



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 06:34 AM GMT

PDB ID : 1O29  
Title : Crystal structure of Thymidylate Synthase Complementing Protein (TM0449) from *Thermotoga maritima* with FAD and FdUMP at 2.0 Å resolution  
Authors : Mathews, I.I.; Deacon, A.M.; Canaves, J.M.; McMullan, D.; Lesley, S.A.; Agarwalla, S.; Kuhn, P.; Joint Center for Structural Genomics (JCSG)  
Deposited on : 2003-02-18  
Resolution : 2.00 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>

---

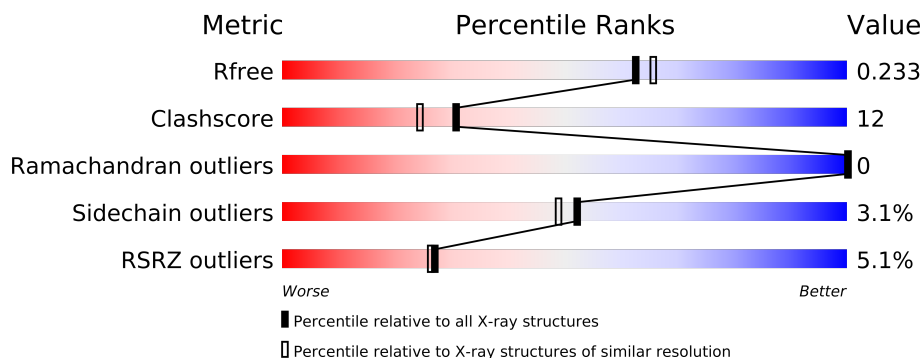
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.00 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	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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	232	
1	B	232	
1	C	232	
1	D	232	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7990 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 thyX.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	216	Total	C	N	O	S	0	0	0
			1809	1177	310	316	6			
1	B	220	Total	C	N	O	S	0	0	0
			1841	1197	314	324	6			
1	C	216	Total	C	N	O	S	0	0	0
			1813	1180	310	318	5			
1	D	215	Total	C	N	O	S	0	0	0
			1803	1173	307	317	6			

There are 48 discrepancies between the modelled and reference sequences:

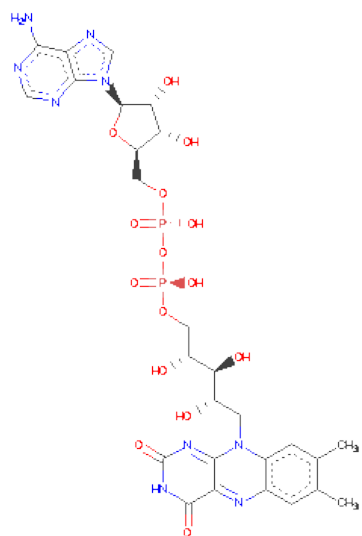
Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	LEADER SEQUENCE	UNP Q9WYT0
A	-10	GLY	-	LEADER SEQUENCE	UNP Q9WYT0
A	-9	SER	-	LEADER SEQUENCE	UNP Q9WYT0
A	-8	ASP	-	LEADER SEQUENCE	UNP Q9WYT0
A	-7	LYS	-	LEADER SEQUENCE	UNP Q9WYT0
A	-6	ILE	-	LEADER SEQUENCE	UNP Q9WYT0
A	-5	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
A	-4	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
A	-3	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
A	-2	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
A	-1	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
A	0	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
B	-11	MET	-	LEADER SEQUENCE	UNP Q9WYT0
B	-10	GLY	-	LEADER SEQUENCE	UNP Q9WYT0
B	-9	SER	-	LEADER SEQUENCE	UNP Q9WYT0
B	-8	ASP	-	LEADER SEQUENCE	UNP Q9WYT0
B	-7	LYS	-	LEADER SEQUENCE	UNP Q9WYT0
B	-6	ILE	-	LEADER SEQUENCE	UNP Q9WYT0
B	-5	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
B	-4	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
B	-3	HIS	-	LEADER SEQUENCE	UNP Q9WYT0

