



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2024 – 06:08 PM EST

PDB ID : 3JUA  
Title : Structural basis of YAP recognition by TEAD4 in the Hippo pathway  
Authors : Chen, L.; Song, H.  
Deposited on : 2009-09-15  
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 : 2.36  
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) : 2.36

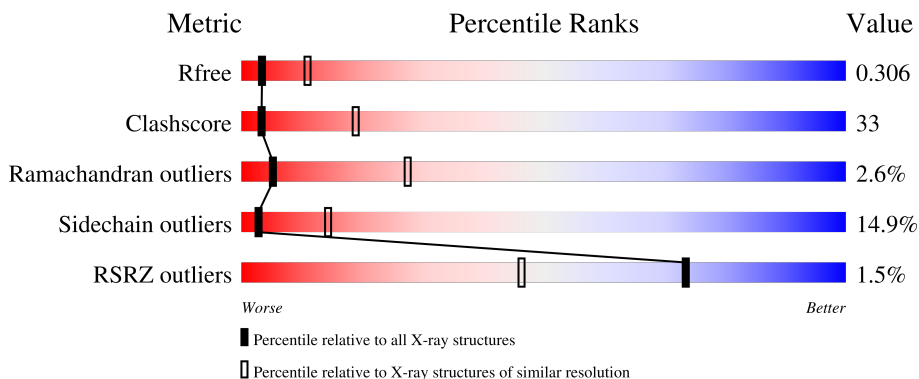
# 1 Overall quality at a glance

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	220	
1	C	220	
1	E	220	
1	G	220	
2	B	39	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	D	39	
2	F	39	
2	H	39	

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 8055 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 Transcriptional enhancer factor TEF-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	204	Total 1692	C 1092	N 281	O 309	S 10	0	0	0
1	C	213	Total 1762	C 1131	N 293	O 328	S 10	0	0	0
1	E	216	Total 1782	C 1144	N 294	O 334	S 10	0	0	0
1	G	190	Total 1586	C 1021	N 263	O 292	S 10	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	208	SER	-	expression tag	UNP Q62296
A	209	MET	-	expression tag	UNP Q62296
C	208	SER	-	expression tag	UNP Q62296
C	209	MET	-	expression tag	UNP Q62296
E	208	SER	-	expression tag	UNP Q62296
E	209	MET	-	expression tag	UNP Q62296
G	208	SER	-	expression tag	UNP Q62296
G	209	MET	-	expression tag	UNP Q62296

- Molecule 2 is a protein called 65 kDa Yes-associated protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	39	Total 311	C 200	N 52	O 57	S 2	0	0	0
2	D	35	Total 279	C 180	N 48	O 49	S 2	0	0	0
2	F	36	Total 286	C 185	N 49	O 50	S 2	0	0	0
2	H	19	Total 158	C 105	N 28	O 24	S 1	0	0	0

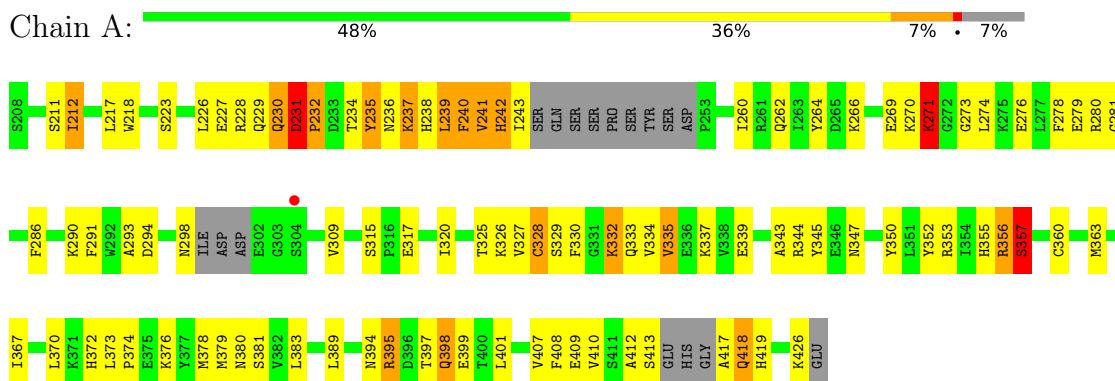
- Molecule 3 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
3	A	41	Total O 41 41	0	0
3	B	7	Total O 7 7	0	0
3	C	45	Total O 45 45	0	0
3	D	6	Total O 6 6	0	0
3	E	51	Total O 51 51	0	0
3	F	7	Total O 7 7	0	0
3	G	42	Total O 42 42	0	0

