wwPDB EM Validation Summary Report

Dec 15, 2024 – 03:06 PM EST

PDB ID	:	1TUB
Title	:	TUBULIN ALPHA-BETA DIMER, ELECTRON DIFFRACTION
Authors	:	Nogales, E.; Downing, K.H.
Deposited on	:	1997-09-23
Resolution	:	3.70 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

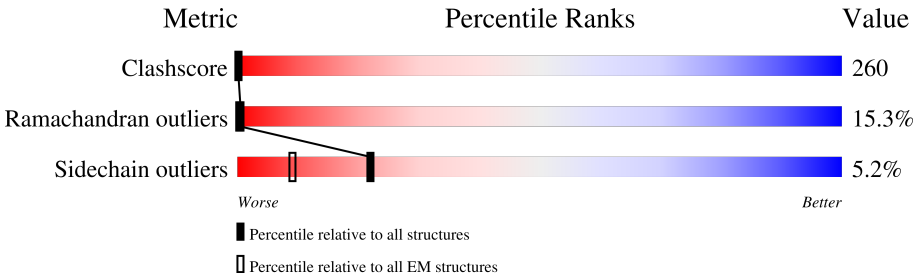
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
ELECTRON CRYSTALLOGRAPHY

The reported resolution of this entry is 3.70 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	440	8% 50% 27% 15%
2	B	427	5% 49% 29% 17%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GTP	A	500	-	-	X	-
4	GDP	B	500	-	-	X	-
5	TXL	B	501	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6907 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	440	3430	2168	583	657	22	0	0

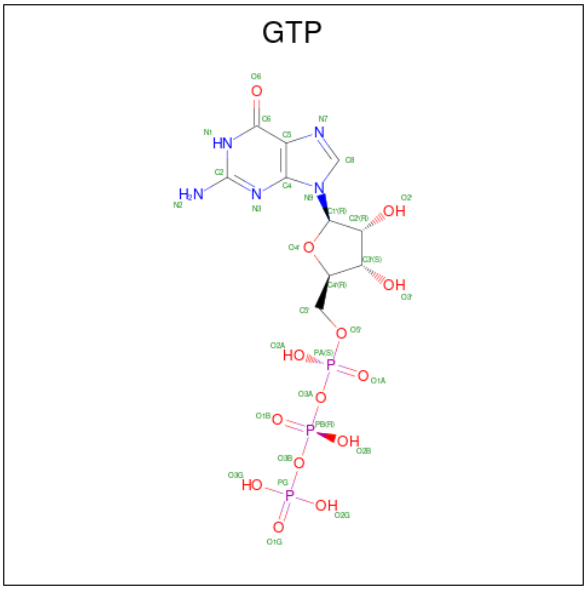
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	265	GLY	ALA	conflict	UNP P02550

- Molecule 2 is a protein called TUBULIN.

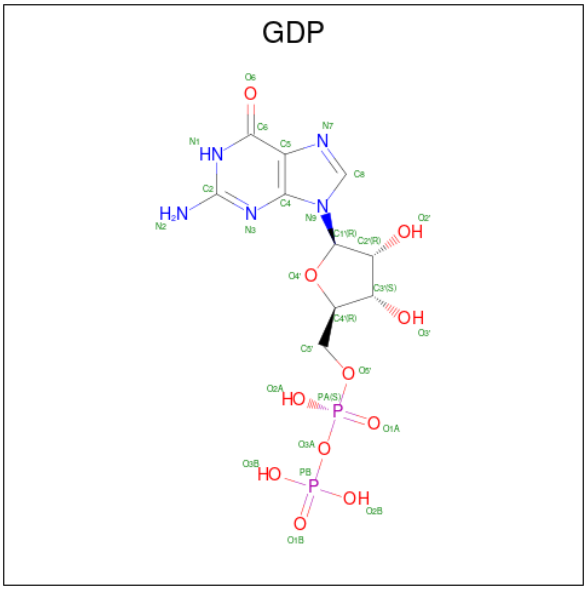
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	427	3359	2110	576	647	26	0	0

