



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:10 PM GMT

PDB ID : 1J3K
Title : Quadruple mutant (N51I+C59R+S108N+I164L) Plasmodium falciparum dihydrofolate reductase-thymidylate synthase (PfDHFR-TS) complexed with WR99210, NADPH, and dUMP
Authors : Yuvaniyama, J.; Chitnumsub, P.; Kamchonwongpaisan, S.; Vanichtanankul, J.; Sirawaraporn, W.; Taylor, P.; Walkinshaw, M.; Yuthavong, Y.
Deposited on : 2003-02-03
Resolution : 2.10 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

<http://wwpdb.org/validation/2016/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	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

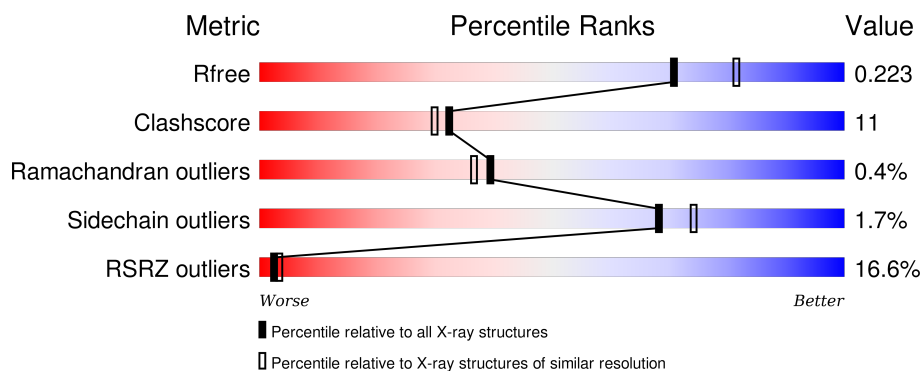
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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}	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (2.10-2.10)

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. 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	280	<div> <div>15%</div> <div> <div>62%</div> <div>15%</div> <div>•</div> <div>21%</div> </div> </div>
1	B	280	<div> <div>35%</div> <div> <div>48%</div> <div>29%</div> <div>•</div> <div>21%</div> </div> </div>
2	C	328	<div> <div>7%</div> <div> <div>91%</div> <div>8%</div> <div>••</div> </div> </div>
2	D	328	<div> <div>6%</div> <div> <div>87%</div> <div>12%</div> <div>••</div> </div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10528 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 Bifunctional dihydrofolate reductase-thymidylate synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	221	Total	C	N	O	S	0	0	0
			1837	1190	297	338	12			
1	B	221	Total	C	N	O	S	0	0	0
			1834	1189	297	336	12			

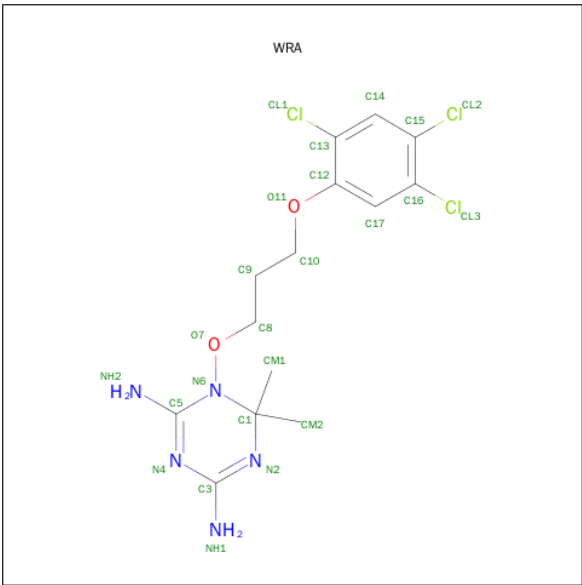
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	51	ILE	ASN	ENGINEERED	UNP P13922
A	59	ARG	CYS	SEE REMARK 999, ENGINEERED	UNP P13922
A	108	ASN	SER	SEE REMARK 999, ENGINEERED	UNP P13922
A	164	LEU	ILE	ENGINEERED	UNP P13922
B	51	ILE	ASN	ENGINEERED	UNP P13922
B	59	ARG	CYS	SEE REMARK 999, ENGINEERED	UNP P13922
B	108	ASN	SER	SEE REMARK 999, ENGINEERED	UNP P13922
B	164	LEU	ILE	ENGINEERED	UNP P13922

- Molecule 2 is a protein called Bifunctional dihydrofolate reductase-thymidylate synthase.

