



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 04:07 am BST

PDB ID : 6DMX  
Title : HBZ56 in complex with KIX and c-Myb  
Authors : Yang, K.; Wright, P.E.; Stanfield, R.L.  
Deposited on : 2018-06-05  
Resolution : 2.80 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

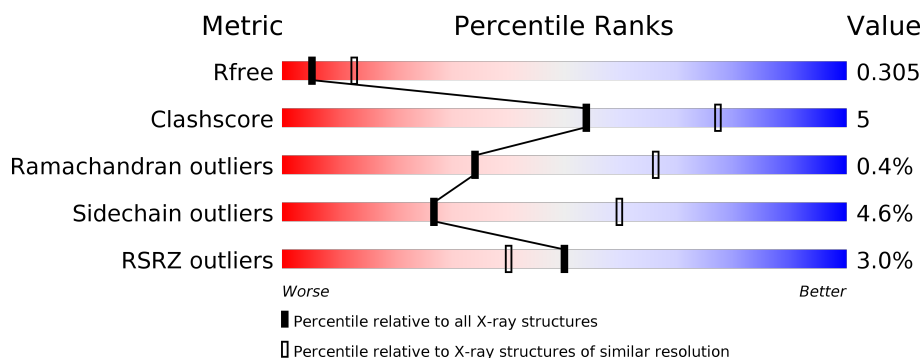
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	E	58	<div> <div>5%</div> <div> <div></div> <div>60%</div> <div>12%</div> <div>•</div> <div>24%</div> </div> </div>
1	J	58	<div> <div>3%</div> <div> <div></div> <div>55%</div> <div>10%</div> <div>•</div> <div>33%</div> </div> </div>
2	A	32	<div> <div></div> <div> <div></div> <div>66%</div> <div>6%</div> <div>•</div> <div>25%</div> </div> </div>
2	C	32	<div> <div>3%</div> <div> <div></div> <div>66%</div> <div>13%</div> <div></div> <div>22%</div> </div> </div>
2	F	32	<div> <div></div> <div> <div></div> <div>50%</div> <div>19%</div> <div></div> <div>31%</div> </div> </div>
2	H	32	<div> <div></div> <div> <div></div> <div>56%</div> <div>16%</div> <div></div> <div>28%</div> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	B	88	<div><div><div></div><div></div><div></div></div><div><div>%</div><div>84%</div><div>10%</div><div>• 5%</div></div></div>
3	D	88	<div><div><div></div><div></div><div></div></div><div><div>2%</div><div>80%</div><div>13%</div><div>8%</div></div></div>
3	G	88	<div><div><div></div><div></div><div></div></div><div><div>2%</div><div>83%</div><div>10%</div><div>• 6%</div></div></div>
3	I	88	<div><div><div></div><div></div><div></div></div><div><div>5%</div><div>77%</div><div>15%</div><div>8%</div></div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4187 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 BZIP factor.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	E	44	Total	C	N	O	0	0	0
			340	210	50	80			
1	J	39	Total	C	N	O	0	0	0
			307	191	42	74			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	-1	GLY	-	expression tag	UNP Q2Q067
E	0	SER	-	expression tag	UNP Q2Q067
E	1	HIS	-	expression tag	UNP Q2Q067
E	2	MET	-	expression tag	UNP Q2Q067
E	9	ALA	CYS	engineered mutation	UNP Q2Q067
E	14	ALA	CYS	engineered mutation	UNP Q2Q067
J	-1	GLY	-	expression tag	UNP Q2Q067
J	0	SER	-	expression tag	UNP Q2Q067
J	1	HIS	-	expression tag	UNP Q2Q067
J	2	MET	-	expression tag	UNP Q2Q067
J	9	ALA	CYS	engineered mutation	UNP Q2Q067
J	14	ALA	CYS	engineered mutation	UNP Q2Q067