- Molecule 3 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



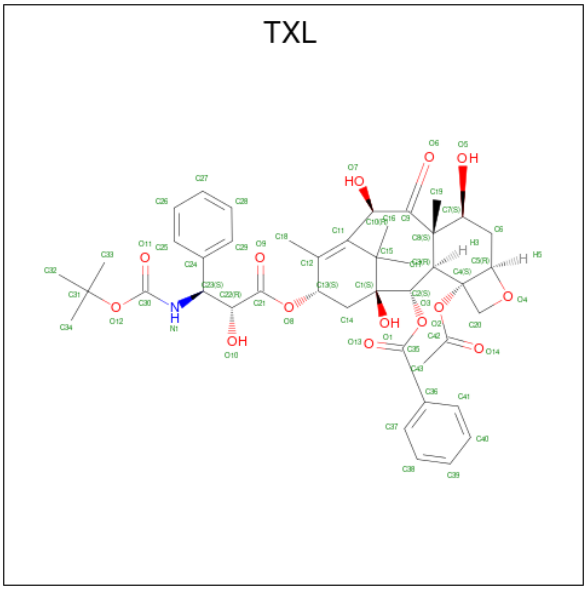
Mol	Chain	Residues	Atoms					AltConf
3	A	1	Total	C	N	O	P	0
			32	10	5	14	3	

- Molecule 4 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).



Mol	Chain	Residues	Atoms					AltConf
4	B	1	Total	C	N	O	P	0
			28	10	5	11	2	

- Molecule 5 is TAXOTERE (three-letter code: TXL) (formula: $C_{43}H_{53}NO_{14}$).



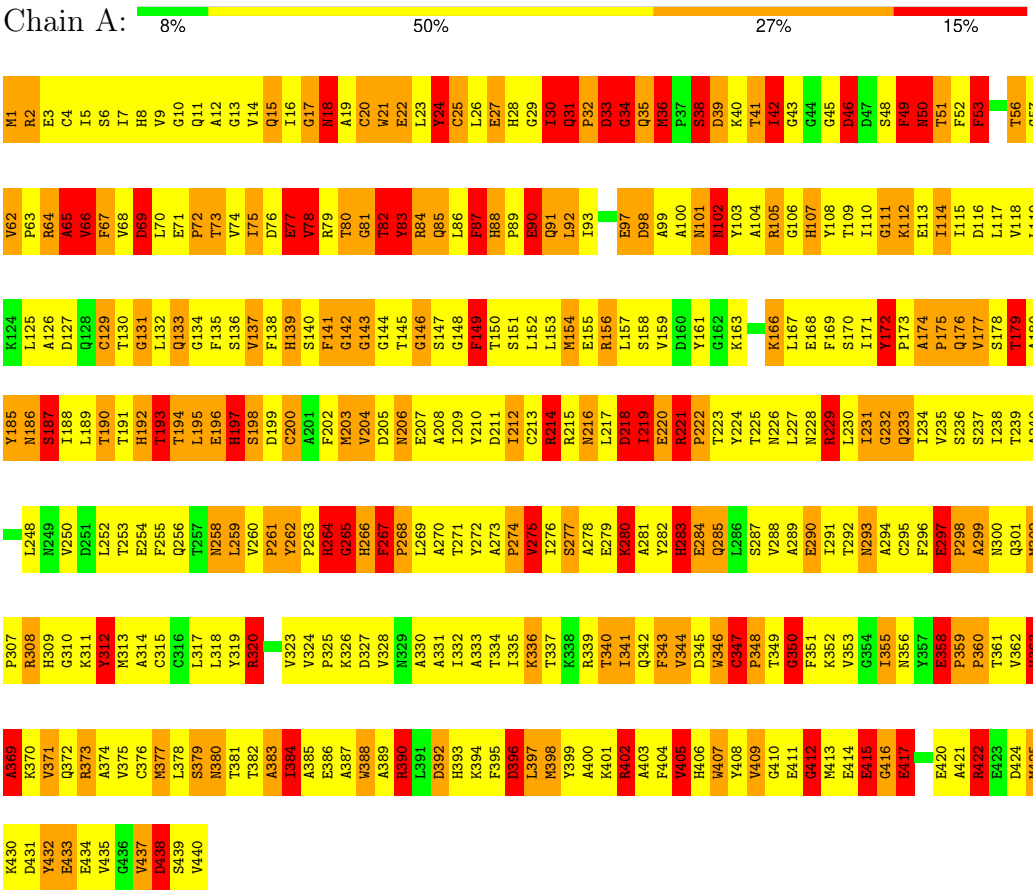
Mol	Chain	Residues	Atoms				AltConf
5	B	1	Total	C	N	O	0
			58	43	1	14	

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. 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. 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.

