



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 09:22 PM GMT

PDB ID : 4E5O  
Title : Crystal structure of mouse thymidylate synthase in complex with dUMP  
Authors : Dowiercial, A.; Jarmula, A.; Rypniewski, W.; Sokolowska, M.; Fraczyk, T.;  
Ciesla, J.; Rode, W.  
Deposited on : 2012-03-14  
Resolution : 1.70 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

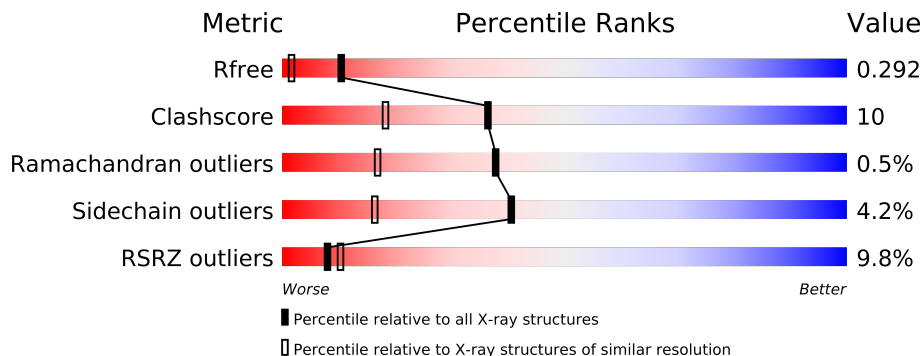
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
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)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.70 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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	2456 (1.70-1.70)
Clashscore	79885	2929 (1.70-1.70)
Ramachandran outliers	78287	2878 (1.70-1.70)
Sidechain outliers	78261	2878 (1.70-1.70)
RSRZ outliers	66119	2456 (1.70-1.70)

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  $\geq 3$ , 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	307	
1	B	307	
1	C	307	
1	D	307	
1	E	307	
1	F	307	

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

Mol	Type	Chain	Res	Geometry	Electron density
2	UMP	F	502	-	X
3	BU1	C	504	-	X

## 2 Entry composition i

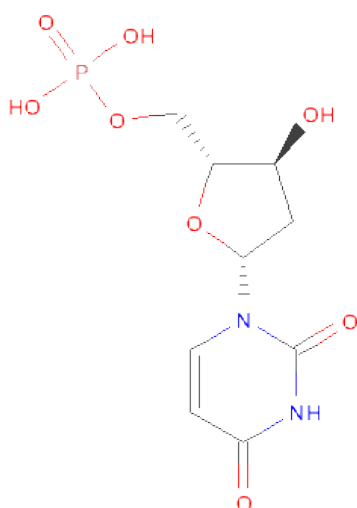
There are 4 unique types of molecules in this entry. The entry contains 15512 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Thymidylate synthase.

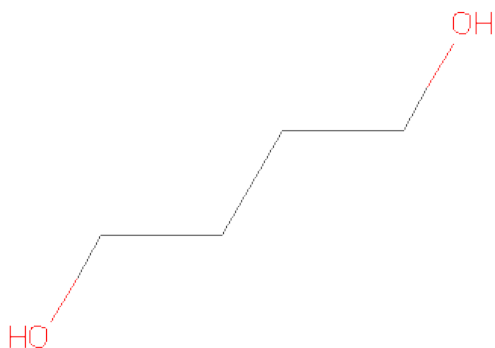
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	287	Total	C	N	O	S	0	3	0
			2346	1501	405	427	13			
1	B	288	Total	C	N	O	S	0	3	0
			2354	1505	409	427	13			
1	C	280	Total	C	N	O	S	0	2	0
			2275	1456	395	413	11			
1	D	283	Total	C	N	O	S	0	1	0
			2298	1472	399	415	12			
1	E	281	Total	C	N	O	S	0	6	0
			2332	1493	407	421	11			
1	F	279	Total	C	N	O	S	0	2	0
			2277	1457	397	412	11			

- Molecule 2 is 2'-DEOXYURIDINE 5'-MONOPHOSPHATE (three-letter code: UMP) (formula: C<sub>9</sub>H<sub>13</sub>N<sub>2</sub>O<sub>8</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
2	B	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
2	C	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
2	D	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
2	E	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
2	F	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
2	F	1	Total	C	N	O	P	0	0
			20	9	2	8	1		

- Molecule 3 is 1,4-BUTANEDIOL (three-letter code: BU1) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			6	4	2		
3	C	1	Total	C	O	0	0
			6	4	2		
3	C	1	Total	C	O	0	0
			6	4	2		

- Molecule 4 is water.