- Molecule 2 is a protein called Transcriptional activator Myb.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	25	Total	C	N	O	S	0	0	0
			202	126	33	42	1			
2	A	24	Total	C	N	O	S	0	0	0
			201	124	31	45	1			
2	H	23	Total	C	N	O	S	0	0	0
			193	120	30	42	1			
2	F	22	Total	C	N	O	S	0	0	0
			181	114	28	38	1			

- Molecule 3 is a protein called CREB-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	81	Total	C	N	O	S	0	0	0
			681	429	125	125	2			
3	B	84	Total	C	N	O	S	0	0	0
			706	446	130	128	2			
3	I	81	Total	C	N	O	S	0	0	0
			685	432	126	125	2			
3	G	83	Total	C	N	O	S	0	0	0
			691	437	126	126	2			

There are 4 discrepancies between the modelled and reference sequences:


Chain	Residue	Modelled	Actual	Comment	Reference
D	585	MET	-	initiating methionine	UNP P45481
B	585	MET	-	initiating methionine	UNP P45481
I	585	MET	-	initiating methionine	UNP P45481
G	585	MET	-	initiating methionine	UNP P45481



Chain F:  50% 19% 31%


TYR  
ASN  
ASP  
GLU  
D288  
P289  
R294  
E297  
L300  
T305  
L309  
LYS  
GLY  
GLN  
GLN  
ALA  
LEU

• Molecule 3: CREB-binding protein

Chain D:  2% 80% 13% 8%


MET  
GLY  
VAL  
ARG  
LYS  
GLY  
H591  
H602  
Q609  
L620  
R623  
R624  
M625  
E626  
M627  
L628  
V629  
A630  
Y631  
M644  
E655  
I660  
R671  
LEU

• Molecule 3: CREB-binding protein

Chain B:  % 84% 10% 5%


MET  
GLY  
VAL  
ARG  
K589  
F617  
L620  
R623  
R624  
M625  
E626  
V629  
M639  
S642  
L652  
I657  
E666  
L672

• Molecule 3: CREB-binding protein

Chain I:  5% 77% 15% 8%

MET  
GLY  
VAL  
ARG  
LYS  
GLY  
H591  
H592  
E593  
H594  
A619  
L620  
K621  
R624  
M625  
E626  
M627  
L628  
V629  
M639  
Y640  
L652  
E655  
Q661  
S670  
R671  
LEU

• Molecule 3: CREB-binding protein

Chain G:  2% 83% 10% 6%

MET  
GLY  
VAL  
R588  
H602  
V608  
Q609  
F613  
THR  
PRO  
R616  
L620  
M625  
V629  
M639  
L652  
K662  
E663  
L664  
S670  
R671  
L672

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.00Å 80.31Å 64.64Å 90.00° 92.60° 90.00°	Depositor
Resolution (Å)	40.97 – 2.80 40.94 – 2.80	Depositor EDS
% Data completeness (in resolution range)	94.6 (40.97-2.80) 94.6 (40.94-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.72 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0222	Depositor
R, $R_{free}$	0.239 , 0.301 0.240 , 0.305	Depositor DCC
$R_{free}$ test set	660 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	84.8	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 63.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.078 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4187	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	98.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.90% 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	E	0.70	0/340	0.60	0/456
1	J	0.70	0/306	0.55	0/410
2	A	0.45	0/201	0.61	0/267
2	C	0.51	0/202	0.61	0/267
2	F	0.55	0/181	0.69	0/240
2	H	0.49	0/193	0.56	0/256
3	B	0.55	0/721	0.58	0/968
3	D	0.55	0/696	0.58	0/937
3	G	0.56	0/704	0.61	0/943
3	I	0.55	0/700	0.59	0/941
All	All	0.57	0/4244	0.59	0/5685

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	E	340	0	333	7	0
1	J	307	0	300	3	0
2	A	201	0	201	3	0
2	C	202	0	209	2	0
2	F	181	0	187	5	0
2	H	193	0	197	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	706	0	707	8	0
3	D	681	0	669	7	0
3	G	691	0	683	7	0
3	I	685	0	680	13	0
All	All	4187	0	4166	44	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 5.