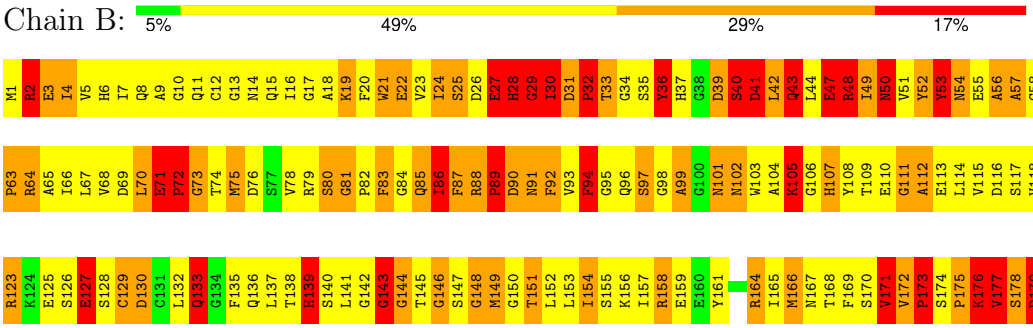
- Molecule 1: TUBULIN

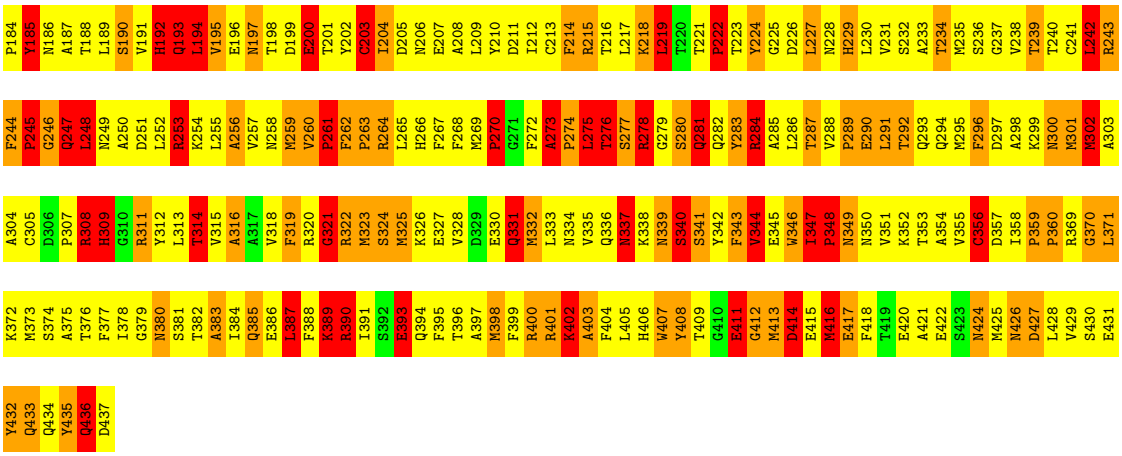
Chain A:



- Molecule 2: TUBULIN

Chain B:




