



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:20 PM GMT

PDB ID : 1F28  
Title : CRYSTAL STRUCTURE OF THYMIDYLATE SYNTHASE FROM PNEUMOCYSTIS CARINII BOUND TO DUMP AND BW1843U89  
Authors : Anderson, A.C.; O'Neil, R.H.; Surti, T.S.; Stroud, R.M.  
Deposited on : 2000-05-23  
Resolution : 1.90 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](mailto: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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

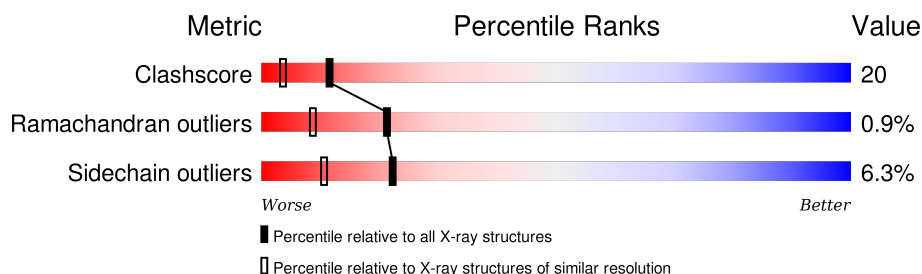
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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))
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	297	 67% 29% . .
1	B	297	 61% 35% . .
1	C	297	 66% 30% . .
1	D	297	 63% 32% 5% .

## 2 Entry composition [i](#)

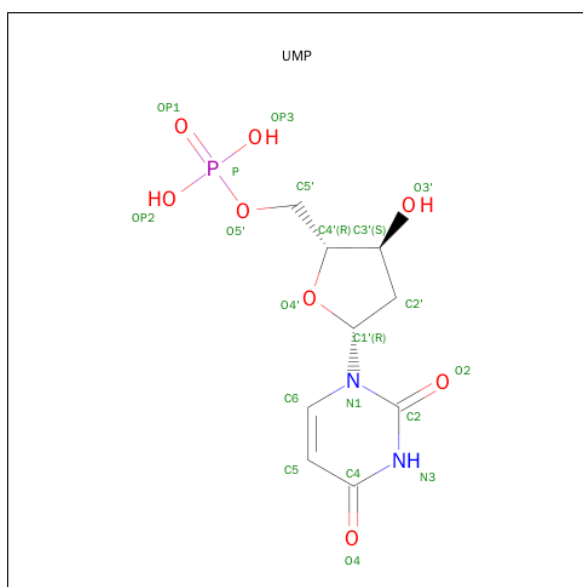
There are 4 unique types of molecules in this entry. The entry contains 10322 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 THYMIDYLATE SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	295	Total	C	N	O	S	0	0	0
			2394	1526	413	440	15			
1	B	295	Total	C	N	O	S	0	0	0
			2403	1530	417	441	15			
1	C	295	Total	C	N	O	S	0	0	0
			2394	1526	413	440	15			
1	D	295	Total	C	N	O	S	0	0	0
			2403	1530	417	441	15			

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



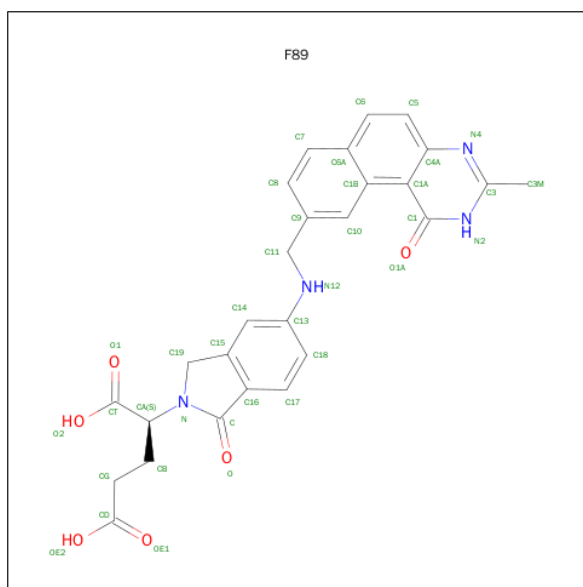
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
2	B	1	Total	C	N	O	P	0	0
			20	9	2	8	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
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		

