



wwPDB X-ray Structure Validation Summary Report

Mar 31, 2014 – 03:15 PM BST

PDB ID : 4O4J
Title : Tubulin-Peloruside A complex
Authors : Prota, A.E.; Bargsten, K.; Northcote, P.T.; Marsh, M.; Altmann, K.H.; Miller, J.H.; Diaz, J.F.; Steinmetz, M.O.
Deposited on : 2013-12-18
Resolution : 2.20 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

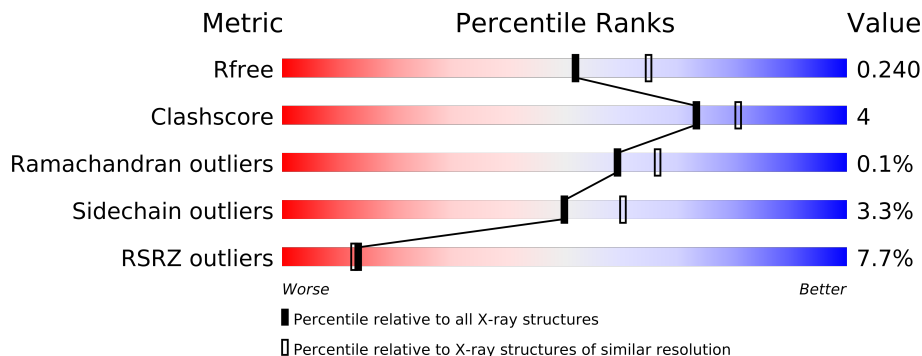
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable23004
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable23004

1 Overall quality at a glance

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}	66092	2938 (2.20-2.20)
Clashscore	79885	3751 (2.20-2.20)
Ramachandran outliers	78287	3681 (2.20-2.20)
Sidechain outliers	78261	3682 (2.20-2.20)
RSRZ outliers	66119	2939 (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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	451	
1	C	451	
2	B	445	
2	D	445	
3	E	143	
4	F	384	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
6	MG	C	502	-	X

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 18124 atoms, of which 0 are hydrogen and 0 are deuterium.

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
1	A	439	Total	C	N	O	S	0	12	0
			3485	2213	587	661	24			
1	C	440	Total	C	N	O	S	0	15	0
			3496	2218	585	667	26			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	428	Total	C	N	O	S	0	12	0
			3424	2151	581	664	28			
2	D	431	Total	C	N	O	S	0	6	0
			3413	2145	580	659	29			

- Molecule 3 is a protein called Stathmin-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	123	Total	C	N	O	S	0	6	0
			1044	645	188	206	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	3	MET	ILE	CLONING ARTIFACT	UNP P63043
E	4	ALA	SER	CLONING ARTIFACT	UNP P63043

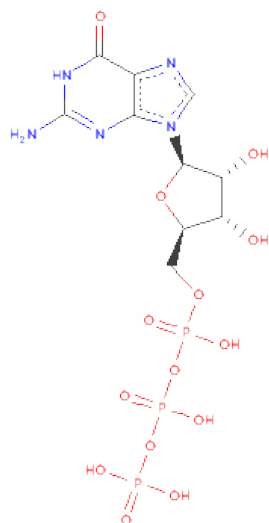
- Molecule 4 is a protein called Tubulin-tyrosine ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	265	Total	C	N	O	S	0	1	0
			2179	1409	370	386	14			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	379	HIS	-	EXPRESSION TAG	UNP E1BQ43
F	380	HIS	-	EXPRESSION TAG	UNP E1BQ43
F	381	HIS	-	EXPRESSION TAG	UNP E1BQ43
F	382	HIS	-	EXPRESSION TAG	UNP E1BQ43
F	383	HIS	-	EXPRESSION TAG	UNP E1BQ43
F	384	HIS	-	EXPRESSION TAG	UNP E1BQ43

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



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

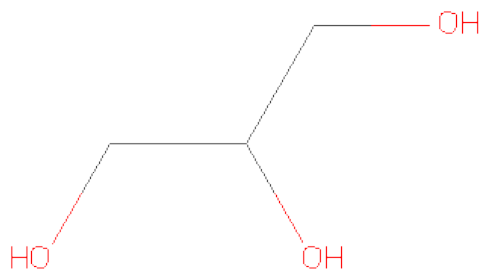
- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