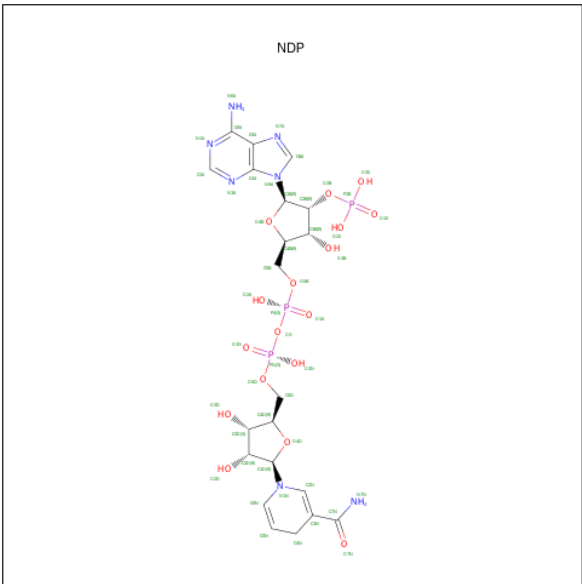
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	326	Total	C	N	O	S	0	0	0
			2713	1747	456	495	15			
2	D	326	Total	C	N	O	S	0	0	0
			2713	1747	456	495	15			

- Molecule 3 is 6,6-DIMETHYL-1-[3-(2,4,5-TRICHLOROPHENOXY)PROPOXY]-1,6-DIHYDRO-1,3,5-TRIAZINE-2,4-DIAMINE (three-letter code: WRA) (formula: C₁₄H₁₈Cl₃N₅O₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	Cl	N	O	0
			24	14	3	5	2	
3	B	1	Total	C	Cl	N	O	0
			24	14	3	5	2	

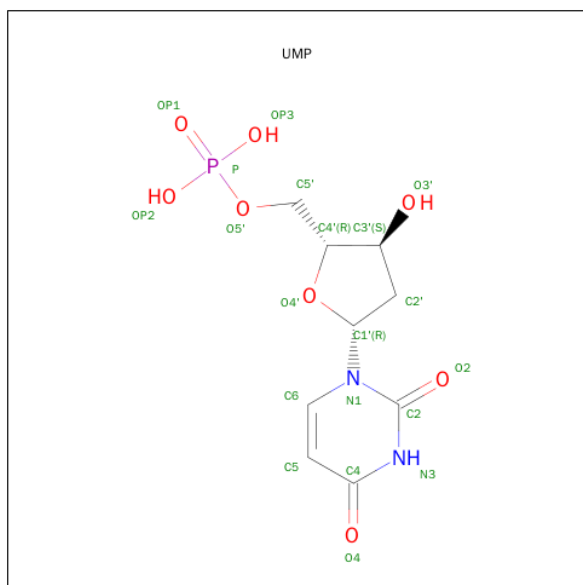
- Molecule 4 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: C₂₁H₃₀N₇O₁₇P₃).



Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 5 is 2'-DEOXYURIDINE 5'-MONOPHOSPHATE (three-letter code: UMP) (formula: $C_9H_{13}N_2O_8P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	C	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
5	D	1	Total	C	N	O	P	0	0
			20	9	2	8	1		

