



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 24, 2022 – 10:07 AM JST

PDB ID : 7X39  
Title : Structure of CIZ1 bound ERH  
Authors : Wang, X.; Xu, C.  
Deposited on : 2022-02-28  
Resolution : 2.85 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	:	2.30
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)	:	2.30

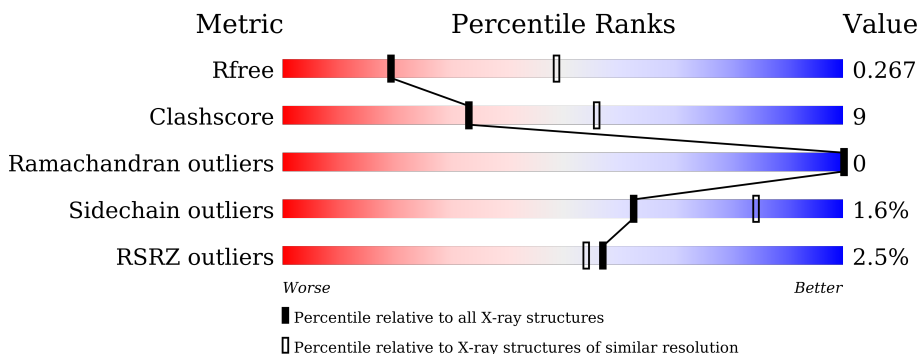
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.85 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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	158	<div> <div>65%</div> <div>15%</div> <div>•</div> <div>19%</div> </div>
1	B	158	<div> <div>58%</div> <div>16%</div> <div>26%</div> </div>
1	C	158	<div> <div>6%</div> <div>54%</div> <div>20%</div> <div>26%</div> </div>
1	D	158	<div> <div>2%</div> <div>54%</div> <div>20%</div> <div>26%</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3828 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 Enhancer of rudimentary homolog,Cip1-interacting zinc finger protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	128	Total	C	N	O	S	0	0	0
			1012	640	167	198	7			
1	B	117	Total	C	N	O	S	0	0	0
			943	595	154	187	7			
1	C	117	Total	C	N	O	S	0	0	0
			933	591	151	184	7			
1	D	117	Total	C	N	O	S	0	0	0
			940	594	154	185	7			

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP P84090
A	0	HIS	-	expression tag	UNP P84090
A	105	GLY	-	linker	UNP P84090
A	106	SER	-	linker	UNP P84090
A	107	SER	-	linker	UNP P84090
A	108	GLY	-	linker	UNP P84090
A	109	SER	-	linker	UNP P84090
A	110	SER	-	linker	UNP P84090
A	111	GLY	-	linker	UNP P84090
A	112	SER	-	linker	UNP P84090
A	113	SER	-	linker	UNP P84090
A	114	GLY	-	linker	UNP P84090
A	115	SER	-	linker	UNP P84090
A	116	SER	-	linker	UNP P84090
A	117	GLY	-	linker	UNP P84090
A	118	SER	-	linker	UNP P84090
A	119	SER	-	linker	UNP P84090
A	120	GLY	-	linker	UNP P84090
A	121	SER	-	linker	UNP P84090
A	122	SER	-	linker	UNP P84090

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	123	GLY	-	linker	UNP P84090
A	124	SER	-	linker	UNP P84090
A	125	SER	-	linker	UNP P84090
B	-1	GLY	-	expression tag	UNP P84090
B	0	HIS	-	expression tag	UNP P84090
B	105	GLY	-	linker	UNP P84090
B	106	SER	-	linker	UNP P84090
B	107	SER	-	linker	UNP P84090
B	108	GLY	-	linker	UNP P84090
B	109	SER	-	linker	UNP P84090
B	110	SER	-	linker	UNP P84090
B	111	GLY	-	linker	UNP P84090
B	112	SER	-	linker	UNP P84090
B	113	SER	-	linker	UNP P84090
B	114	GLY	-	linker	UNP P84090
B	115	SER	-	linker	UNP P84090
B	116	SER	-	linker	UNP P84090
B	117	GLY	-	linker	UNP P84090
B	118	SER	-	linker	UNP P84090
B	119	SER	-	linker	UNP P84090
B	120	GLY	-	linker	UNP P84090
B	121	SER	-	linker	UNP P84090
B	122	SER	-	linker	UNP P84090
B	123	GLY	-	linker	UNP P84090
B	124	SER	-	linker	UNP P84090
B	125	SER	-	linker	UNP P84090
C	-1	GLY	-	expression tag	UNP P84090
C	0	HIS	-	expression tag	UNP P84090
C	105	GLY	-	linker	UNP P84090
C	106	SER	-	linker	UNP P84090
C	107	SER	-	linker	UNP P84090
C	108	GLY	-	linker	UNP P84090
C	109	SER	-	linker	UNP P84090
C	110	SER	-	linker	UNP P84090
C	111	GLY	-	linker	UNP P84090
C	112	SER	-	linker	UNP P84090
C	113	SER	-	linker	UNP P84090
C	114	GLY	-	linker	UNP P84090
C	115	SER	-	linker	UNP P84090
C	116	SER	-	linker	UNP P84090
C	117	GLY	-	linker	UNP P84090
C	118	SER	-	linker	UNP P84090