*Continued on next page...*

*Continued from previous page...*

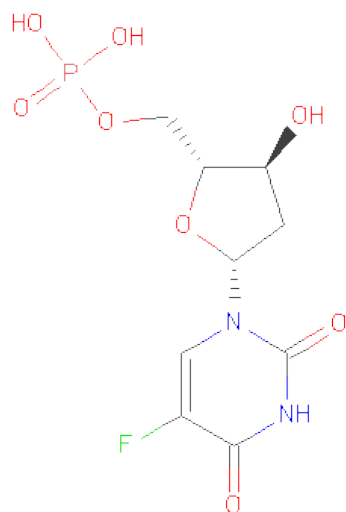
Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
B	-1	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
B	0	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
C	-11	MET	-	LEADER SEQUENCE	UNP Q9WYT0
C	-10	GLY	-	LEADER SEQUENCE	UNP Q9WYT0
C	-9	SER	-	LEADER SEQUENCE	UNP Q9WYT0
C	-8	ASP	-	LEADER SEQUENCE	UNP Q9WYT0
C	-7	LYS	-	LEADER SEQUENCE	UNP Q9WYT0
C	-6	ILE	-	LEADER SEQUENCE	UNP Q9WYT0
C	-5	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
C	-4	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
C	-3	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
C	-2	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
C	-1	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
C	0	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
D	-11	MET	-	LEADER SEQUENCE	UNP Q9WYT0
D	-10	GLY	-	LEADER SEQUENCE	UNP Q9WYT0
D	-9	SER	-	LEADER SEQUENCE	UNP Q9WYT0
D	-8	ASP	-	LEADER SEQUENCE	UNP Q9WYT0
D	-7	LYS	-	LEADER SEQUENCE	UNP Q9WYT0
D	-6	ILE	-	LEADER SEQUENCE	UNP Q9WYT0
D	-5	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
D	-4	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
D	-3	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
D	-2	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
D	-1	HIS	-	LEADER SEQUENCE	UNP Q9WYT0
D	0	HIS	-	LEADER SEQUENCE	UNP Q9WYT0

- Molecule 2 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula:  $C_{27}H_{33}N_9O_{15}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	D	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
2	C	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
2	B	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
2	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 3 is 5-FLUORO-2'-DEOXYURIDINE-5'-MONOPHOSPHATE (three-letter code: UFP) (formula: C<sub>9</sub>H<sub>12</sub>FN<sub>2</sub>O<sub>8</sub>P).

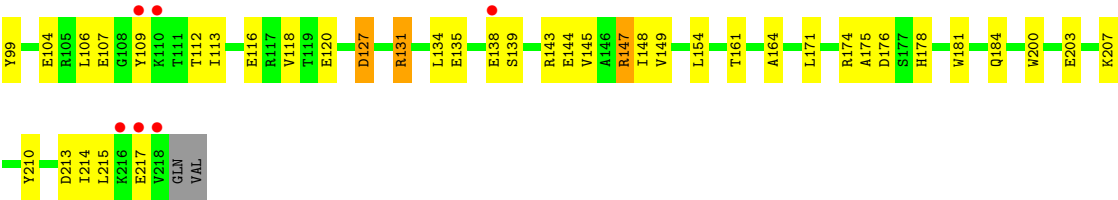


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total 21	C 9	F 1	N 2	O 8	P 1	0	0
3	B	1	Total 21	C 9	F 1	N 2	O 8	P 1	0	0
3	C	1	Total 21	C 9	F 1	N 2	O 8	P 1	0	0
3	D	1	Total 21	C 9	F 1	N 2	O 8	P 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	111	Total 111	O 111	0	0
4	B	97	Total 97	O 97	0	0
4	C	116	Total 116	O 116	0	0
4	D	104	Total 104	O 104	0	0







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.64Å 116.95Å 141.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00 45.13 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.0 (20.00-2.00) 99.2 (45.13-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.04 (at 1.89Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.194 , 0.232 0.198 , 0.233	Depositor DCC
$R_{free}$ test set	3649 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.6	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 42.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 71914 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7990	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% of the height of the origin peak. No significant pseudotranslation is detected.*

<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: UFP, FAD

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	0.34	0/1859	0.57	0/2514
1	B	0.33	0/1890	0.57	0/2554
1	C	0.34	0/1861	0.57	0/2514
1	D	0.33	0/1851	0.56	0/2502
All	All	0.34	0/7461	0.57	0/10084

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	1809	0	1801	52	0
1	B	1841	0	1839	47	0
1	C	1813	0	1811	39	0
1	D	1803	0	1794	51	0
2	A	53	0	30	1	0
2	B	53	0	30	1	0
2	C	53	0	30	2	0
2	D	53	0	30	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	21	0	10	1	0
3	B	21	0	10	2	0
3	C	21	0	10	2	0
3	D	21	0	10	1	0
4	A	111	0	0	1	0
4	B	97	0	0	4	0
4	C	116	0	0	2	0
4	D	104	0	0	3	0
All	All	7990	0	7405	177	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 12.