P184 V185 N186 A187 T188 L189 S190 V191 H192 Q193 L194 V195 E196 N197 T198 D199 E200 T201 Y202 C203 I204 D205 N206 E207 A208 L209 Y210 D211 I212 C213 F214 R215 T216 L217 K218 L219 T220 T221 P222 T223 Y224 G225 D226 L227 M228 H229 L230 V231 S232 A233 T234 M235 S236 G237 V238 T239 T240 C241 L242 R243
F244 P245 G246 Q247 L248 N249 A250 D251 L252 R253 K254 L255 A256 V257 N258 M259 V260 P261 F262 P263 R264 L265 H266 F267 F268 M269 P270 G271 F272 A273 P274 L275 T276 S277 R278 G279 S280 Q281 Q282 Y283 R284 A285 L286 T287 V288 P289 E290 L291 T292 Q293 Q294 M295 F296 D297 A298 K299 N300 M301 M302 A303
A304 C305 D306 P307 R308 H309 G310 R311 Y312 L313 T314 V315 A316 A317 V318 F319 R320 G321 R322 M323 S324 M325 K326 E327 V328 D329 E330 Q331 M332 L333 N334 V335 Q336 N337 K338 N339 S340 S341 Y342 F343 V344 E345 W346 I347 P348 N349 N350 V351 K352 T353 A354 V355 C356 D357 I358 P359 P360 R369 G370 L371
K372 M373 S374 A375 T376 F377 I378 G379 N380 S381 T382 A383 I384 Q385 E386 L387 F388 K389 R390 I391 S392 E393 Q394 F395 T396 A397 M398 F399 R400 R401 K402 A403 F404 L405 H406 W407 Y408 T409 G410 E411 G412 M413 D414 E415 M416 E417 F418 T419 E420 A421 E422 S423 N424 M425 N426 D427 L428 V429 S430 E431
Y432 Q433 Q434 Y435 Q436 D437

4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	80.00Å 92.00Å 90.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 3.70	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-3.70)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Wilson B-factor (Å^2)	144.0	Xtriage
Anisotropy	0.504	Xtriage
L-test for twinning[1]	$< \|L\| > = 0.53$, $< L^2 > = 0.36$	Xtriage
Estimated twinning fraction	0.000 for -h,-l,-k 0.000 for -h,l,k 0.006 for h,-k,-l	Xtriage
Total number of atoms	6907	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 61.95 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2123e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	2.34	110/3508 (3.1%)	2.76	214/4762 (4.5%)
2	B	2.47	111/3434 (3.2%)	3.07	266/4652 (5.7%)
All	All	2.41	221/6942 (3.2%)	2.92	480/9414 (5.1%)

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	56
2	B	0	59
All	All	0	115

The worst 5 of 221 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	278	ARG	CA-CB	34.96	2.30	1.53
2	B	105	LYS	C-N	-29.38	0.80	1.33
2	B	73	GLY	C-N	-28.00	0.69	1.34
1	A	38	SER	C-N	-27.48	0.70	1.34
1	A	347	CYS	C-N	-23.43	0.89	1.34

The worst 5 of 480 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	368	LEU	O-C-N	-54.19	36.00	122.70
1	A	363	VAL	C-N-CD	-48.68	13.51	120.60
2	B	273	ALA	C-N-CD	-46.48	18.35	120.60
2	B	105	LYS	O-C-N	-44.51	47.53	123.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	88	ARG	C-N-CD	-44.07	23.65	120.60

There are no chirality outliers.

5 of 115 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	18	ASN	Mainchain
1	A	24	TYR	Mainchain,Peptide
1	A	30	ILE	Mainchain,Peptide
1	A	34	GLY	Peptide
1	A	36	MET	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	3430	0	3257	1657	34
2	B	3359	0	3172	1921	15
3	A	32	0	11	30	0
4	B	28	0	12	17	0
5	B	58	0	51	59	0
All	All	6907	0	6503	3480	35

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 260.

The worst 5 of 3480 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:229:HIS:CE1	5:B:501:TXL:H343	1.28	1.68
2:B:346:TRP:CE3	2:B:347:ILE:HG13	1.25	1.66
1:A:212:ILE:HD11	1:A:230:LEU:CD2	1.25	1.65
2:B:151:THR:CB	2:B:192:HIS:CD2	1.75	1.62
1:A:115:ILE:CD1	1:A:152:LEU:HG	1.15	1.62

The worst 5 of 35 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:283:HIS:CG	1:A:420:GLU:CD[2_444]	0.73	1.47
1:A:1:MET:CE	2:B:72:PRO:CG[1_655]	1.00	1.20
1:A:283:HIS:CG	1:A:420:GLU:OE1[2_444]	1.01	1.19
1:A:283:HIS:ND1	1:A:420:GLU:CD[2_444]	1.11	1.09
1:A:283:HIS:CD2	1:A:420:GLU:CA[2_444]	1.18	1.02

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 PDB entries followed by that with respect to all EM entries.