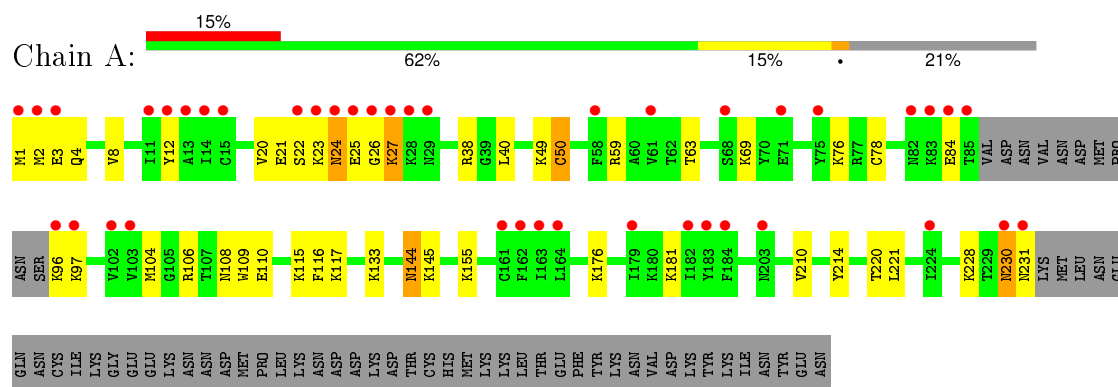
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	238	Total	O	0	0
			238	238		
6	B	102	Total	O	0	0
			102	102		
6	C	428	Total	O	0	0
			428	428		
6	D	479	Total	O	0	0
			479	479		

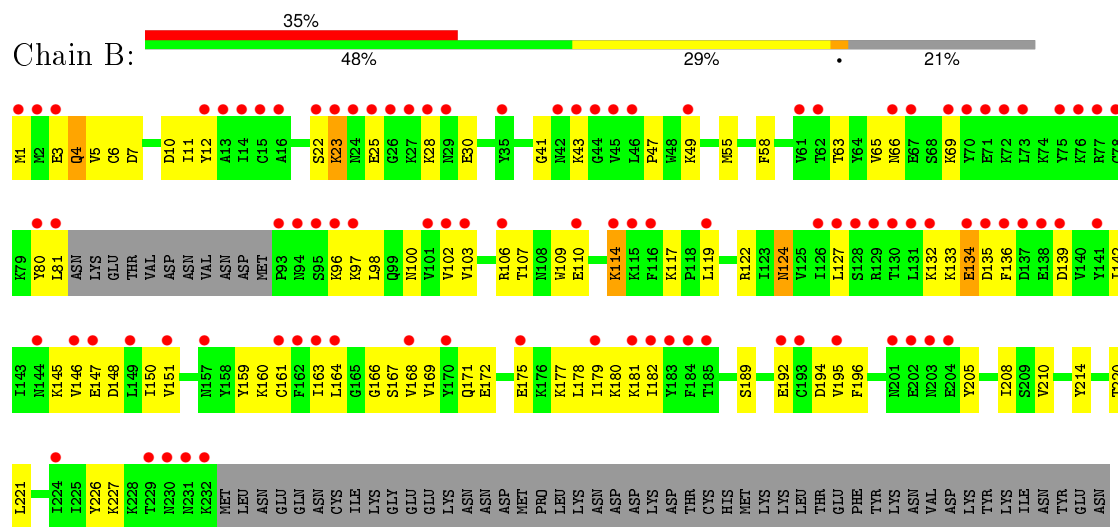
3 Residue-property plots [i](#)

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.

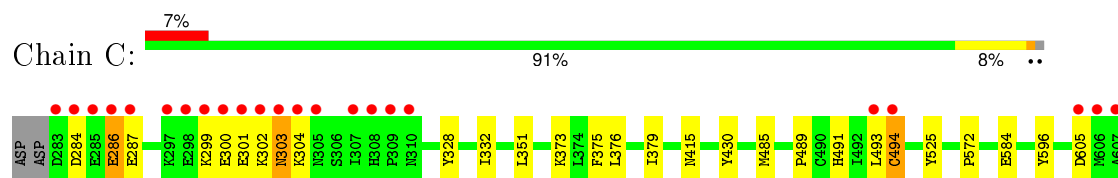
- Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase