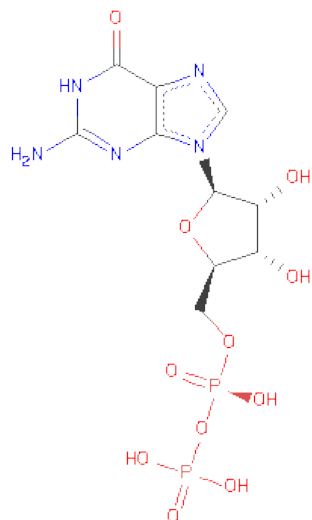
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Ca	0	0
			1	1		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



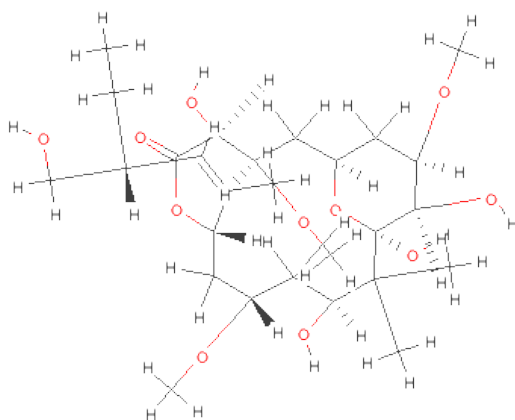
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 9 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: C₁₀H₁₅N₅O₁₁P₂).



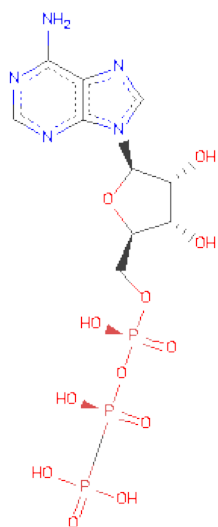
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	B	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
9	D	1	Total	C	N	O	P	0	0
			28	10	5	11	2		

- Molecule 10 is PELORUSIDE A (three-letter code: POU) (formula: $C_{27}H_{48}O_{11}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	B	1	Total	C	O	0	0
			38	27	11		
10	D	1	Total	C	O	0	0
			38	27	11		

- Molecule 11 is PHOSPHOMETHYLPHOSPHONICACID ADENYLATE ESTER (three-letter code: ACP) (formula: $C_{11}H_{18}N_5O_{12}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
11	F	1	Total	C	N	O	P	0	0
			31	11	5	12	3		

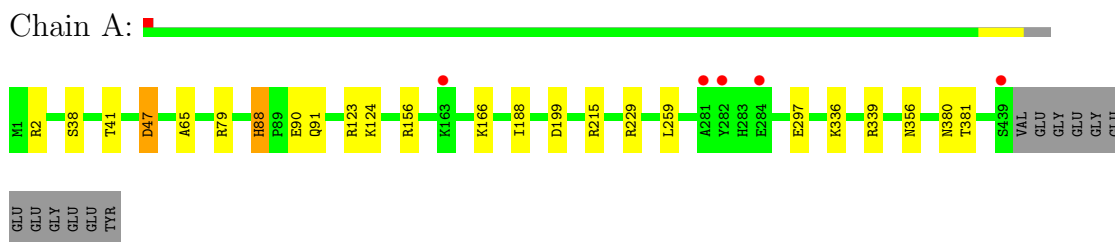
- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	191	Total	O	0	0
			191	191		
12	B	210	Total	O	0	0
			210	210		
12	C	264	Total	O	0	0
			264	264		
12	D	116	Total	O	0	0
			116	116		
12	E	43	Total	O	0	0
			43	43		
12	F	21	Total	O	0	0
			21	21		

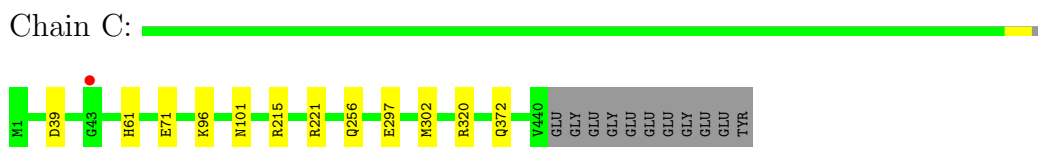
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