*Continued on next page...*

*Continued from previous page...*

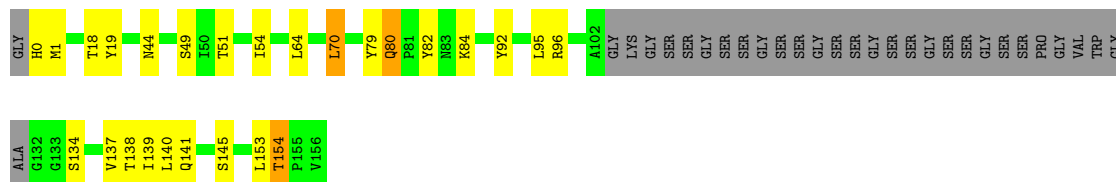
Chain	Residue	Modelled	Actual	Comment	Reference
C	119	SER	-	linker	UNP P84090
C	120	GLY	-	linker	UNP P84090
C	121	SER	-	linker	UNP P84090
C	122	SER	-	linker	UNP P84090
C	123	GLY	-	linker	UNP P84090
C	124	SER	-	linker	UNP P84090
C	125	SER	-	linker	UNP P84090
D	-1	GLY	-	expression tag	UNP P84090
D	0	HIS	-	expression tag	UNP P84090
D	105	GLY	-	linker	UNP P84090
D	106	SER	-	linker	UNP P84090
D	107	SER	-	linker	UNP P84090
D	108	GLY	-	linker	UNP P84090
D	109	SER	-	linker	UNP P84090
D	110	SER	-	linker	UNP P84090
D	111	GLY	-	linker	UNP P84090
D	112	SER	-	linker	UNP P84090
D	113	SER	-	linker	UNP P84090
D	114	GLY	-	linker	UNP P84090
D	115	SER	-	linker	UNP P84090
D	116	SER	-	linker	UNP P84090
D	117	GLY	-	linker	UNP P84090
D	118	SER	-	linker	UNP P84090
D	119	SER	-	linker	UNP P84090
D	120	GLY	-	linker	UNP P84090
D	121	SER	-	linker	UNP P84090
D	122	SER	-	linker	UNP P84090
D	123	GLY	-	linker	UNP P84090
D	124	SER	-	linker	UNP P84090
D	125	SER	-	linker	UNP P84090

### 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.

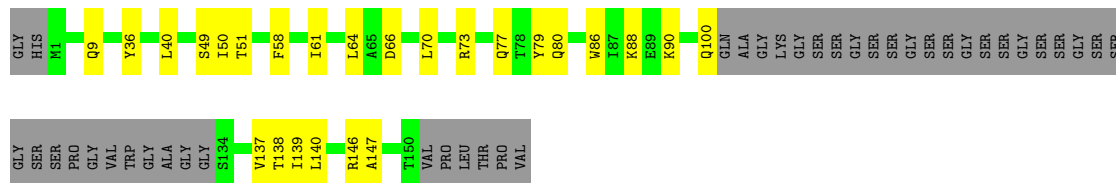
- Molecule 1: Enhancer of rudimentary homolog,Cip1-interacting zinc finger protein

Chain A: 