- Molecule 3 is S)-2-(5(((1,2-DIHYDRO-3-METHYL-1-OXOBENZO(F)QUINAZOLIN-9-YL) METHYL)AMINO)1-OXO-2-ISOINDOLINYL)GLUTARIC ACID (three-letter code: F89) (formula: C<sub>27</sub>H<sub>24</sub>N<sub>4</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			37	27	4	6		
3	B	1	Total	C	N	O	0	0
			37	27	4	6		
3	C	1	Total	C	N	O	0	0
			37	27	4	6		
3	D	1	Total	C	N	O	0	0
			37	27	4	6		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	121	Total	O	0	0
			121	121		
4	B	125	Total	O	0	0
			125	125		

*Continued on next page...*

*Continued from previous page...*

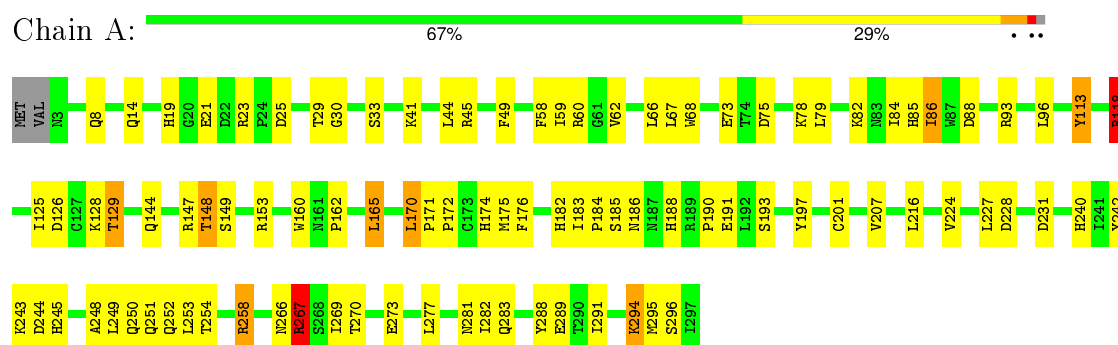
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	118	Total 118	O 118	0	0
4	D	136	Total 136	O 136	0	0

### 3 Residue-property plots

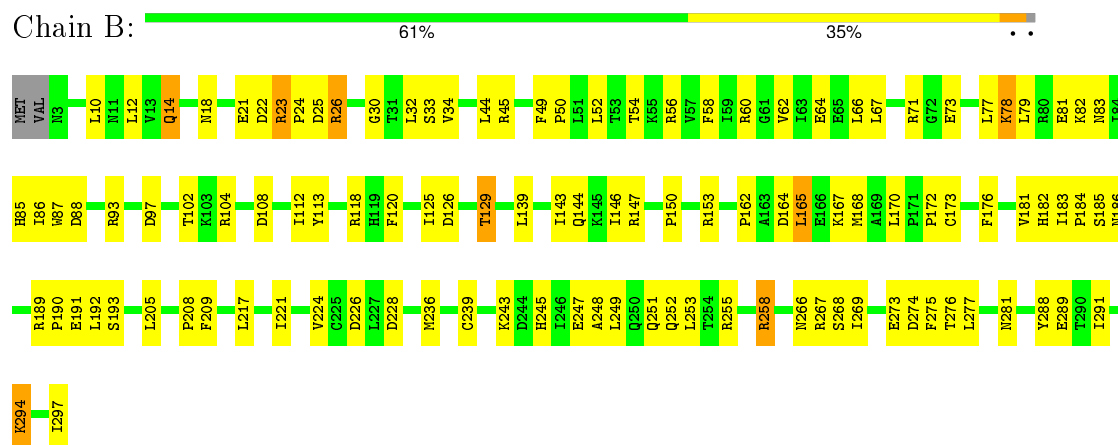
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

Note EDS was not executed.