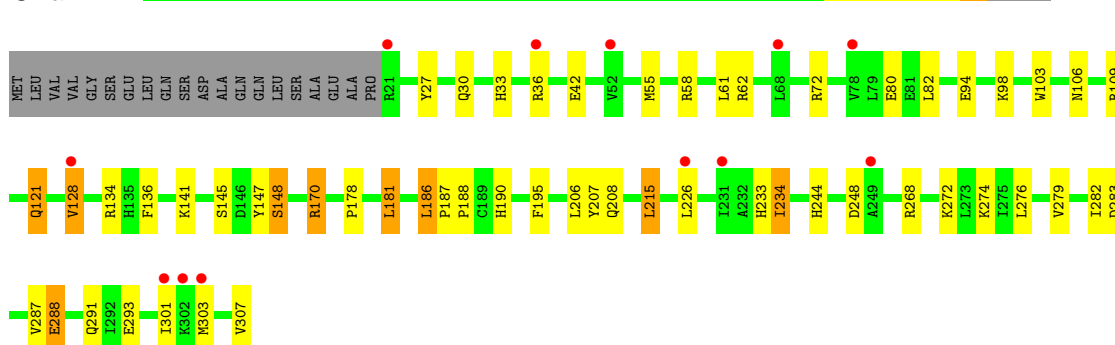
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	344	Total 344	O 344	0	0
4	B	315	Total 315	O 315	0	0
4	C	279	Total 279	O 279	0	0
4	D	242	Total 242	O 242	0	0
4	E	116	Total 116	O 116	0	0
4	F	176	Total 176	O 176	0	0

### 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.

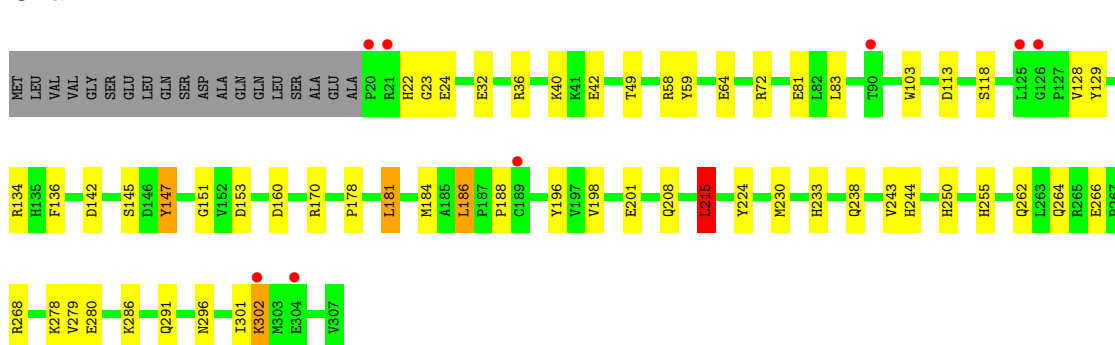
- Molecule 1: Thymidylate synthase

Chain A:



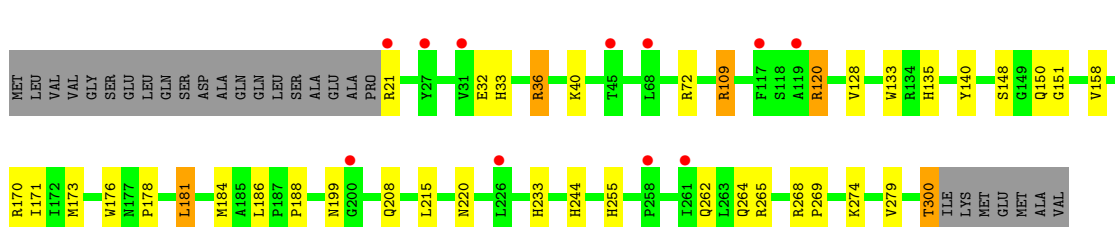
- Molecule 1: Thymidylate synthase

Chain B:



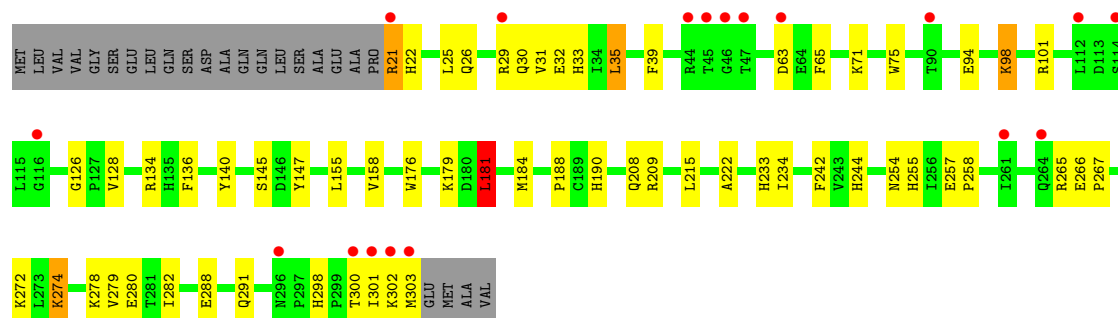
- Molecule 1: Thymidylate synthase

Chain C:



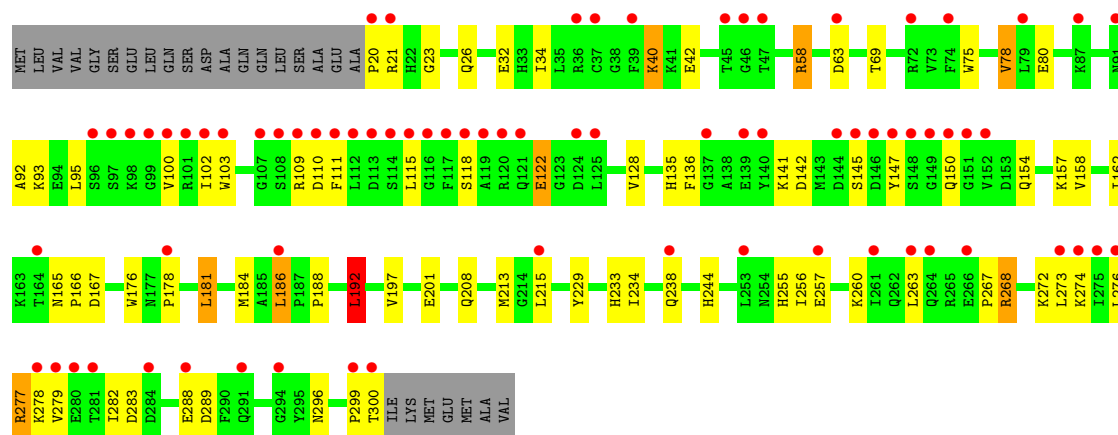
- Molecule 1: Thymidylate synthase

Chain D:



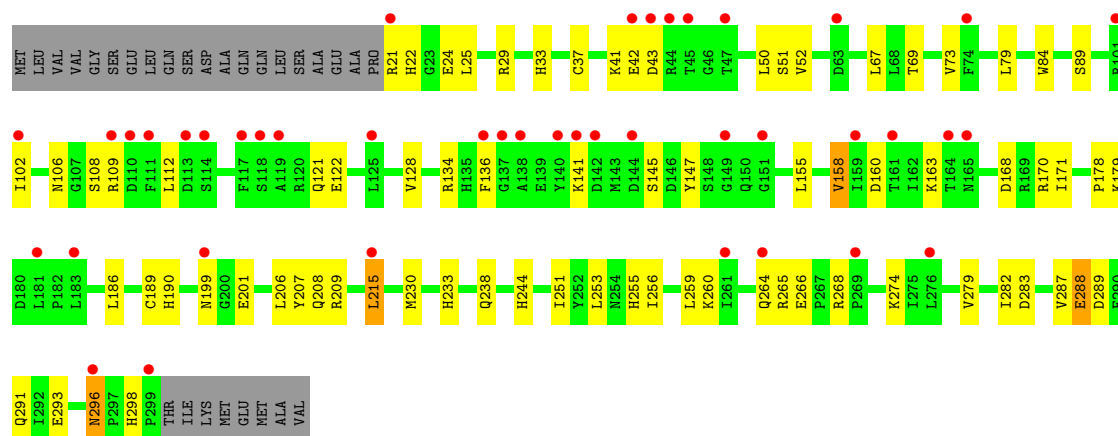
• Molecule 1: Thymidylate synthase

Chain E:



• Molecule 1: Thymidylate synthase

Chain F:



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	160.35Å 88.54Å 136.76Å 90.00° 95.99° 90.00°	Depositor
Resolution (Å)	19.95 – 1.70 19.95 – 1.70	Depositor EDS
% Data completeness (in resolution range)	98.6 (19.95-1.70) 98.6 (19.95-1.70)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.86 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.235 , 0.292 0.234 , 0.292	Depositor DCC
$R_{free}$ test set	4109 reflections (2.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.6	Xtriage
Anisotropy	0.054	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 48.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 205349 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15512	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	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 40.64 % 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 2.6537e-04.*