The worst 5 of 177 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:174:ARG:HH12	3:D:818:UFP:HN3	1.09	0.99
1:A:17:MET:HB2	1:B:17:MET:HB2	1.46	0.97
3:A:803:UFP:HN3	1:D:174:ARG:HH12	1.11	0.94
1:B:174:ARG:HH12	3:C:813:UFP:HN3	1.09	0.90
3:B:808:UFP:HN3	1:C:174:ARG:HH12	1.12	0.89

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	214/232 (92%)	209 (98%)	5 (2%)	0	100	100
1	B	218/232 (94%)	212 (97%)	6 (3%)	0	100	100
1	C	212/232 (91%)	204 (96%)	8 (4%)	0	100	100
1	D	211/232 (91%)	205 (97%)	6 (3%)	0	100	100
All	All	855/928 (92%)	830 (97%)	25 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	192/207 (93%)	185 (96%)	7 (4%)	47	42
1	B	196/207 (95%)	190 (97%)	6 (3%)	52	49
1	C	193/207 (93%)	186 (96%)	7 (4%)	47	42
1	D	192/207 (93%)	188 (98%)	4 (2%)	66	67
All	All	773/828 (93%)	749 (97%)	24 (3%)	52	49

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	147	ARG
1	C	38	GLU
1	D	131	ARG
1	B	192	ARG
1	B	196	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	185	GLN
1	D	178	HIS
1	C	185	GLN
1	A	178	HIS
1	C	219	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 ⓘ

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	$\# Z  > 2$
3	UFP	A	803	-	22,22,22	3.44	10 (45%)	28,33,33	3.70	7 (25%)
2	FAD	A	815	-	58,58,58	2.83	27 (46%)	85,89,89	2.73	23 (27%)
3	UFP	B	808	-	22,22,22	3.34	10 (45%)	28,33,33	3.69	7 (25%)
2	FAD	B	810	-	58,58,58	2.81	26 (44%)	85,89,89	2.77	23 (27%)
2	FAD	C	805	-	58,58,58	2.73	25 (43%)	85,89,89	2.74	23 (27%)
3	UFP	C	813	-	22,22,22	3.37	11 (50%)	28,33,33	3.64	8 (28%)
2	FAD	D	800	-	58,58,58	2.80	28 (48%)	85,89,89	2.72	23 (27%)
3	UFP	D	818	-	22,22,22	3.37	10 (45%)	28,33,33	3.47	8 (28%)

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	UFP	A	803	-	-	0/7/22/22	0/2/2/2
2	FAD	A	815	-	1/1/9/9	0/34/50/50	0/1/6/6
3	UFP	B	808	-	-	0/7/22/22	0/2/2/2
2	FAD	B	810	-	1/1/9/9	0/34/50/50	0/1/6/6
2	FAD	C	805	-	1/1/9/9	0/34/50/50	0/1/6/6

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	UFP	C	813	-	-	0/7/22/22	0/2/2/2
2	FAD	D	800	-	1/1/9/9	0/34/50/50	0/1/6/6
3	UFP	D	818	-	-	0/7/22/22	0/2/2/2

The worst 5 of 147 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	803	UFP	C2-N1	10.80	1.50	1.38
3	D	818	UFP	C2-N1	10.20	1.49	1.38
3	C	813	UFP	C2-N1	10.12	1.49	1.38
3	B	808	UFP	C2-N1	10.07	1.49	1.38
2	B	810	FAD	C4A-N3A	6.60	1.45	1.35

The worst 5 of 122 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	808	UFP	C6-N1-C2	-16.69	117.67	122.41
3	A	803	UFP	C6-N1-C2	-16.65	117.68	122.41
3	C	813	UFP	C6-N1-C2	-16.16	117.82	122.41
3	D	818	UFP	C6-N1-C2	-15.22	118.08	122.41
2	C	805	FAD	C4B-O4B-C1B	-9.58	99.34	109.75

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	D	800	FAD	C1B
2	C	805	FAD	C1B
2	B	810	FAD	C1B
2	A	815	FAD	C1B

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	216/232 (93%)	0.09	8 (3%) 39 39	15, 27, 58, 66	0
1	B	220/232 (94%)	0.25	16 (7%) 15 14	15, 30, 64, 67	0
1	C	216/232 (93%)	0.00	8 (3%) 39 39	16, 26, 55, 65	0
1	D	215/232 (92%)	0.16	13 (6%) 21 20	15, 28, 58, 66	0
All	All	867/928 (93%)	0.13	45 (5%) 27 26	15, 28, 59, 67	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	218	VAL	6.3
1	B	40	ARG	4.2
1	C	220	VAL	4.2
1	B	38	GLU	3.8
1	B	107	GLU	3.7

### 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( $\text{\AA}^2$ )	Q<0.9
3	UFP	B	808	21/21	0.12	0.21	18,20,22,24	0
3	UFP	C	813	21/21	0.12	-0.10	20,23,28,29	0
2	FAD	C	805	53/53	0.12	-0.17	16,21,25,26	0
2	FAD	D	800	53/53	0.13	-0.29	18,23,28,29	0
2	FAD	B	810	53/53	0.14	-0.32	17,23,27,30	0
2	FAD	A	815	53/53	0.12	-0.61	17,21,24,25	0
3	UFP	D	818	21/21	0.10	-0.71	19,23,25,27	0
3	UFP	A	803	21/21	0.10	-0.76	21,23,25,29	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.