- Molecule 1: Bifunctional dihydrofolate reductase-thymidylate synthase

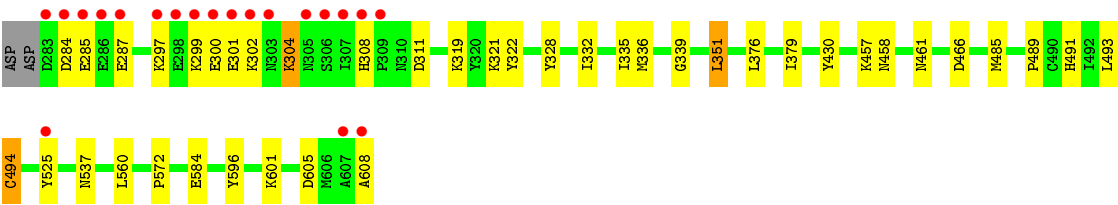
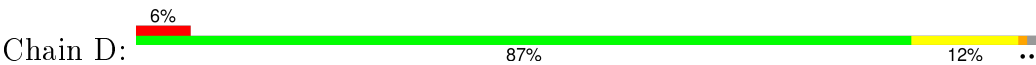


- Molecule 2: Bifunctional dihydrofolate reductase-thymidylate synthase



A608

● Molecule 2: Bifunctional dihydrofolate reductase-thymidylate synthase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	59.68Å 158.09Å 165.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.98 – 2.10 28.98 – 2.10	Depositor EDS
% Data completeness (in resolution range)	96.8 (28.98-2.10) 96.9 (28.98-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.43 (at 2.10Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.182 , 0.222 0.183 , 0.223	Depositor DCC
R_{free} test set	4520 reflections (5.03%)	DCC
Wilson B-factor (Å ²)	34.0	Xtriage
Anisotropy	0.446	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 65.5	EDS
Estimated twinning fraction	0.006 for -h,l,k	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 89989 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	10528	wwPDB-VP
Average B, all atoms (Å ²)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.27% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: NDP, UMP, WRA

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	0.64	0/1871	0.88	5/2515 (0.2%)
1	B	0.46	0/1869	0.76	1/2512 (0.0%)
2	C	0.74	1/2784 (0.0%)	0.85	2/3766 (0.1%)
2	D	0.73	1/2784 (0.0%)	0.84	1/3766 (0.0%)
All	All	0.67	2/9308 (0.0%)	0.83	9/12559 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	494	CYS	CB-SG	-9.88	1.65	1.82
2	D	494	CYS	CB-SG	-7.79	1.69	1.82

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	22	SER	N-CA-C	-8.83	87.17	111.00
2	C	304	LYS	N-CA-C	6.75	129.22	111.00
2	C	351	LEU	N-CA-C	-6.23	94.19	111.00
2	D	351	LEU	N-CA-C	-6.21	94.23	111.00
1	B	30	GLU	N-CA-C	5.85	126.79	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	1837	0	1873	42	0
1	B	1834	0	1873	86	0
2	C	2713	0	2638	22	0
2	D	2713	0	2638	44	0
3	A	24	0	18	1	0
3	B	24	0	18	0	0
4	A	48	0	26	4	0
4	B	48	0	26	6	0
5	C	20	0	11	1	0
5	D	20	0	11	0	0
6	A	238	0	0	12	1
6	B	102	0	0	11	0
6	C	428	0	0	4	1
6	D	479	0	0	13	0
All	All	10528	0	9132	194	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 11.

The worst 5 of 194 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:A:1:MET:HG3	1:A:2:MET:H	0.97	1.11
1:A:1:MET:HG3	1:A:2:MET:N	1.78	0.98
2:D:300:GLU:HG3	2:D:304:LYS:NZ	1.78	0.98
2:D:605:ASP:HB3	2:D:608:ALA:HB2	1.46	0.98
1:B:210:VAL:HB	6:B:2028:HOH:O	1.64	0.95

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)
6:A:1592:HOH:O	6:C:1593:HOH:O[3_645]	2.18	0.02

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	217/280 (78%)	208 (96%)	8 (4%)	1 (0%)	34	30
1	B	217/280 (78%)	198 (91%)	18 (8%)	1 (0%)	34	30
2	C	324/328 (99%)	312 (96%)	11 (3%)	1 (0%)	46	45
2	D	324/328 (99%)	307 (95%)	16 (5%)	1 (0%)	46	45
All	All	1082/1216 (89%)	1025 (95%)	53 (5%)	4 (0%)	39	37

