



# wwPDB X-ray Structure Validation Summary Report

Nov 9, 2021 – 10:09 am GMT

PDB ID : 7PJF  
Title : Inhibiting parasite proliferation using a rationally designed anti-tubulin agent  
Authors : Sharma, A.; Gaillard, N.; Ehrhard, V.A.; Steinmetz, M.O.  
Deposited on : 2021-08-24  
Resolution : 1.86 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  
Mogul : 1.8.4 (270009), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

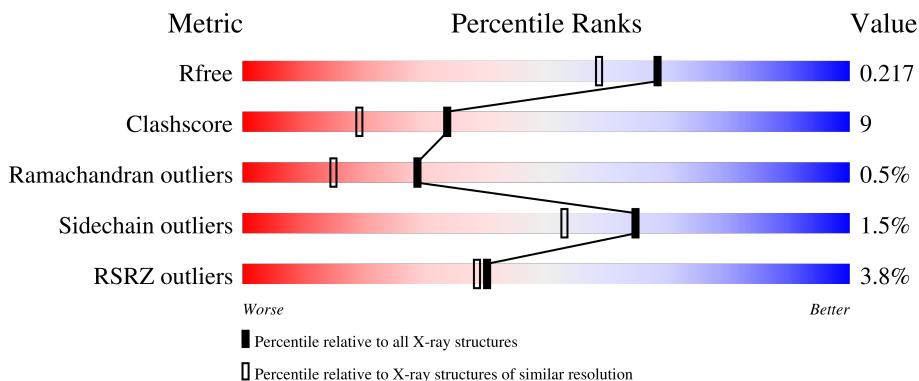
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.86 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	451	 4% 78% 13% • 7%
2	B	450	 3% 78% 16% • 5%
3	F	157	 3% 89% 9% ..

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 8262 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 Tubulin alpha-1B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	419	3268	2074	555	619	20	0	2	0

- Molecule 2 is a protein called Tubulin beta-3 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	428	3355	2110	572	647	26	0	2	0

- Molecule 3 is a protein called Designed ankyrin repeat protein (DARPIN) D1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	F	155	1135	713	195	224	3	0	0	0

- Molecule 4 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
4	B	1	Total	C	N	O	P	0	0
			32	10	5	14	3		

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mg	0	0
			1	1		
5	B	1	Total	Mg	0	0
			1	1		

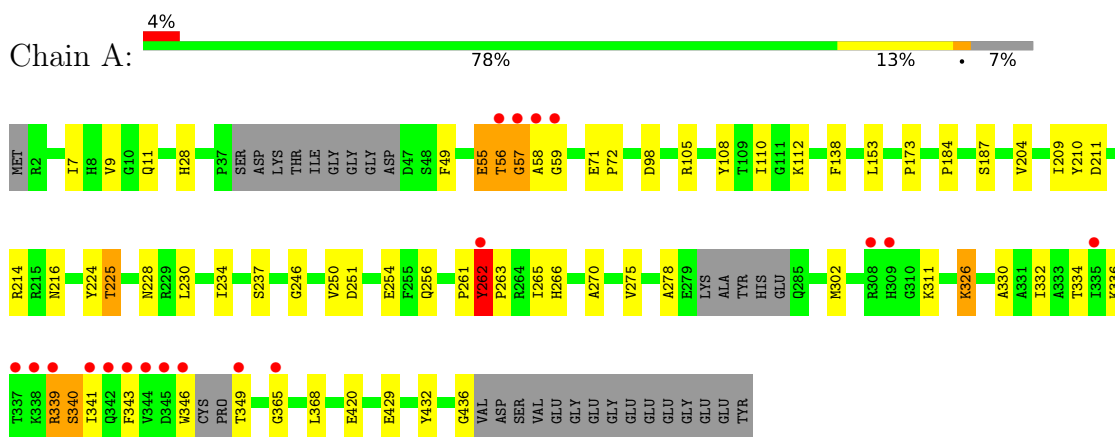
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	194	Total	O	0	0
			194	194		
6	B	150	Total	O	0	0
			150	150		
6	F	94	Total	O	0	0
			94	94		

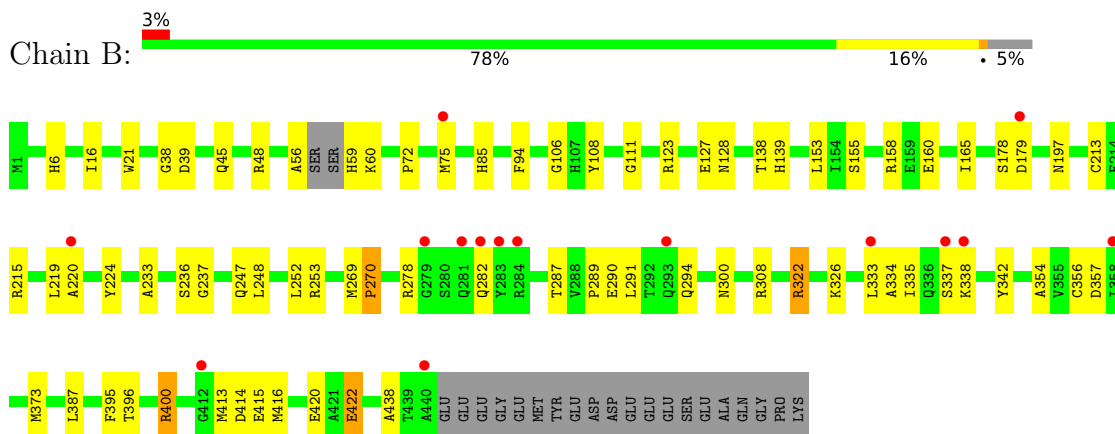
### 3 Residue-property plots [i](#)

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.

