



wwPDB X-ray Structure Validation Summary Report

Nov 9, 2021 – 10:12 am GMT

PDB ID : 7PJE
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 : 1.75 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

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

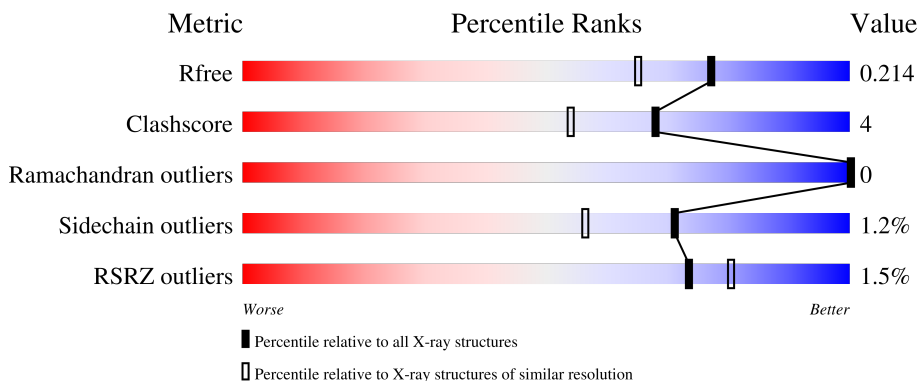
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.75 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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 ≥ 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	449	 87% 6% • 6%
1	C	449	 85% 8% 6%
2	B	443	 90% 5% • •
2	D	443	 86% 9% • •
3	E	155	 75% 6% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	155	 79% 19%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 31317 atoms, of which 14958 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 chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	422	6574	2096	3265	561	628	24	0	9	0
1	C	424	6551	2089	3250	561	629	22	0	2	0

- Molecule 2 is a protein called Tubulin beta chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	425	6673	2124	3295	582	643	29	0	11	0
2	D	424	6553	2091	3227	570	637	28	0	4	0

- Molecule 3 is a protein called Darpin D1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
3	F	126	1867	588	937	161	178	3	0	0	0
3	E	125	1877	589	946	162	177	3	0	2	0

- Molecule 4 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: C₁₀H₁₆N₅O₁₄P₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	H	N	O			P
4	A	1	Total	C	H	N	O	P	0	0
			41	10	9	5	14	3		
4	B	1	Total	C	H	N	O	P	0	0
			42	10	10	5	14	3		
4	C	1	Total	C	H	N	O	P	0	0
			42	10	10	5	14	3		
4	D	1	Total	C	H	N	O	P	0	0
			41	10	9	5	14	3		

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

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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	241	Total	O	0	0
			241	241		
6	B	280	Total	O	0	0
			280	280		

Continued on next page...


Continued from previous page...

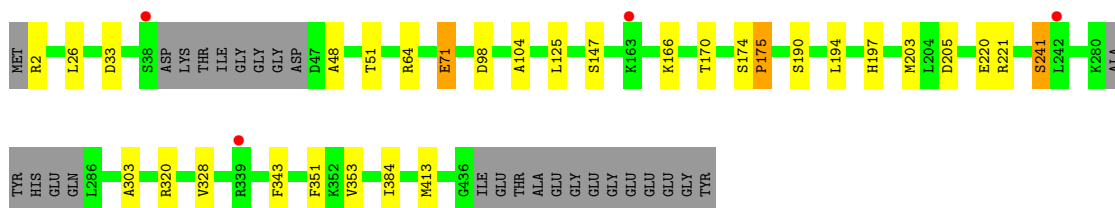
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	F	67	Total 67	O 67	0	0
6	C	228	Total 228	O 228	0	0
6	D	160	Total 160	O 160	0	0
6	E	77	Total 77	O 77	0	0

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.