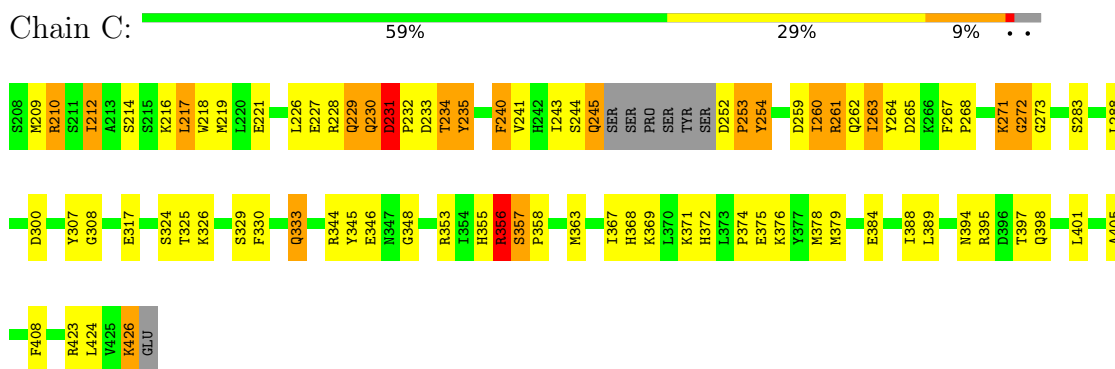
### 3 Residue-property plots

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.

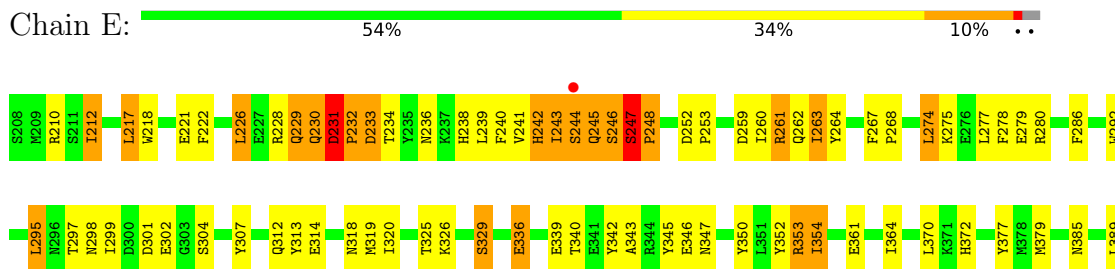
- Molecule 1: Transcriptional enhancer factor TEF-3



- Molecule 1: Transcriptional enhancer factor TEF-3

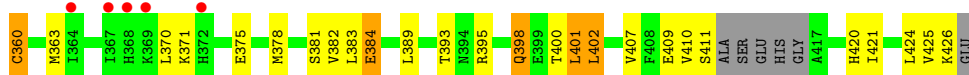
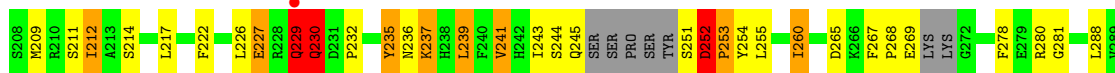


- Molecule 1: Transcriptional enhancer factor TEF-3





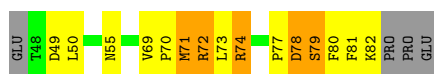
- Molecule 1: Transcriptional enhancer factor TEF-3



- Molecule 2: 65 kDa Yes-associated protein



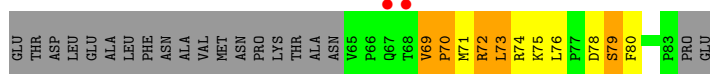
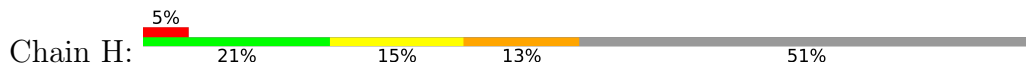
- Molecule 2: 65 kDa Yes-associated protein



- Molecule 2: 65 kDa Yes-associated protein



- Molecule 2: 65 kDa Yes-associated protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.98Å 146.91Å 165.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00 74.35 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (20.00-3.00) 99.9 (74.35-2.80)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.04 (at 2.82Å)	Xtrriage
Refinement program	REFMAC 5.4.0077	Depositor
R, $R_{free}$	0.236 , 0.288 0.256 , 0.306	Depositor DCC
$R_{free}$ test set	3103 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	68.7	Xtrriage
Anisotropy	0.507	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 66.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	8055	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.24% 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	0.58	0/1733	0.69	1/2335 (0.0%)
1	C	0.61	0/1806	0.85	8/2438 (0.3%)
1	E	0.60	0/1827	0.83	4/2468 (0.2%)
1	G	0.49	0/1623	0.95	8/2189 (0.4%)
2	B	0.49	0/319	0.67	0/434
2	D	0.65	0/285	0.83	1/386 (0.3%)
2	F	0.61	0/293	0.98	2/398 (0.5%)
2	H	0.46	0/163	0.81	1/220 (0.5%)
All	All	0.57	0/8049	0.83	25/10868 (0.2%)

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	A	0	1
1	C	0	3
1	E	0	1
1	G	0	1
2	F	0	1
All	All	0	7

There are no bond length outliers.