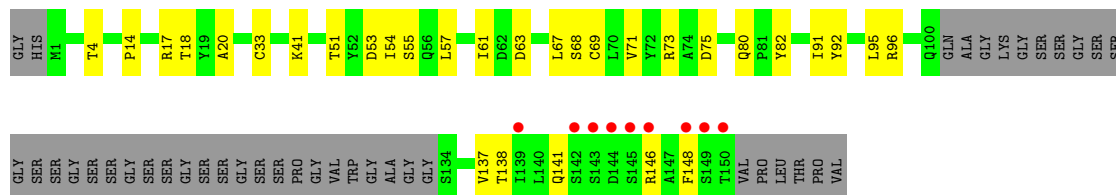
- Molecule 1: Enhancer of rudimentary homolog,Cip1-interacting zinc finger protein

Chain B: 



- Molecule 1: Enhancer of rudimentary homolog,Cip1-interacting zinc finger protein

Chain C: 



- Molecule 1: Enhancer of rudimentary homolog,Cip1-interacting zinc finger protein

Chain D: 



GLY	SER	SER	GLY	SER	SER	GLY	SER	GLY	SER	SER	PRO	GLY	VAL	TRP	GLY	ALA	GLY	GLY	GLY	S134	V137	S143	D144	S145	R146	T150	VAL	PRO	LEU	THR	PRO	VAL
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	44.45Å 48.67Å 73.88Å 85.83° 75.08° 63.31°	Depositor
Resolution (Å)	38.08 – 2.85 38.08 – 2.85	Depositor EDS
% Data completeness (in resolution range)	96.5 (38.08-2.85) 96.5 (38.08-2.85)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.80 (at 2.85Å)	Xtriage
Refinement program	PHENIX 1.14 _3260	Depositor
R, $R_{free}$	0.218 , 0.267 0.218 , 0.267	Depositor DCC
$R_{free}$ test set	694 reflections (5.77%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.6	Xtriage
Anisotropy	0.857	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 31.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.055 for h,h-k,h-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3828	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.27% 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	0.38	0/1033	0.48	0/1402
1	B	0.26	0/961	0.43	0/1302
1	C	0.30	0/951	0.45	0/1290
1	D	0.28	0/958	0.45	0/1298
All	All	0.31	0/3903	0.46	0/5292

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	1012	0	979	16	0
1	B	943	0	911	14	0
1	C	933	0	895	18	0
1	D	940	0	909	19	0
All	All	3828	0	3694	65	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 9.

The worst 5 of 65 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:D:58:PHE:HZ	1:D:92:TYR:HB2	1.65	0.60
1:C:95:LEU:HD22	1:C:148:PHE:HZ	1.67	0.59
1:A:51:THR:HG22	1:A:138:THR:HG23	1.84	0.59
1:B:70:LEU:HB3	1:B:79:TYR:HB3	1.86	0.58
1:A:80:GLN:HG2	1:A:82:TYR:CE2	2.40	0.57

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	124/158 (78%)	121 (98%)	3 (2%)	0	100	100
1	B	113/158 (72%)	110 (97%)	3 (3%)	0	100	100
1	C	113/158 (72%)	111 (98%)	2 (2%)	0	100	100
1	D	113/158 (72%)	111 (98%)	2 (2%)	0	100	100
All	All	463/632 (73%)	453 (98%)	10 (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	113/136 (83%)	109 (96%)	4 (4%)	36	67
1	B	107/136 (79%)	107 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	104/136 (76%)	103 (99%)	1 (1%)	76	91
1	D	106/136 (78%)	104 (98%)	2 (2%)	57	81
All	All	430/544 (79%)	423 (98%)	7 (2%)	62	84

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

Mol	Chain	Res	Type
1	A	154	THR
1	C	63	ASP
1	D	143	SER
1	D	49	SER
1	A	153	LEU

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	128/158 (81%)	-0.00	0	100 100	28, 40, 66, 79	0
1	B	117/158 (74%)	-0.12	0	100 100	30, 40, 54, 72	0
1	C	117/158 (74%)	0.50	9 (7%)	13 10	38, 56, 93, 106	0
1	D	117/158 (74%)	0.14	3 (2%)	56 52	38, 51, 71, 87	0
All	All	479/632 (75%)	0.12	12 (2%)	57 54	28, 48, 74, 106	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	142	SER	5.5
1	C	148	PHE	5.5
1	C	150	THR	5.4
1	C	146	ARG	4.3
1	C	145	SER	4.2

### 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 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.