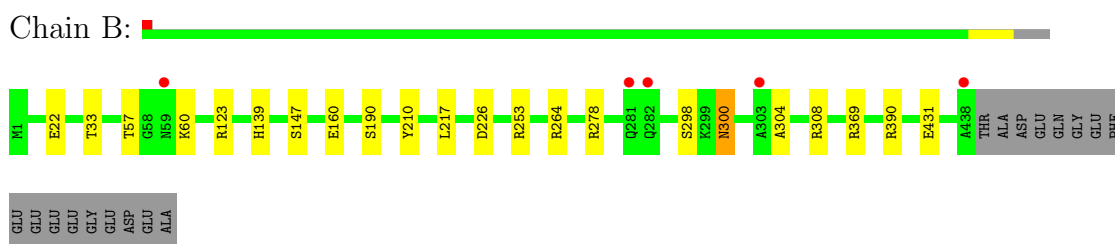
- Molecule 1: Tubulin alpha-1B chain



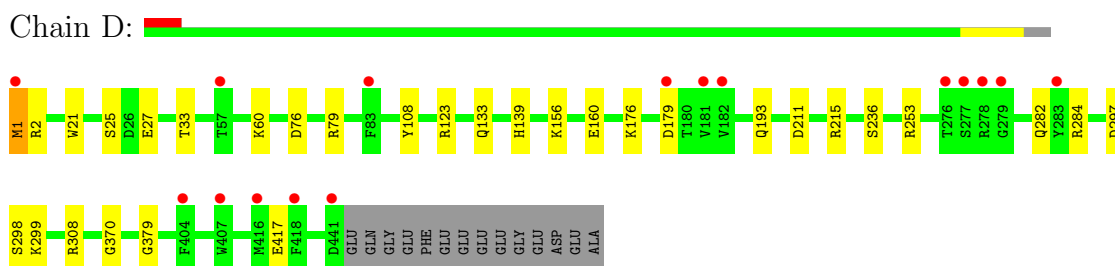
- Molecule 1: Tubulin alpha-1B chain



- Molecule 2: Tubulin beta-2B chain

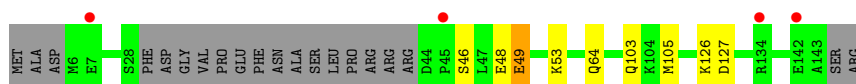


- Molecule 2: Tubulin beta-2B chain



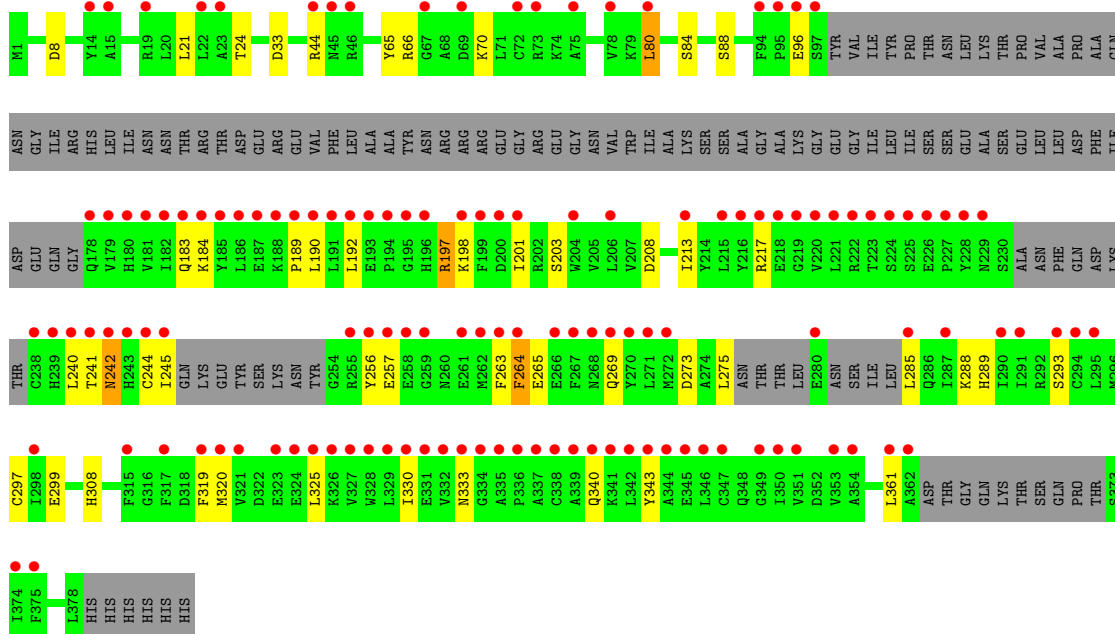
- Molecule 3: Stathmin-4





● Molecule 4: Tubulin-tyrosine ligase

Chain F:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	104.37Å 157.84Å 180.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	72.30 – 2.20 72.30 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (72.30-2.20) 99.7 (72.30-2.20)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.53 (at 2.20Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.8.2_1309)	Depositor
R, R_{free}	0.201 , 0.239 0.201 , 0.240	Depositor DCC
R_{free} test set	7566 reflections (5.02%)	DCC
Wilson B-factor (Å ²)	39.0	Xtriage
Anisotropy	0.131	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 33.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 150728 reflections (0.001%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	18124	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.31% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GDP, GOL, MG, CA, POU, GTP, ACP

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	0.29	0/3599	0.47	0/4887
1	C	0.35	0/3619	0.51	0/4917
2	B	0.31	0/3532	0.48	0/4783
2	D	0.27	0/3506	0.45	0/4751
3	E	0.28	0/1071	0.38	0/1423
4	F	0.25	0/2230	0.46	0/3006
All	All	0.30	0/17557	0.47	0/23767

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3485	0	0	12	0
1	C	3496	0	0	7	0
2	B	3424	0	0	13	0
2	D	3413	0	0	18	0
3	E	1044	0	0	5	0
4	F	2179	0	0	15	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	32	0	12	0	0