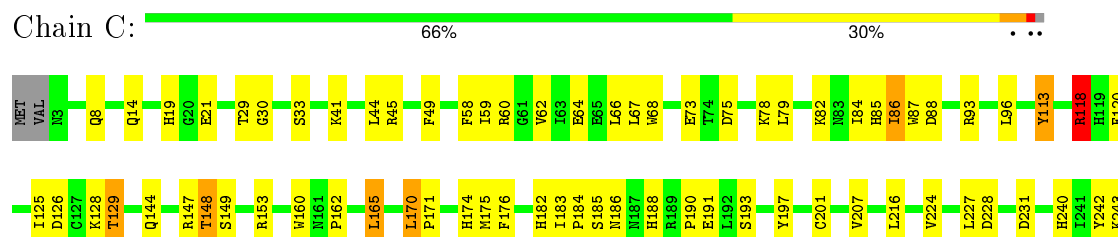
#### • Molecule 1: THYMIDYLATE SYNTHASE

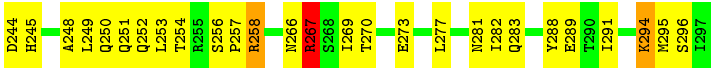


#### • Molecule 1: THYMIDYLATE SYNTHASE

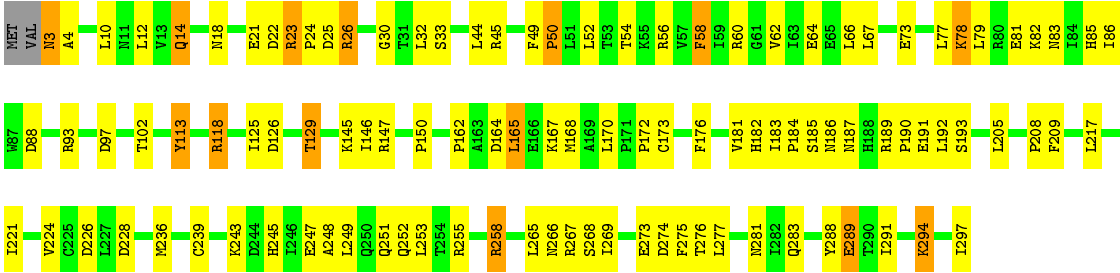


#### • Molecule 1: THYMIDYLATE SYNTHASE





● Molecule 1: THYMIDYLATE SYNTHASE



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.44Å 65.90Å 184.94Å 90.00° 90.24° 90.00°	Depositor
Resolution (Å)	22.40 – 1.90	Depositor
% Data completeness (in resolution range)	83.5 (22.40-1.90)	Depositor
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 0.9	Depositor
R, $R_{free}$	0.217 , 0.247	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10322	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP



## 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: F89, 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	0.52	0/2456	0.83	4/3331 (0.1%)
1	B	0.51	0/2465	0.76	0/3342
1	C	0.51	0/2456	0.83	4/3331 (0.1%)
1	D	0.51	0/2465	0.75	1/3342 (0.0%)
All	All	0.51	0/9842	0.79	9/13346 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	267	ARG	NE-CZ-NH2	-13.19	113.70	120.30
1	C	267	ARG	NE-CZ-NH1	12.65	126.63	120.30
1	A	267	ARG	NE-CZ-NH1	-12.50	114.05	120.30
1	A	267	ARG	NE-CZ-NH2	12.35	126.47	120.30
1	C	267	ARG	CD-NE-CZ	6.46	132.64	123.60

There are no chirality outliers.

There are no planarity outliers.

### 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	2394	0	2335	102	0
1	B	2403	0	2351	111	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2394	0	2335	105	0
1	D	2403	0	2351	112	0
2	A	20	0	10	0	0
2	B	20	0	11	2	0
2	C	20	0	10	0	0
2	D	20	0	11	2	0
3	A	37	0	22	2	0
3	B	37	0	22	5	0
3	C	37	0	22	4	0
3	D	37	0	22	5	0
4	A	121	0	0	2	0
4	B	125	0	0	9	0
4	C	118	0	0	0	0
4	D	136	0	0	8	0
All	All	10322	0	9502	394	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 20.