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	357	SER	C-N-CD	-19.15	78.46	120.60
1	G	252	ASP	C-N-CD	-16.01	85.37	120.60
1	G	230	GLN	N-CA-C	-10.08	83.78	111.00
1	G	358	PRO	N-CA-C	-9.87	86.45	112.10
1	C	272	GLY	N-CA-C	-9.60	89.10	113.10

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	357	SER	Peptide
1	C	229	GLN	Peptide
1	C	231	ASP	Peptide
1	C	357	SER	Peptide
1	E	229	GLN	Peptide

## 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	1692	0	1657	121	0
1	C	1762	0	1710	100	0
1	E	1782	0	1730	111	0
1	G	1586	0	1537	95	0
2	B	311	0	315	23	0
2	D	279	0	289	31	0
2	F	286	0	296	26	0
2	H	158	0	170	29	0
3	A	41	0	0	3	0
3	B	7	0	0	2	0
3	C	45	0	0	10	0
3	D	6	0	0	3	0
3	E	51	0	0	6	0
3	F	7	0	0	1	0
3	G	42	0	0	9	0
All	All	8055	0	7704	511	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 33.

The worst 5 of 511 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)
2:H:69:VAL:HG12	2:H:70:PRO:CD	1.54	1.38

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:252:ASP:CB	1:G:253:PRO:HD3	1.30	1.37
2:H:75:LYS:O	2:H:76:LEU:HD23	1.18	1.32
1:C:210:ARG:HD3	1:C:252:ASP:OD1	1.27	1.25
1:C:271:LYS:CD	1:C:271:LYS:O	1.87	1.23

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	196/220 (89%)	173 (88%)	16 (8%)	7 (4%)	3	19
1	C	209/220 (95%)	189 (90%)	18 (9%)	2 (1%)	15	53
1	E	212/220 (96%)	189 (89%)	16 (8%)	7 (3%)	4	21
1	G	178/220 (81%)	164 (92%)	10 (6%)	4 (2%)	6	31
2	B	37/39 (95%)	32 (86%)	4 (11%)	1 (3%)	5	26
2	D	33/39 (85%)	30 (91%)	3 (9%)	0	100	100
2	F	34/39 (87%)	25 (74%)	8 (24%)	1 (3%)	4	24
2	H	17/39 (44%)	10 (59%)	5 (29%)	2 (12%)	0	1
All	All	916/1036 (88%)	812 (89%)	80 (9%)	24 (3%)	5	27

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	230	GLN
1	A	231	ASP
1	A	271	LYS
1	C	358	PRO
1	E	244	SER

### 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	188/203 (93%)	162 (86%)	26 (14%)	3	17
1	C	196/203 (97%)	169 (86%)	27 (14%)	3	17
1	E	200/203 (98%)	171 (86%)	29 (14%)	3	15
1	G	178/203 (88%)	152 (85%)	26 (15%)	3	15
2	B	36/36 (100%)	28 (78%)	8 (22%)	1	4
2	D	32/36 (89%)	26 (81%)	6 (19%)	1	8
2	F	33/36 (92%)	27 (82%)	6 (18%)	1	9
2	H	19/36 (53%)	16 (84%)	3 (16%)	2	12
All	All	882/956 (92%)	751 (85%)	131 (15%)	3	14

5 of 131 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	295	LEU
1	G	360	CYS
2	H	79	SER
1	C	324	SER
1	C	283	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	385	ASN
1	E	333	GLN
1	G	368	HIS
1	E	262	GLN
1	E	365	ASN

### 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 [i](#)

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	204/220 (92%)	-0.05	1 (0%) 91 75	15, 25, 35, 43	0
1	C	213/220 (96%)	-0.13	0 100 100	18, 28, 39, 53	0
1	E	216/220 (98%)	-0.00	1 (0%) 91 75	19, 28, 40, 60	0
1	G	190/220 (86%)	0.32	8 (4%) 36 14	14, 28, 39, 43	0
2	B	39/39 (100%)	0.06	1 (2%) 56 27	21, 25, 36, 36	0
2	D	35/39 (89%)	-0.19	0 100 100	25, 28, 37, 39	0
2	F	36/39 (92%)	0.18	1 (2%) 53 25	26, 32, 39, 39	0
2	H	19/39 (48%)	0.38	2 (10%) 6 2	21, 27, 34, 34	0
All	All	952/1036 (91%)	0.03	14 (1%) 73 46	14, 28, 39, 60	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	229	GLN	3.9
2	F	61	LYS	3.4
2	H	67	GLN	3.0
1	G	367	ILE	3.0
1	G	359	LEU	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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.