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	438/440 (100%)	323 (74%)	62 (14%)	53 (12%)	0 4
2	B	425/427 (100%)	297 (70%)	49 (12%)	79 (19%)	0 1
All	All	863/867 (100%)	620 (72%)	111 (13%)	132 (15%)	0 3

5 of 132 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	31	GLN
1	A	33	ASP
1	A	35	GLN
1	A	39	ASP
1	A	42	ILE

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 PDB entries followed by that with respect to all EM entries.

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	369/369 (100%)	351 (95%)	18 (5%)	21	47
2	B	368/368 (100%)	348 (95%)	20 (5%)	18	46
All	All	737/737 (100%)	699 (95%)	38 (5%)	22	46

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	300	ASN
2	B	426	ASN
2	B	302	MET
2	B	390	ARG
2	B	436	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	336	GLN
2	B	337	ASN
2	B	406	HIS
1	A	380	ASN
1	A	356	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	GDP	B	500	-	25,30,30	1.59	5 (20%)	30,47,47	1.06	3 (10%)
3	GTP	A	500	-	29,34,34	2.32	9 (31%)	35,54,54	1.77	7 (20%)
5	TXL	B	501	-	63,63,63	4.48	48 (76%)	100,100,100	3.47	55 (55%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GDP	B	500	-	-	4/12/32/32	0/3/3/3
3	GTP	A	500	-	-	1/18/38/38	0/3/3/3
5	TXL	B	501	-	-	0/38/124/124	0/6/6/6

The worst 5 of 62 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	501	TXL	O3-C4	-10.05	1.26	1.46
5	B	501	TXL	C41-C36	9.35	1.53	1.39
5	B	501	TXL	C37-C36	-8.74	1.26	1.39
5	B	501	TXL	C8-C3	-8.13	1.38	1.57
5	B	501	TXL	O6-C9	-8.05	1.08	1.21

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	501	TXL	C39-C38-C37	11.06	133.89	120.24
5	B	501	TXL	C15-C1-C2	10.89	123.89	111.93
5	B	501	TXL	C38-C37-C36	-9.75	110.78	120.36
5	B	501	TXL	O7-C10-C9	7.21	121.53	109.51
5	B	501	TXL	C39-C40-C41	-6.27	112.51	120.24

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	500	GDP	PA-O3A-PB-O2B
4	B	500	GDP	C5'-O5'-PA-O1A
4	B	500	GDP	PA-O3A-PB-O1B
4	B	500	GDP	PA-O3A-PB-O3B
3	A	500	GTP	PA-O3A-PB-O2B

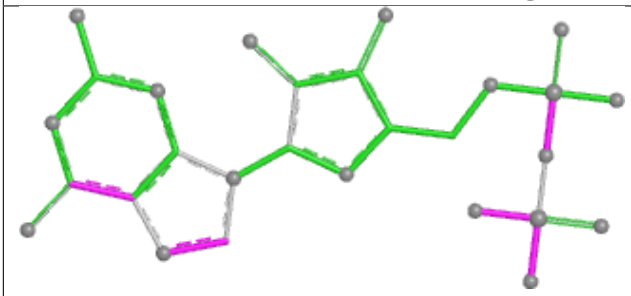
There are no ring outliers.

3 monomers are involved in 106 short contacts:

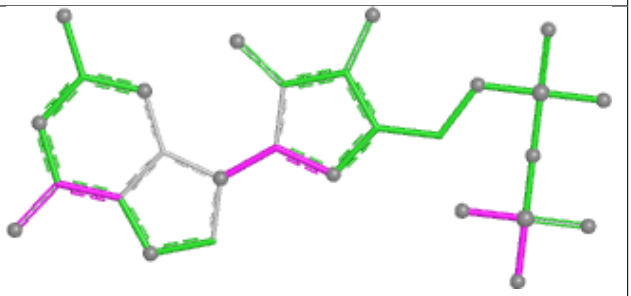
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	500	GDP	17	0
3	A	500	GTP	30	0
5	B	501	TXL	59	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

