



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 12:38 pm BST

PDB ID : 1J1J  
Title : Crystal Structure of human Translin  
Authors : Sugiura, I.; Sasaki, C.; Hasegawa, T.; Kohno, T.; Sugio, S.; Moriyama, H.;  
Kasai, M.; Matsuzaki, T.  
Deposited on : 2002-12-06  
Resolution : 2.20 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

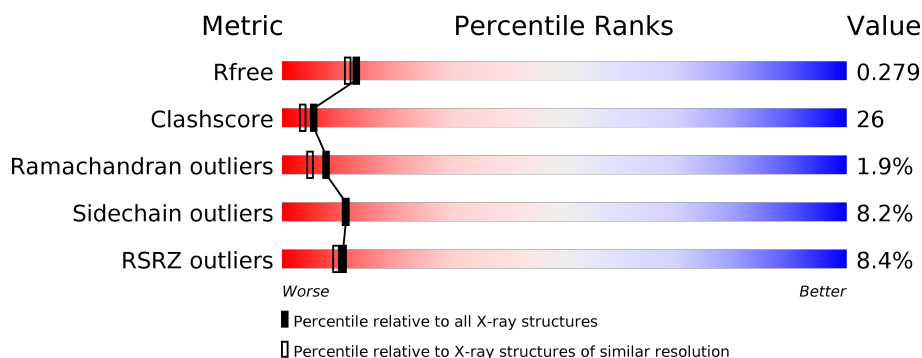
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	240	<div> <div>7%</div> <div> <div></div> <div>59%</div> <div>27%</div> <div>•</div> <div>10%</div> </div> </div>
1	B	240	<div> <div>8%</div> <div> <div></div> <div>53%</div> <div>33%</div> <div>5%</div> <div>10%</div> </div> </div>
1	C	240	<div> <div>10%</div> <div> <div></div> <div>50%</div> <div>36%</div> <div>•</div> <div>10%</div> </div> </div>
1	D	240	<div> <div>5%</div> <div> <div></div> <div>54%</div> <div>32%</div> <div>•</div> <div>10%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7494 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 Translin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	217	Total	C	N	O	S	0	0	0
			1770	1135	305	328	2			
1	B	217	Total	C	N	O	S	0	0	0
			1770	1135	305	328	2			
1	C	217	Total	C	N	O	S	0	0	0
			1770	1135	305	328	2			
1	D	217	Total	C	N	O	S	0	0	0
			1770	1135	305	328	2			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	EXPRESSION TAG	UNP Q15631
A	-10	ARG	-	EXPRESSION TAG	UNP Q15631
A	-9	GLY	-	EXPRESSION TAG	UNP Q15631
A	-8	SER	-	EXPRESSION TAG	UNP Q15631
A	-7	HIS	-	EXPRESSION TAG	UNP Q15631
A	-6	HIS	-	EXPRESSION TAG	UNP Q15631
A	-5	HIS	-	EXPRESSION TAG	UNP Q15631
A	-4	HIS	-	EXPRESSION TAG	UNP Q15631
A	-3	HIS	-	EXPRESSION TAG	UNP Q15631
A	-2	HIS	-	EXPRESSION TAG	UNP Q15631
A	-1	GLY	-	EXPRESSION TAG	UNP Q15631
A	0	SER	-	EXPRESSION TAG	UNP Q15631
B	-11	MET	-	EXPRESSION TAG	UNP Q15631
B	-10	ARG	-	EXPRESSION TAG	UNP Q15631
B	-9	GLY	-	EXPRESSION TAG	UNP Q15631
B	-8	SER	-	EXPRESSION TAG	UNP Q15631
B	-7	HIS	-	EXPRESSION TAG	UNP Q15631
B	-6	HIS	-	EXPRESSION TAG	UNP Q15631
B	-5	HIS	-	EXPRESSION TAG	UNP Q15631
B	-4	HIS	-	EXPRESSION TAG	UNP Q15631
B	-3	HIS	-	EXPRESSION TAG	UNP Q15631