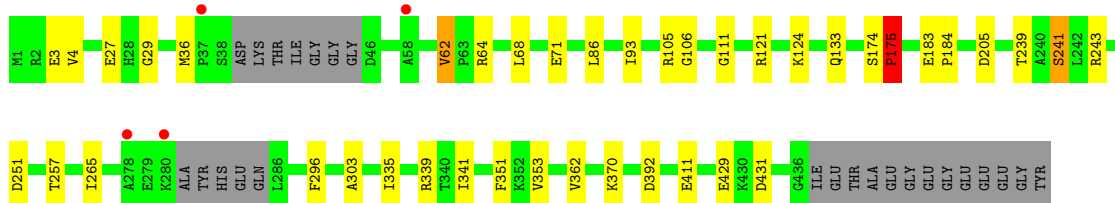
- Molecule 1: Tubulin alpha chain

Chain A: 

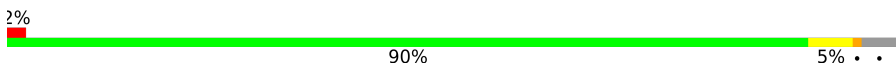


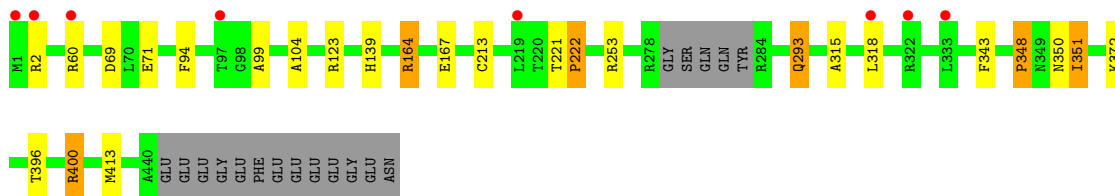
- Molecule 1: Tubulin alpha chain

Chain C: 




- Molecule 2: Tubulin beta chain

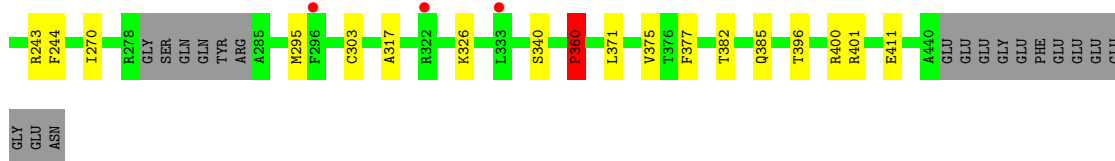
Chain B: 



- Molecule 2: Tubulin beta chain

Chain D: 

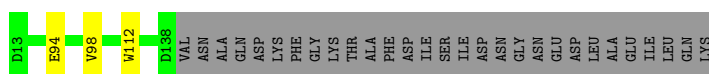




GLY
GLU
ASN

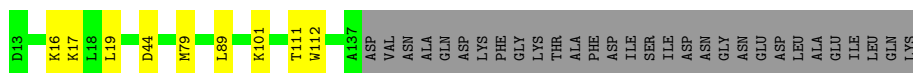
- Molecule 3: Darpin D1

Chain F: 79% 19%



- Molecule 3: Darpin D1

Chain E: 75% 6% 19%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	52.13Å 183.92Å 118.32Å 90.00° 92.44° 90.00°	Depositor
Resolution (Å)	46.83 – 1.75 46.83 – 1.75	Depositor EDS
% Data completeness (in resolution range)	98.4 (46.83-1.75) 98.4 (46.83-1.75)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.33 (at 1.75Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.181 , 0.213 0.182 , 0.214	Depositor DCC
R_{free} test set	10973 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	30.5	Xtrriage
Anisotropy	0.630	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	0.064 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	31317	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: MG, GTP

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	0.66	3/3429 (0.1%)	0.70	1/4642 (0.0%)
1	C	0.61	3/3380 (0.1%)	0.69	1/4577 (0.0%)
2	B	0.73	9/3495 (0.3%)	0.80	3/4728 (0.1%)
2	D	0.71	6/3417 (0.2%)	0.78	10/4625 (0.2%)
3	E	0.51	0/953	0.66	1/1298 (0.1%)
3	F	0.47	0/944	0.64	0/1286
All	All	0.66	21/15618 (0.1%)	0.74	16/21156 (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
2	D	0	1

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	175	PRO	N-CA	13.16	1.69	1.47
2	D	175	PRO	N-CA	12.90	1.69	1.47
2	B	348	PRO	N-CA	12.87	1.69	1.47
1	A	175	PRO	N-CA	12.51	1.68	1.47
2	D	162	PRO	N-CA	12.44	1.68	1.47

