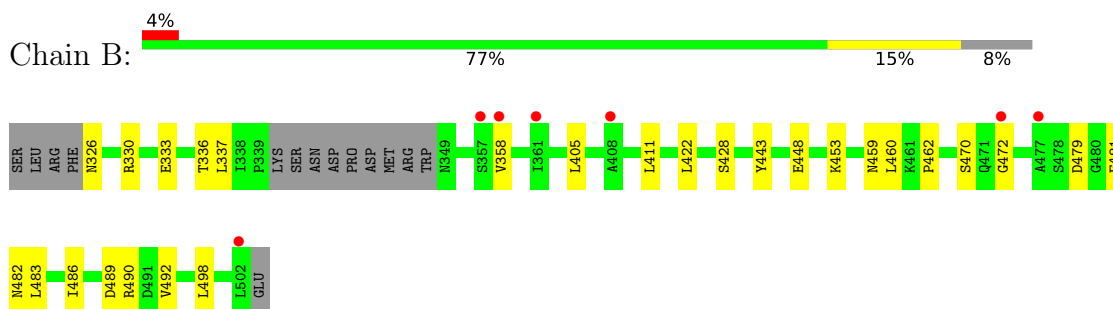
### 3 Residue-property plots [i](#)

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.

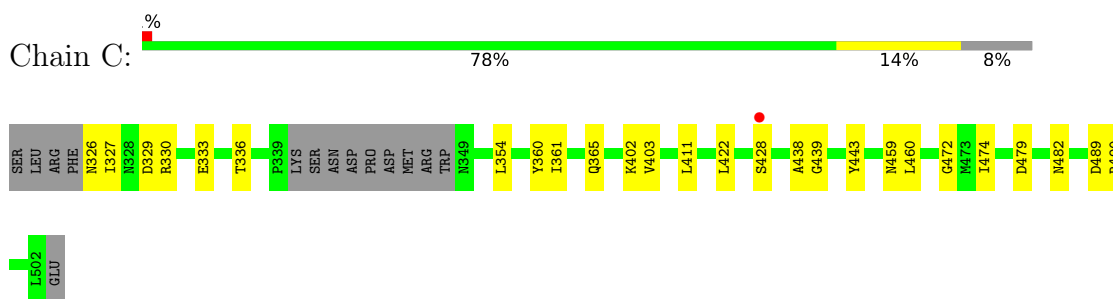
- Molecule 1: Microphthalmia-associated transcription factor,Methionyl-tRNA synthetase beta subunit



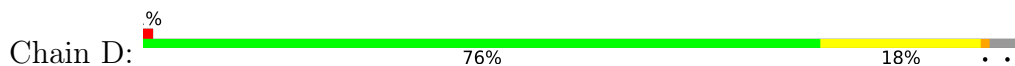
- Molecule 1: Microphthalmia-associated transcription factor,Methionyl-tRNA synthetase beta subunit

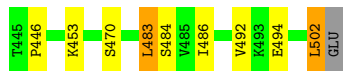


- Molecule 1: Microphthalmia-associated transcription factor,Methionyl-tRNA synthetase beta subunit



- Molecule 1: Microphthalmia-associated transcription factor,Methionyl-tRNA synthetase beta subunit





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	41.91Å 78.12Å 80.59Å 93.18° 105.03° 105.62°	Depositor
Resolution (Å)	38.69 – 3.00 38.69 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.2 (38.69-3.00) 98.0 (38.69-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.51 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.285 , 0.318 0.281 , 0.315	Depositor DCC
$R_{free}$ test set	949 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	106.1	Xtriage
Anisotropy	0.636	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 123.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.459 for h,-h-k,-h-l 0.016 for -h,h+l,h+k 0.011 for -h,-l,-k	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4798	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.94% 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	0.23	0/1227	0.39	0/1674
1	B	0.24	0/1180	0.43	0/1610
1	C	0.23	0/1163	0.43	0/1590
1	D	0.23	0/1227	0.42	0/1677
All	All	0.23	0/4797	0.42	0/6551

There are no bond length outliers.

There are no bond angle outliers.

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	1218	0	1146	17	0
1	B	1171	0	1107	24	0
1	C	1154	0	1084	17	0
1	D	1217	0	1141	29	0
2	A	8	0	0	0	0
2	B	12	0	0	2	0
2	C	11	0	0	2	0
2	D	7	0	0	1	0
All	All	4798	0	4478	69	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 7.



The worst 5 of 69 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:358:VAL:HG23	1:B:337:LEU:HD11	1.63	0.80
1:B:459:ASN:HD22	1:D:395:LEU:HB2	1.45	0.80
1:D:381:LEU:O	1:D:385:ASN:ND2	2.23	0.70
1:B:411:LEU:HD11	1:B:428:SER:HB2	1.75	0.68
1:C:402:LYS:NZ	2:C:601:HOH:O	2.26	0.68

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	170/182 (93%)	165 (97%)	5 (3%)	0	100	100
1	B	164/182 (90%)	162 (99%)	2 (1%)	0	100	100
1	C	164/182 (90%)	162 (99%)	2 (1%)	0	100	100
1	D	170/182 (93%)	168 (99%)	2 (1%)	0	100	100
All	All	668/728 (92%)	657 (98%)	11 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	112/158 (71%)	109 (97%)	3 (3%)	44	77
1	B	106/158 (67%)	105 (99%)	1 (1%)	78	92
1	C	102/158 (65%)	100 (98%)	2 (2%)	55	83
1	D	113/158 (72%)	109 (96%)	4 (4%)	36	71
All	All	433/632 (68%)	423 (98%)	10 (2%)	50	80

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	389	LEU
1	D	483	LEU
1	D	502	LEU
1	B	448	GLU
1	C	330	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 monosaccharides 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

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	174/182 (95%)	-0.20	4 (2%) 60 31	49, 111, 168, 221	0
1	B	168/182 (92%)	-0.27	7 (4%) 36 14	58, 126, 188, 214	0
1	C	168/182 (92%)	-0.31	1 (0%) 89 72	58, 122, 182, 220	0
1	D	174/182 (95%)	-0.29	1 (0%) 89 72	46, 111, 174, 212	0
All	All	684/728 (93%)	-0.27	13 (1%) 66 37	46, 118, 182, 221	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	472	GLY	5.7
1	A	360	TYR	3.7
1	A	423	LEU	3.6
1	A	422	LEU	3.5
1	B	477	ALA	2.8

### 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 monosaccharides 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.