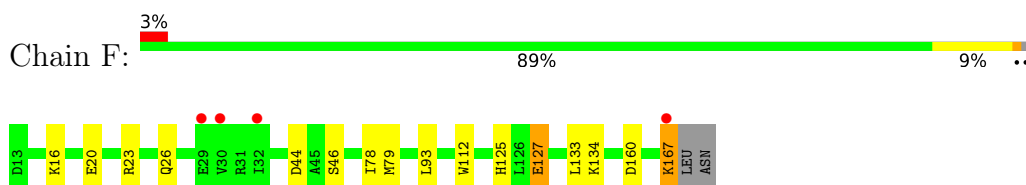
- Molecule 1: Tubulin alpha-1B chain



- Molecule 2: Tubulin beta-3 chain



- Molecule 3: Designed ankyrin repeat protein (DARPIN) D1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.78Å 91.31Å 82.66Å 90.00° 97.55° 90.00°	Depositor
Resolution (Å)	45.66 – 1.86 49.27 – 1.86	Depositor EDS
% Data completeness (in resolution range)	99.6 (45.66-1.86) 99.9 (49.27-1.86)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.03 (at 1.87Å)	Xtrriage
Refinement program	PHENIX 1.14_3260	Depositor
R, $R_{free}$	0.184 , 0.218 0.181 , 0.217	Depositor DCC
$R_{free}$ test set	4538 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.2	Xtrriage
Anisotropy	0.342	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8262	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.52% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GTP, MG

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	0.80	3/3343 (0.1%)	0.79	3/4541 (0.1%)
2	B	0.75	3/3431 (0.1%)	0.81	7/4648 (0.2%)
3	F	0.74	0/1150	0.79	1/1565 (0.1%)
All	All	0.77	6/7924 (0.1%)	0.80	11/10754 (0.1%)

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	2

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	270	PRO	N-CA	12.22	1.68	1.47
1	A	9	VAL	CB-CG1	6.75	1.67	1.52
2	B	422	GLU	CG-CD	6.44	1.61	1.51
1	A	138	PHE	CD1-CE1	6.29	1.51	1.39
1	A	429	GLU	CB-CG	5.32	1.62	1.52

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	400	ARG	NE-CZ-NH1	9.49	125.04	120.30
2	B	270	PRO	CA-N-CD	-7.56	100.92	111.50
2	B	400	ARG	NE-CZ-NH2	-7.55	116.53	120.30
2	B	253	ARG	NE-CZ-NH2	-7.40	116.60	120.30
3	F	44	ASP	CB-CG-OD1	6.97	124.57	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	55	GLU	Peptide
1	A	57	GLY	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	3268	0	3159	65	1
2	B	3355	0	3229	52	0
3	F	1135	0	1125	18	0
4	A	32	0	12	2	0
4	B	32	0	12	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	194	0	0	20	0
6	B	150	0	0	15	0
6	F	94	0	0	7	0
All	All	8262	0	7537	131	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 9.

The worst 5 of 131 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:B:270:PRO:N	2:B:270:PRO:CA	1.68	1.47
2:B:233:ALA:O	6:B:601:HOH:O	1.67	1.11
3:F:127:GLU:OE1	6:F:201:HOH:O	1.68	1.09
1:A:105:ARG:O	6:A:601:HOH:O	1.72	1.05
1:A:105:ARG:NH1	6:A:603:HOH:O	1.94	0.99

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)
1:A:262:TYR:OH	1:A:365:GLY:N[2_1055]	2.15	0.05

### 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	413/451 (92%)	398 (96%)	12 (3%)	3 (1%)	22	9
2	B	426/450 (95%)	415 (97%)	9 (2%)	2 (0%)	29	15
3	F	153/157 (98%)	152 (99%)	1 (1%)	0	100	100
All	All	992/1058 (94%)	965 (97%)	22 (2%)	5 (0%)	29	15

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	56	THR
1	A	225	THR
1	A	278	ALA
2	B	282	GLN
2	B	60	LYS

#### 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	348/379 (92%)	343 (99%)	5 (1%)	67	55
2	B	365/386 (95%)	360 (99%)	5 (1%)	67	55

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	F	116/122 (95%)	114 (98%)	2 (2%)	60	47
All	All	829/887 (94%)	817 (99%)	12 (1%)	65	55

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	158	ARG
2	B	322	ARG
3	F	167	LYS
2	B	326	LYS
1	A	339	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	45	GLN
2	B	59	HIS
3	F	140	ASN
3	F	26	GLN
3	F	59	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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

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	419/451 (92%)	-0.04	19 (4%) 33 32	22, 35, 70, 131	1 (0%)
2	B	428/450 (95%)	-0.03	15 (3%) 44 41	23, 37, 71, 138	0
3	F	155/157 (98%)	-0.33	4 (2%) 56 54	21, 30, 51, 91	0
All	All	1002/1058 (94%)	-0.08	38 (3%) 40 38	21, 35, 67, 138	1 (0%)

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	57	GLY	6.5
2	B	283	TYR	6.4
2	B	282	GLN	6.4
2	B	440	ALA	6.0
1	A	262	TYR	5.2

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	MG	A	502	1/1	0.91	0.14	26,26,26,26	0
4	GTP	B	501	32/32	0.97	0.08	25,28,34,36	0
4	GTP	A	501	32/32	0.97	0.11	17,25,27,35	0
5	MG	B	502	1/1	0.97	0.04	32,32,32,32	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.