5	C	32	0	12	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
7	A	1	0	0	0	0
8	A	6	0	8	0	0
9	B	28	0	12	0	0
9	D	28	0	12	0	0
10	B	38	0	48	3	0
10	D	38	0	48	2	0
11	F	31	0	14	1	0
12	A	191	0	0	6	0
12	B	210	0	0	8	0
12	C	264	0	0	5	0
12	D	116	0	0	4	0
12	E	43	0	0	2	0
12	F	21	0	0	1	0
All	All	18124	0	166	73	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 4.

The worst 5 of 73 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:253:ARG:NE	12:B:735:HOH:O	2.20	0.72
4:F:197:ARG:NH1	4:F:257:GLU:OE1	2.23	0.71
2:B:22:GLU:OE2	12:B:802:HOH:O	2.09	0.70
2:D:176:LYS:NZ	2:D:211:ASP:OD1	2.24	0.70
1:C:372:GLN:NE2	12:C:786:HOH:O	2.25	0.70

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	449/451 (100%)	440 (98%)	9 (2%)	0	100	100
1	C	453/451 (100%)	441 (97%)	12 (3%)	0	100	100
2	B	438/445 (98%)	430 (98%)	8 (2%)	0	100	100
2	D	435/445 (98%)	421 (97%)	14 (3%)	0	100	100
3	E	125/143 (87%)	123 (98%)	2 (2%)	0	100	100
4	F	253/384 (66%)	239 (94%)	12 (5%)	2 (1%)	27	24
All	All	2153/2319 (93%)	2094 (97%)	57 (3%)	2 (0%)	59	66

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	F	264	PHE
4	F	189	PRO

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	382/379 (101%)	372 (97%)	10 (3%)	59	70
1	C	386/379 (102%)	382 (99%)	4 (1%)	85	93
2	B	382/383 (100%)	376 (98%)	6 (2%)	75	85
2	D	378/383 (99%)	372 (98%)	6 (2%)	75	85
3	E	116/127 (91%)	112 (97%)	4 (3%)	49	59
4	F	241/342 (70%)	211 (88%)	30 (12%)	7	5
All	All	1885/1993 (95%)	1825 (97%)	60 (3%)	50	62

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	49	GLU
4	F	183	GLN
4	F	330	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	F	33	ASP
4	F	192	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 13 ligands modelled in this entry, 5 are monoatomic - leaving 8 for Mogul analysis.