The worst 5 of 394 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:B:150:PRO:HB2	1:B:181:VAL:HG21	1.41	1.01
1:D:150:PRO:HB2	1:D:181:VAL:HG21	1.41	0.99
1:B:146:ILE:HD12	1:B:221:ILE:HD11	1.49	0.94
1:D:146:ILE:HD12	1:D:221:ILE:HD11	1.48	0.93
1:C:58:PHE:CE2	1:C:60:ARG:HB3	2.11	0.85

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	293/297 (99%)	277 (94%)	13 (4%)	3 (1%)	19	7
1	B	293/297 (99%)	277 (94%)	15 (5%)	1 (0%)	46	35
1	C	293/297 (99%)	278 (95%)	12 (4%)	3 (1%)	19	7
1	D	293/297 (99%)	275 (94%)	14 (5%)	4 (1%)	14	4
All	All	1172/1188 (99%)	1107 (94%)	54 (5%)	11 (1%)	21	9

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	113	TYR
1	A	267	ARG
1	C	113	TYR
1	C	267	ARG
1	A	244	ASP

### 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	264/268 (98%)	250 (95%)	14 (5%)	28	16
1	B	266/268 (99%)	247 (93%)	19 (7%)	18	8
1	C	264/268 (98%)	250 (95%)	14 (5%)	28	16
1	D	266/268 (99%)	246 (92%)	20 (8%)	17	7
All	All	1060/1072 (99%)	993 (94%)	67 (6%)	22	10

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	258	ARG
1	C	86	ILE
1	D	170	LEU
1	B	277	LEU
1	C	66	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	252	GLN
1	C	14	GLN
1	D	182	HIS
1	C	8	GLN
1	C	19	HIS

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

8 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	301	1	16,21,21	3.74	7 (43%)	23,31,31	3.04	7 (30%)
3	F89	A	401	-	35,41,41	3.37	24 (68%)	44,60,60	1.86	9 (20%)
2	UMP	B	302	-	16,21,21	1.98	5 (31%)	23,31,31	3.16	6 (26%)
3	F89	B	402	-	35,41,41	3.24	24 (68%)	44,60,60	1.92	9 (20%)
2	UMP	C	303	1	16,21,21	3.90	7 (43%)	23,31,31	2.95	7 (30%)
3	F89	C	403	-	35,41,41	3.38	23 (65%)	44,60,60	1.92	10 (22%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	UMP	D	304	-	16,21,21	1.86	5 (31%)	23,31,31	3.16	6 (26%)
3	F89	D	404	-	35,41,41	3.24	23 (65%)	44,60,60	2.01	10 (22%)

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	301	1	-	0/6/22/22	0/2/2/2
3	F89	A	401	-	-	0/12/30/30	0/5/5/5
2	UMP	B	302	-	-	0/6/22/22	0/2/2/2
3	F89	B	402	-	-	0/12/30/30	0/5/5/5
2	UMP	C	303	1	-	0/6/22/22	0/2/2/2
3	F89	C	403	-	-	0/12/30/30	0/5/5/5
2	UMP	D	304	-	-	0/6/22/22	0/2/2/2
3	F89	D	404	-	-	0/12/30/30	0/5/5/5

The worst 5 of 118 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	404	F89	C19-C15	-3.26	1.45	1.50
3	B	402	F89	C19-C15	-3.20	1.45	1.50
2	A	301	UMP	P-OP3	-2.59	1.45	1.54
2	C	303	UMP	P-OP3	-2.44	1.45	1.54
2	D	304	UMP	P-OP3	-2.37	1.46	1.54

The worst 5 of 64 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	403	F89	C1A-C1-N2	-5.61	119.94	124.19
3	D	404	F89	C1A-C1-N2	-5.45	120.07	124.19
3	A	401	F89	C1A-C1-N2	-5.43	120.08	124.19
3	B	402	F89	C1A-C1-N2	-5.36	120.14	124.19
2	A	301	UMP	C5-C6-N1	-4.15	110.40	120.58

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	401	F89	2	0
2	B	302	UMP	2	0
3	B	402	F89	5	0
3	C	403	F89	4	0
2	D	304	UMP	2	0
3	D	404	F89	5	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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.