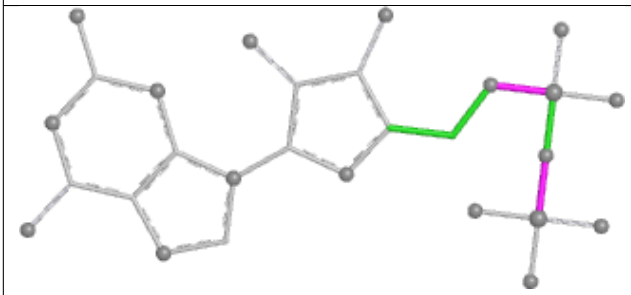
Ligand GDP B 500



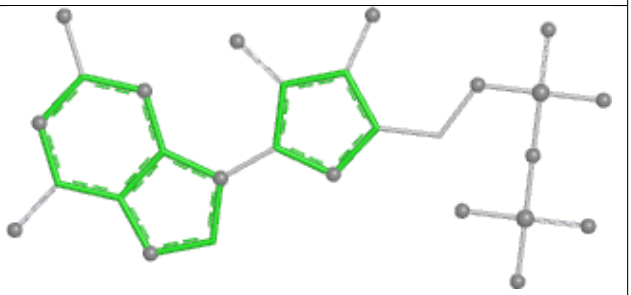
Bond lengths



Bond angles

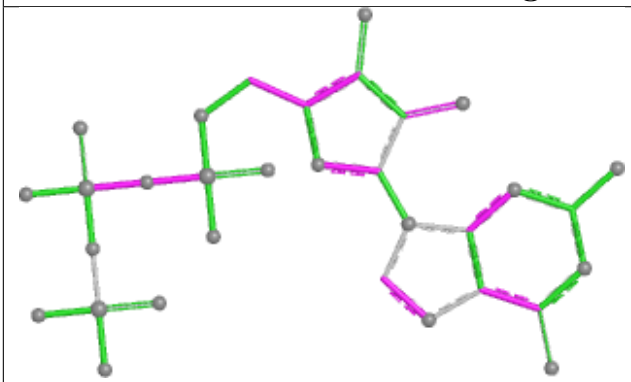


Torsions

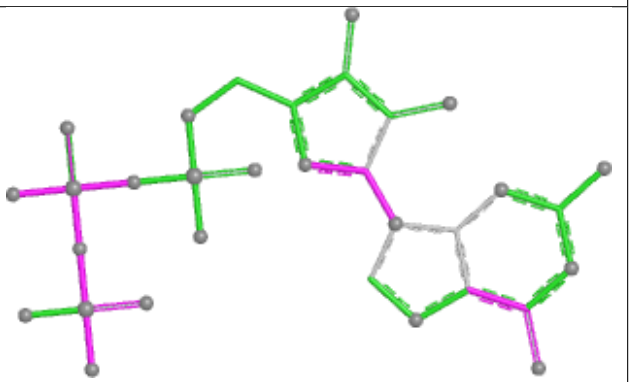


Rings

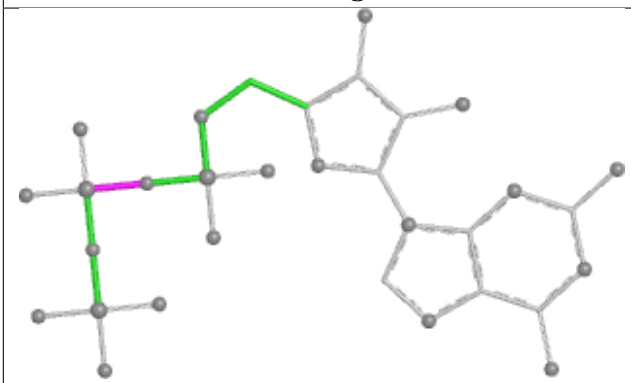
Ligand GTP A 500



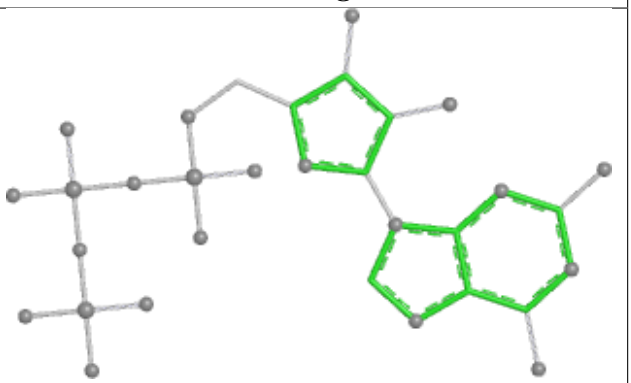
Bond lengths



Bond angles