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$
5	GTP	A	501	6	34,34,34	1.09	1 (2%)	51,54,54	1.88	11 (21%)
8	GOL	A	504	-	5,5,5	0.34	0	5,5,5	0.27	0
9	GDP	B	501	6	30,30,30	1.48	5 (16%)	44,47,47	3.05	7 (15%)
10	POU	B	503	-	39,39,39	2.50	7 (17%)	57,57,57	2.10	15 (26%)
5	GTP	C	501	6	34,34,34	0.94	1 (2%)	51,54,54	2.43	8 (15%)
9	GDP	D	501	-	30,30,30	1.39	4 (13%)	44,47,47	3.15	7 (15%)
10	POU	D	503	-	39,39,39	2.37	7 (17%)	57,57,57	1.90	13 (22%)
11	ACP	F	401	-	33,33,33	1.52	7 (21%)	52,52,52	1.96	11 (21%)

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
5	GTP	A	501	6	-	0/22/38/38	0/1/3/3
8	GOL	A	504	-	-	0/4/4/4	0/0/0/0
9	GDP	B	501	6	-	0/16/32/32	0/1/3/3
10	POU	B	503	-	-	1/56/76/76	0/0/2/2
5	GTP	C	501	6	-	0/22/38/38	0/1/3/3
9	GDP	D	501	-	-	0/16/32/32	0/1/3/3
10	POU	D	503	-	-	1/56/76/76	0/0/2/2
11	ACP	F	401	-	-	0/20/38/38	0/1/3/3

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	B	503	POU	C22-C21	10.97	1.56	1.33
10	D	503	POU	C22-C21	9.82	1.54	1.33
10	B	503	POU	C2-C1	-7.95	1.39	1.52
10	D	503	POU	C2-C1	-7.70	1.39	1.52
9	B	501	GDP	C6-C5	4.62	1.48	1.41

The worst 5 of 72 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	D	501	GDP	C6-C5-N7	17.83	136.54	134.14
9	B	501	GDP	C6-C5-N7	16.13	136.31	134.14
5	C	501	GTP	C6-C5-N7	-14.88	132.14	134.14
5	A	501	GTP	C6-C5-N7	-8.67	132.97	134.14
10	B	503	POU	C12-C11-C10	6.36	121.42	114.79

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
10	D	503	POU	C23-C22-C21-C15
10	B	503	POU	C23-C22-C21-C15

There are no ring outliers.

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	439/451 (97%)	-0.15	5 (1%) 77 78	17, 32, 66, 112	0
1	C	440/451 (97%)	-0.22	1 (0%) 93 94	11, 22, 45, 68	0
2	B	428/445 (96%)	-0.02	5 (1%) 75 76	13, 27, 60, 112	2 (0%)
2	D	431/445 (96%)	0.16	16 (3%) 39 40	18, 44, 77, 102	6 (1%)
3	E	123/143 (86%)	0.23	4 (3%) 44 45	25, 44, 85, 121	0
4	F	265/384 (69%)	2.32	132 (49%) 1 0	39, 79, 123, 153	0
All	All	2126/2319 (91%)	0.26	163 (7%) 13 13	11, 35, 92, 153	8 (0%)

The worst 5 of 163 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	F	194	PRO	12.1
4	F	182	ILE	9.2
4	F	192	LEU	7.8
4	F	240	LEU	7.6
4	F	195	GLY	7.6

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
6	MG	C	502	1/1	0.15	2.78	18,18,18,18	0
8	GOL	A	504	6/6	0.24	1.63	57,60,60,62	0
6	MG	B	502	1/1	0.15	1.19	9,9,9,9	0
10	POU	D	503	38/38	0.19	1.03	35,59,74,77	0
6	MG	D	502	1/1	0.16	0.77	67,67,67,67	0
5	GTP	C	501	32/32	0.13	0.62	4,13,19,30	0
7	CA	A	503	1/1	0.13	0.46	68,68,68,68	0
5	GTP	A	501	32/32	0.12	0.26	10,17,22,41	0
10	POU	B	503	38/38	0.12	0.22	23,44,77,81	0
9	GDP	B	501	28/28	0.14	0.21	11,16,20,22	0
9	GDP	D	501	28/28	0.10	-0.50	27,36,45,53	0
11	ACP	F	401	31/31	0.27	-0.96	67,106,145,166	0
6	MG	A	502	1/1	0.10	-1.06	14,14,14,14	0

6.5 Other polymers ⓘ

There are no such residues in this entry.