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	430	TYR
1	A	26	GLY
1	B	28	LYS
2	D	430	TYR

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	210/268 (78%)	203 (97%)	7 (3%)	45	47
1	B	210/268 (78%)	203 (97%)	7 (3%)	45	47
2	C	300/302 (99%)	298 (99%)	2 (1%)	88	92
2	D	300/302 (99%)	299 (100%)	1 (0%)	94	97
All	All	1020/1140 (90%)	1003 (98%)	17 (2%)	68	74

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	4	GLN
1	B	23	LYS
1	B	189	SER
1	A	230	ASN
2	C	286	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	157	ASN
1	B	171	GLN
2	D	394	ASN
1	B	99	GLN
2	C	424	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

6 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
3	WRA	A	609	-	22,25,25	2.77	13 (59%)	28,36,36	1.87	8 (28%)
4	NDP	A	610	-	42,52,52	2.71	17 (40%)	55,80,80	2.35	19 (34%)
3	WRA	B	709	-	22,25,25	3.00	12 (54%)	28,36,36	1.81	5 (17%)
4	NDP	B	710	-	42,52,52	2.74	21 (50%)	55,80,80	2.13	17 (30%)
5	UMP	C	611	-	16,21,21	1.89	3 (18%)	23,31,31	2.94	8 (34%)
5	UMP	D	711	-	16,21,21	1.92	4 (25%)	23,31,31	3.27	8 (34%)

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
3	WRA	A	609	-	-	0/7/27/27	0/2/2/2
4	NDP	A	610	-	-	0/30/77/77	0/5/5/5
3	WRA	B	709	-	-	0/7/27/27	0/2/2/2
4	NDP	B	710	-	-	0/30/77/77	0/5/5/5
5	UMP	C	611	-	-	0/6/22/22	0/2/2/2
5	UMP	D	711	-	-	0/6/22/22	0/2/2/2

The worst 5 of 70 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	710	NDP	C3B-C2B	-4.51	1.42	1.53
4	B	710	NDP	PA-O2A	-4.41	1.36	1.54
4	B	710	NDP	O4B-C1B	-3.91	1.36	1.41
4	A	610	NDP	P2B-O2X	-3.75	1.41	1.54
4	B	710	NDP	C3B-C4B	-3.70	1.43	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	610	NDP	C4B-O4B-C1B	-5.06	104.16	109.72
4	B	710	NDP	O3X-P2B-O1X	-5.01	94.44	110.58
4	A	610	NDP	C1D-N1N-C2N	-4.84	112.47	120.91
4	B	710	NDP	N3A-C2A-N1A	-4.80	125.22	128.89
4	A	610	NDP	N3A-C2A-N1A	-4.22	125.67	128.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	609	WRA	1	0
4	A	610	NDP	4	0
4	B	710	NDP	6	0
5	C	611	UMP	1	0

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, 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	221/280 (78%)	0.99	41 (18%) 2 2	23, 41, 87, 95	0
1	B	221/280 (78%)	2.34	97 (43%) 0 0	29, 73, 95, 95	0
2	C	326/328 (99%)	0.23	24 (7%) 17 24	21, 31, 90, 95	0
2	D	326/328 (99%)	0.14	20 (6%) 25 33	19, 32, 86, 95	0
All	All	1094/1216 (89%)	0.78	182 (16%) 2 3	19, 37, 94, 95	0

The worst 5 of 182 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	MET	11.9
1	A	2	MET	10.9
1	B	75	TYR	10.0
1	A	26	GLY	9.8
1	A	23	LYS	9.7

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](#)

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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	WRA	A	609	24/24	0.55	0.27	1.69	29,41,47,49	0
3	WRA	B	709	24/24	0.40	0.32	0.70	57,64,65,67	0
4	NDP	B	710	48/48	0.75	0.27	0.04	76,85,95,95	0
4	NDP	A	610	48/48	0.95	0.10	-0.70	32,40,49,50	0
5	UMP	C	611	20/20	0.98	0.08	-0.95	23,27,32,32	0
5	UMP	D	711	20/20	0.98	0.08	-1.26	27,31,36,37	0

6.5 Other polymers [i](#)

There are no such residues in this entry.