<sup>1</sup>Intensities 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: BU1, UMP

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	1.22	6/2405 (0.2%)	1.20	15/3248 (0.5%)
1	B	1.22	9/2414 (0.4%)	1.12	12/3261 (0.4%)
1	C	1.07	0/2337	1.01	3/3161 (0.1%)
1	D	0.99	1/2357 (0.0%)	0.94	2/3186 (0.1%)
1	E	0.78	0/2394	0.84	2/3237 (0.1%)
1	F	0.91	0/2336	0.92	4/3158 (0.1%)
All	All	1.05	16/14243 (0.1%)	1.01	38/19251 (0.2%)

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	D	0	1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	64	GLU	CB-CG	-6.85	1.39	1.52
1	A	207	TYR	CE1-CZ	6.16	1.46	1.38
1	B	129	TYR	CE2-CZ	6.09	1.46	1.38
1	A	27	TYR	CE1-CZ	6.01	1.46	1.38
1	D	242	PHE	CE1-CZ	5.99	1.48	1.37

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	134	ARG	NE-CZ-NH1	9.12	124.86	120.30
1	A	109	ARG	NE-CZ-NH2	-8.45	116.08	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	72	ARG	NE-CZ-NH1	8.14	124.37	120.30
1	B	170	ARG	NE-CZ-NH1	7.70	124.15	120.30
1	A	268	ARG	NE-CZ-NH2	-7.59	116.51	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	300	THR	Peptide

## 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	2346	0	2312	42	0
1	B	2354	0	2320	35	1
1	C	2275	0	2238	50	0
1	D	2298	0	2268	57	0
1	E	2332	0	2288	57	0
1	F	2277	0	2238	55	0
2	A	20	0	11	0	0
2	B	20	0	11	0	0
2	C	20	0	11	0	0
2	D	20	0	11	0	0
2	E	20	0	11	0	0
2	F	40	0	22	8	0
3	C	18	0	30	9	0
4	A	344	0	0	11	1
4	B	315	0	0	7	0
4	C	279	0	0	18	0
4	D	242	0	0	16	0
4	E	116	0	0	10	0
4	F	176	0	0	13	0
All	All	15512	0	13771	287	1

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 10.

The worst 5 of 287 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:F:502:UMP:C5	4:F:776:HOH:O	1.72	1.32
1:D:65:PHE:CE2	4:D:654:HOH:O	1.93	1.20
1:F:136:PHE:HE2	4:F:620:HOH:O	1.28	1.16
2:F:502:UMP:O4	4:F:774:HOH:O	1.64	1.16
1:D:188:PRO:HB2	4:D:642:HOH:O	1.47	1.11

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	Distance(Å)	Clash(Å)
1:B:36[B]:ARG:NH1	4:A:858:HOH:O[4.545]	2.00	0.20

## 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	288/307 (94%)	279 (97%)	7 (2%)	2 (1%)	30	10
1	B	289/307 (94%)	283 (98%)	5 (2%)	1 (0%)	50	27
1	C	280/307 (91%)	273 (98%)	6 (2%)	1 (0%)	43	22
1	D	282/307 (92%)	272 (96%)	9 (3%)	1 (0%)	43	22
1	E	285/307 (93%)	268 (94%)	15 (5%)	2 (1%)	30	10
1	F	279/307 (91%)	271 (97%)	7 (2%)	1 (0%)	43	22
All	All	1703/1842 (92%)	1646 (97%)	49 (3%)	8 (0%)	38	17

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	23	GLY
1	A	148	SER
1	E	128	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	128	VAL
1	A	128	VAL

### 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	253/266 (95%)	244 (96%)	9 (4%)	47	22
1	B	254/266 (96%)	247 (97%)	7 (3%)	56	32
1	C	246/266 (92%)	239 (97%)	7 (3%)	56	32
1	D	248/266 (93%)	237 (96%)	11 (4%)	39	15
1	E	251/266 (94%)	234 (93%)	17 (7%)	22	6
1	F	245/266 (92%)	233 (95%)	12 (5%)	35	13
All	All	1497/1596 (94%)	1434 (96%)	63 (4%)	40	16

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	181	LEU
1	E	78	VAL
1	F	215	LEU
1	D	274	LYS
1	D	280	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	233	HIS
1	D	205	GLN
1	F	233	HIS
1	C	244	HIS
1	C	264	GLN