*Continued on next page...*

*Continued from previous page...*

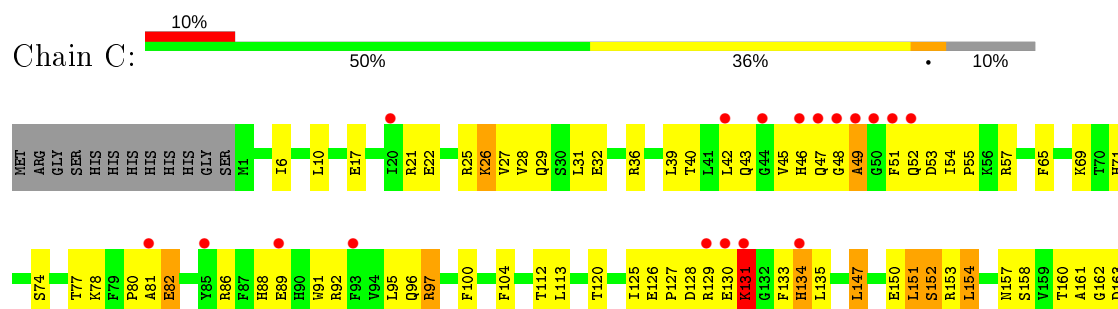
Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	HIS	-	EXPRESSION TAG	UNP Q15631
B	-1	GLY	-	EXPRESSION TAG	UNP Q15631
B	0	SER	-	EXPRESSION TAG	UNP Q15631
C	-11	MET	-	EXPRESSION TAG	UNP Q15631
C	-10	ARG	-	EXPRESSION TAG	UNP Q15631
C	-9	GLY	-	EXPRESSION TAG	UNP Q15631
C	-8	SER	-	EXPRESSION TAG	UNP Q15631
C	-7	HIS	-	EXPRESSION TAG	UNP Q15631
C	-6	HIS	-	EXPRESSION TAG	UNP Q15631
C	-5	HIS	-	EXPRESSION TAG	UNP Q15631
C	-4	HIS	-	EXPRESSION TAG	UNP Q15631
C	-3	HIS	-	EXPRESSION TAG	UNP Q15631
C	-2	HIS	-	EXPRESSION TAG	UNP Q15631
C	-1	GLY	-	EXPRESSION TAG	UNP Q15631
C	0	SER	-	EXPRESSION TAG	UNP Q15631
D	-11	MET	-	EXPRESSION TAG	UNP Q15631
D	-10	ARG	-	EXPRESSION TAG	UNP Q15631
D	-9	GLY	-	EXPRESSION TAG	UNP Q15631
D	-8	SER	-	EXPRESSION TAG	UNP Q15631
D	-7	HIS	-	EXPRESSION TAG	UNP Q15631
D	-6	HIS	-	EXPRESSION TAG	UNP Q15631
D	-5	HIS	-	EXPRESSION TAG	UNP Q15631
D	-4	HIS	-	EXPRESSION TAG	UNP Q15631
D	-3	HIS	-	EXPRESSION TAG	UNP Q15631
D	-2	HIS	-	EXPRESSION TAG	UNP Q15631
D	-1	GLY	-	EXPRESSION TAG	UNP Q15631
D	0	SER	-	EXPRESSION TAG	UNP Q15631

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	126	Total O 126 126	0	0
2	B	86	Total O 86 86	0	0
2	C	91	Total O 91 91	0	0
2	D	111	Total O 111 111	0	0



- Molecule 1: Translin





● Molecule 1: Translin



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	129.42Å 135.27Å 134.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 2.20 15.04 – 2.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (15.00-2.20) 96.2 (15.04-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.04 (at 2.20Å)	Xtriage
Refinement program	CNX	Depositor
R, $R_{free}$	0.229 , 0.265 0.246 , 0.279	Depositor DCC
$R_{free}$ test set	2908 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.3	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 57.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.017 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7494	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.74% 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	0.40	0/1803	0.54	0/2433
1	B	0.36	0/1803	0.57	2/2433 (0.1%)
1	C	0.36	0/1803	0.55	0/2433
1	D	0.37	0/1803	0.57	0/2433
All	All	0.37	0/7212	0.56	2/9732 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	54	ILE	N-CA-C	-5.68	95.65	111.00
1	B	54	ILE	C-N-CD	-5.64	108.19	120.60

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	1770	0	1787	72	0
1	B	1770	0	1787	102	0
1	C	1770	0	1787	123	0
1	D	1770	0	1787	95	0
2	A	126	0	0	12	0
2	B	86	0	0	8	1
2	C	91	0	0	20	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	111	0	0	14	0
All	All	7494	0	7148	366	1

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 26.

The worst 5 of 366 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:97:ARG:HB3	1:C:97:ARG:HH11	1.12	1.09
1:D:187:LYS:HE3	1:D:187:LYS:H	1.10	1.08
1:B:54:ILE:HA	2:B:229:HOH:O	1.55	1.05
1:B:97:ARG:HH11	1:B:97:ARG:HB3	1.21	1.04
1:A:97:ARG:HH11	1:A:97:ARG:HB3	1.18	1.03

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:268:HOH:O	2:B:268:HOH:O[4_555]	1.74	0.46

## 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	215/240 (90%)	205 (95%)	7 (3%)	3 (1%)	11	8
1	B	215/240 (90%)	205 (95%)	8 (4%)	2 (1%)	17	16
1	C	215/240 (90%)	191 (89%)	19 (9%)	5 (2%)	6	3
1	D	215/240 (90%)	204 (95%)	5 (2%)	6 (3%)	5	2
All	All	860/960 (90%)	805 (94%)	39 (4%)	16 (2%)	8	5

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	130	GLU
1	A	131	LYS
1	B	128	ASP
1	B	131	LYS
1	C	49	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	A	194/212 (92%)	179 (92%)	15 (8%)	13	13
1	B	194/212 (92%)	174 (90%)	20 (10%)	7	6
1	C	194/212 (92%)	180 (93%)	14 (7%)	14	15
1	D	194/212 (92%)	179 (92%)	15 (8%)	13	13
All	All	776/848 (92%)	712 (92%)	64 (8%)	11	11

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	182	ARG
1	C	26	LYS
1	D	185	ASN
1	B	185	ASN
1	B	206	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	188	ASN
1	C	47	GLN
1	D	47	GLN
1	B	175	ASN
1	B	185	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 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 [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	217/240 (90%)	0.51	17 (7%) 13 11	29, 47, 67, 96	0
1	B	217/240 (90%)	0.46	20 (9%) 9 7	31, 51, 78, 93	0
1	C	217/240 (90%)	0.98	25 (11%) 4 4	30, 51, 91, 106	0
1	D	217/240 (90%)	0.32	11 (5%) 28 26	31, 46, 73, 99	0
All	All	868/960 (90%)	0.56	73 (8%) 11 9	29, 49, 81, 106	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	49	ALA	15.2
1	C	48	GLY	11.6
1	C	50	GLY	10.0
1	C	47	GLN	8.8
1	C	51	PHE	8.2

### 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 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.