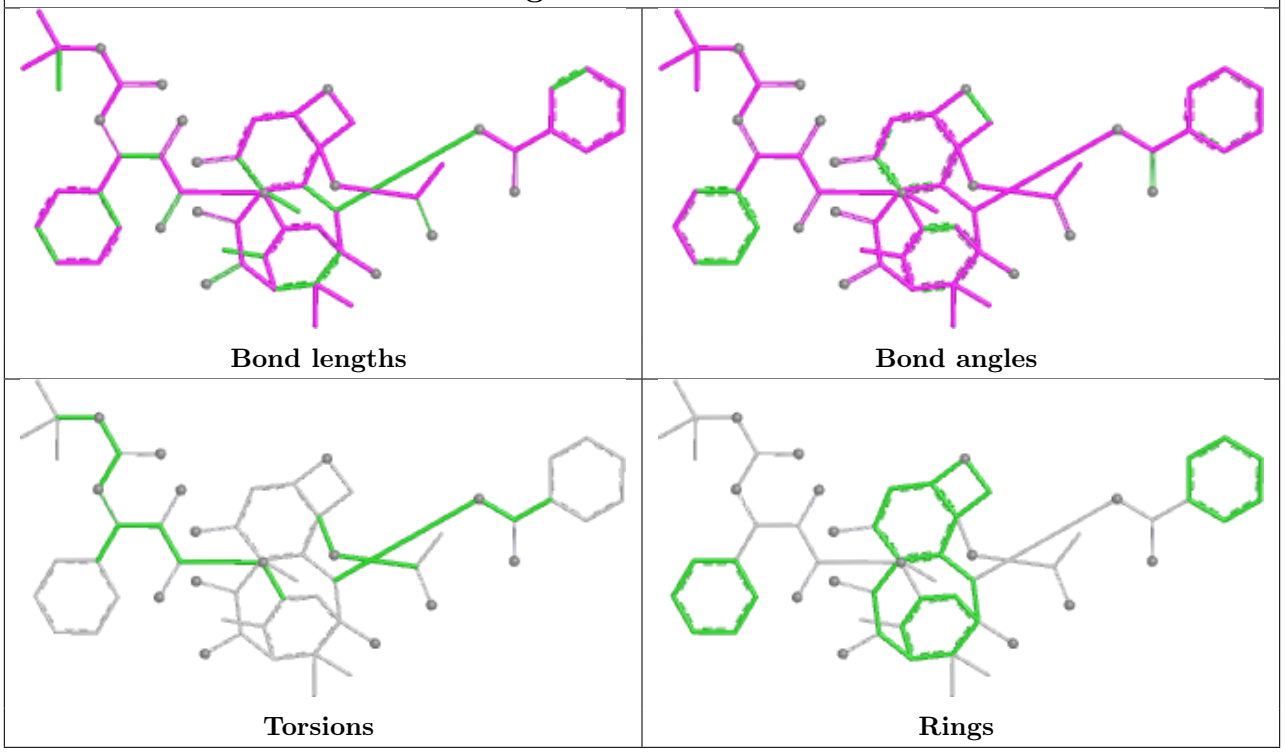


Torsions



Rings

Ligand TXL B 501



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	52
1	A	41

The worst 5 of 93 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	387:LEU	C	388:PHE	N	1.63
1	A	283:HIS	C	284:GLU	N	1.20
1	A	380:ASN	C	381:THR	N	1.20
1	A	437:VAL	C	438:ASP	N	1.20
1	B	52:TYR	C	53:TYR	N	1.20