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	360	PRO	O-C-N	-10.01	106.69	122.70
2	B	222	PRO	CA-N-CD	-9.85	97.71	111.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	162	PRO	CA-N-CD	-7.75	100.65	111.50
2	D	360	PRO	CA-C-N	7.39	133.46	117.20
2	D	243	ARG	NE-CZ-NH2	-7.29	116.65	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	360	PRO	Mainchain

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	3309	3265	3215	20	0
1	C	3301	3250	3234	29	0
2	B	3378	3295	3248	23	0
2	D	3326	3227	3200	35	0
3	E	931	946	940	8	0
3	F	930	937	940	2	0
4	A	32	9	12	0	0
4	B	32	10	12	1	0
4	C	32	10	12	0	0
4	D	32	9	12	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
6	A	241	0	0	1	0
6	B	280	0	0	5	0
6	C	228	0	0	3	0
6	D	160	0	0	2	0
6	E	77	0	0	1	0
6	F	67	0	0	0	0
All	All	16359	14958	14825	114	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 4.

The worst 5 of 114 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:348:PRO:CA	2:B:348:PRO:N	1.69	1.46
1:C:175:PRO:CA	1:C:175:PRO:N	1.69	1.43
2:B:222:PRO:CA	2:B:222:PRO:N	1.68	1.41
1:A:175:PRO:N	1:A:175:PRO:CA	1.68	1.38
2:D:175:PRO:N	2:D:175:PRO:CA	1.69	1.32

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	425/449 (95%)	413 (97%)	12 (3%)	0	100	100
1	C	420/449 (94%)	406 (97%)	14 (3%)	0	100	100
2	B	432/443 (98%)	424 (98%)	8 (2%)	0	100	100
2	D	424/443 (96%)	415 (98%)	9 (2%)	0	100	100
3	E	125/155 (81%)	125 (100%)	0	0	100	100
3	F	124/155 (80%)	124 (100%)	0	0	100	100
All	All	1950/2094 (93%)	1907 (98%)	43 (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	366/376 (97%)	362 (99%)	4 (1%)	73	60
1	C	361/376 (96%)	357 (99%)	4 (1%)	73	60
2	B	372/376 (99%)	367 (99%)	5 (1%)	69	54
2	D	363/376 (96%)	357 (98%)	6 (2%)	60	42
3	E	97/120 (81%)	96 (99%)	1 (1%)	76	63
3	F	96/120 (80%)	96 (100%)	0	100	100
All	All	1655/1744 (95%)	1635 (99%)	20 (1%)	71	56

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	39	ASP
2	D	326	LYS
3	E	17	LYS
2	D	340	SER
2	B	167	GLU

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](#)

Of 7 ligands modelled in this entry, 3 are monoatomic - leaving 4 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, 95th 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 ²)	Q<0.9
1	A	422/449 (93%)	-0.02	4 (0%) 84 89	26, 41, 69, 100	0
1	C	424/449 (94%)	-0.05	4 (0%) 84 89	28, 42, 75, 126	0
2	B	425/443 (95%)	-0.05	8 (1%) 66 74	25, 36, 60, 84	0
2	D	424/443 (95%)	0.01	13 (3%) 49 55	30, 46, 77, 117	0
3	E	125/155 (80%)	-0.14	0 100 100	32, 44, 62, 85	0
3	F	126/155 (81%)	-0.16	0 100 100	31, 41, 62, 78	0
All	All	1946/2094 (92%)	-0.04	29 (1%) 73 80	25, 42, 71, 126	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	322	ARG	4.7
2	D	221	THR	3.9
2	D	296	PHE	3.6
2	B	1	MET	3.5
2	D	58	GLY	3.0

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, 95th 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(Å ²)	Q<0.9
5	MG	B	502	1/1	0.82	0.11	45,45,45,45	0
4	GTP	D	501	32/32	0.96	0.10	31,39,66,77	0
5	MG	A	502	1/1	0.97	0.10	31,31,31,31	0
4	GTP	B	501	32/32	0.97	0.09	26,31,38,49	0
4	GTP	C	502	32/32	0.98	0.10	29,33,40,45	0
5	MG	C	501	1/1	0.98	0.14	37,37,37,37	0
4	GTP	A	501	32/32	0.99	0.10	26,30,35,38	0

6.5 Other polymers [i](#)

There are no such residues in this entry.