### 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 ⓘ

10 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$
2	UMP	A	501	-	21,21,21	1.09	1 (4%)	26,31,31	1.35	2 (7%)
2	UMP	B	501	-	21,21,21	1.29	1 (4%)	26,31,31	1.06	2 (7%)
2	UMP	C	501	-	21,21,21	1.34	3 (14%)	26,31,31	1.92	3 (11%)
3	BU1	C	502	-	5,5,5	0.29	0	4,4,4	0.51	0
3	BU1	C	503	-	5,5,5	0.34	0	4,4,4	0.49	0
3	BU1	C	504	-	5,5,5	0.31	0	4,4,4	0.48	0
2	UMP	D	501	-	21,21,21	1.10	2 (9%)	26,31,31	1.24	2 (7%)
2	UMP	E	501	-	21,21,21	0.88	1 (4%)	26,31,31	1.29	2 (7%)
2	UMP	F	501	-	21,21,21	1.12	3 (14%)	26,31,31	1.30	2 (7%)
2	UMP	F	502	-	21,21,21	0.87	1 (4%)	26,31,31	1.88	4 (15%)

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
2	UMP	A	501	-	-	0/7/22/22	0/2/2/2
2	UMP	B	501	-	-	0/7/22/22	0/2/2/2
2	UMP	C	501	-	-	0/7/22/22	0/2/2/2
3	BU1	C	502	-	-	0/3/3/3	0/0/0/0
3	BU1	C	503	-	-	0/3/3/3	0/0/0/0
3	BU1	C	504	-	-	0/3/3/3	0/0/0/0
2	UMP	D	501	-	-	0/7/22/22	0/2/2/2
2	UMP	E	501	-	-	0/7/22/22	0/2/2/2
2	UMP	F	501	-	-	0/7/22/22	0/2/2/2
2	UMP	F	502	-	-	0/7/22/22	0/2/2/2

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	UMP	C6-C5	3.93	1.42	1.36
2	C	501	UMP	C6-C5	3.62	1.41	1.36
2	A	501	UMP	C6-C5	3.35	1.41	1.36
2	F	501	UMP	C6-C5	2.98	1.40	1.36
2	D	501	UMP	C6-C5	2.60	1.40	1.36

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	501	UMP	N3-C2-N1	7.63	122.34	115.97
2	F	502	UMP	N3-C2-N1	5.44	120.51	115.97
2	F	502	UMP	O4'-C1'-N1	5.32	117.68	107.68
2	A	501	UMP	N3-C2-N1	5.28	120.38	115.97
2	E	501	UMP	N3-C2-N1	4.73	119.92	115.97

There are no chirality outliers.

There are no torsion outliers.

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, 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	287/307 (93%)	0.46	12 (4%) 35 39	12, 20, 33, 43	0
1	B	288/307 (93%)	0.39	8 (2%) 50 56	13, 21, 31, 54	0
1	C	280/307 (91%)	0.47	11 (3%) 37 42	16, 24, 36, 43	0
1	D	283/307 (92%)	0.69	18 (6%) 19 23	18, 27, 44, 69	0
1	E	281/307 (91%)	1.69	76 (27%) 1 2	23, 39, 69, 86	0
1	F	279/307 (90%)	1.02	42 (15%) 3 4	20, 31, 50, 55	0
All	All	1698/1842 (92%)	0.78	167 (9%) 8 10	12, 26, 49, 86	0

The worst 5 of 167 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	117	PHE	10.3
1	E	115	LEU	10.3
1	E	112	LEU	10.0
1	E	119	ALA	9.4
1	E	116	GLY	8.7

### 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, 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	BU1	C	504	6/6	0.27	6.73	40,47,51,54	0
2	UMP	F	502	20/20	0.30	3.13	23,41,46,46	20
3	BU1	C	503	6/6	0.16	1.03	37,45,47,48	0
2	UMP	E	501	20/20	0.17	0.51	23,29,35,36	0
2	UMP	F	501	20/20	0.15	-0.05	24,31,32,34	0
2	UMP	D	501	20/20	0.10	-0.63	21,24,29,29	0
2	UMP	C	501	20/20	0.08	-0.95	19,22,29,30	0
2	UMP	B	501	20/20	0.09	-0.98	13,17,21,23	0
2	UMP	A	501	20/20	0.06	-2.01	13,16,24,24	0
3	BU1	C	502	6/6	0.27	-	56,57,57,58	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.