The worst 5 of 44 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)
3:B:623:ARG:HH11	3:B:623:ARG:HG3	1.50	0.77
2:A:307:ASN:N	2:A:307:ASN:HD22	1.88	0.72
3:I:594:HIS:O	3:I:594:HIS:CG	2.49	0.65
3:I:591:TRP:HZ2	3:I:640:TYR:CZ	2.15	0.64
1:E:14:ALA:HB3	1:E:15:PRO:HD3	1.81	0.62

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	E	42/58 (72%)	40 (95%)	1 (2%)	1 (2%)	6	20
1	J	37/58 (64%)	37 (100%)	0	0	100	100
2	A	22/32 (69%)	20 (91%)	2 (9%)	0	100	100
2	C	23/32 (72%)	21 (91%)	2 (9%)	0	100	100
2	F	20/32 (62%)	16 (80%)	3 (15%)	1 (5%)	2	6
2	H	21/32 (66%)	19 (90%)	2 (10%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	B	82/88 (93%)	80 (98%)	2 (2%)	0	100	100
3	D	79/88 (90%)	78 (99%)	1 (1%)	0	100	100
3	G	79/88 (90%)	79 (100%)	0	0	100	100
3	I	79/88 (90%)	76 (96%)	3 (4%)	0	100	100
All	All	484/596 (81%)	466 (96%)	16 (3%)	2 (0%)	34	66

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	289	PRO
1	E	14	ALA

### 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	E	39/49 (80%)	35 (90%)	4 (10%)	7	21
1	J	36/49 (74%)	32 (89%)	4 (11%)	6	19
2	A	24/30 (80%)	22 (92%)	2 (8%)	11	32
2	C	23/30 (77%)	21 (91%)	2 (9%)	10	30
2	F	21/30 (70%)	20 (95%)	1 (5%)	25	58
2	H	23/30 (77%)	21 (91%)	2 (9%)	10	30
3	B	75/78 (96%)	73 (97%)	2 (3%)	44	78
3	D	72/78 (92%)	71 (99%)	1 (1%)	67	90
3	G	72/78 (92%)	71 (99%)	1 (1%)	67	90
3	I	73/78 (94%)	71 (97%)	2 (3%)	44	78
All	All	458/530 (86%)	437 (95%)	21 (5%)	27	60

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	623	ARG
1	J	17	ASP
3	I	592	HIS
2	A	308	GLU
3	I	621	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	307	ASN
3	I	594	HIS

### 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 carbohydrates 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 [i](#)

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	E	44/58 (75%)	0.07	3 (6%) 17 10	53, 84, 153, 172	0
1	J	39/58 (67%)	-0.06	2 (5%) 28 19	74, 106, 148, 151	0
2	A	24/32 (75%)	-0.26	0 100 100	59, 95, 131, 137	0
2	C	25/32 (78%)	0.15	1 (4%) 38 28	67, 96, 147, 152	0
2	F	22/32 (68%)	-0.27	0 100 100	58, 76, 104, 121	0
2	H	23/32 (71%)	-0.03	0 100 100	69, 101, 136, 154	0
3	B	84/88 (95%)	-0.12	1 (1%) 79 73	50, 75, 120, 174	0
3	D	81/88 (92%)	0.04	2 (2%) 57 47	69, 101, 143, 158	0
3	G	83/88 (94%)	0.04	2 (2%) 59 49	52, 86, 152, 169	0
3	I	81/88 (92%)	0.20	4 (4%) 29 20	71, 109, 154, 188	0
All	All	506/596 (84%)	0.01	15 (2%) 50 40	50, 94, 151, 188	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	644	ASN	5.0
2	C	312	GLN	4.4
1	E	14	ALA	4.3
3	G	620	LEU	3.0
3	D	623	ARG	2.9

### 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 [i](#)

There are no carbohydrates 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.