W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 16, 2023 – 05:02 PM EDT

PDB ID	:	2A1F
Title	:	Crystal Structure of Uridylate kinase
Authors	:	Gorman, J.; Shapiro, L.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2005-06-20
Resolution	:	2.10 Å(reported)

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

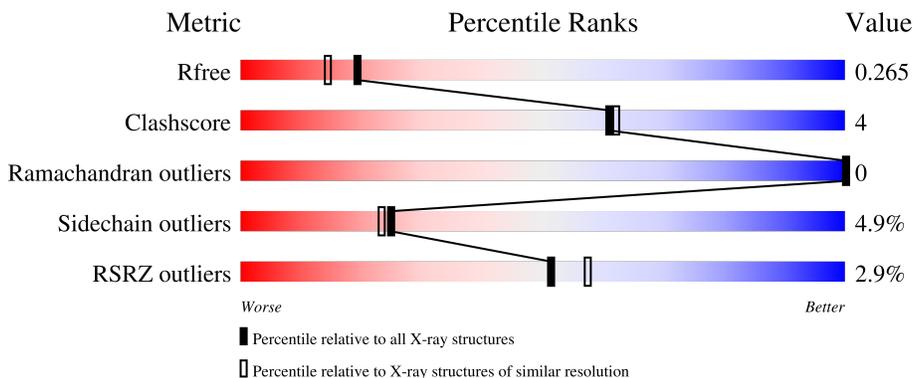|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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	Percentile Ranks	Value
Rfree		0.265
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		4.9%
RSRZ outliers		2.9%

Worse　　　　　　　　　　　　　　　　　　*Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=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	247	3%　82%　11%　• 6%
1	B	247	3%　85%　8%　6%
1	C	247	2%　84%　11%　• •
1	D	247	3%　83%　12%　•
1	E	247	3%　84%　11%　• 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	247	

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 11567 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 Uridylate kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	232	Total	C	N	O	S	0	0	0
			1755	1108	301	330	16			
1	B	232	Total	C	N	O	S	0	0	0
			1755	1108	301	330	16			
1	C	237	Total	C	N	O	S	0	0	0
			1807	1138	318	335	16			
1	D	236	Total	C	N	O	S	0	0	0
			1797	1132	315	334	16			
1	E	235	Total	C	N	O	S	0	0	0
			1787	1126	312	333	16			
1	F	234	Total	C	N	O	S	0	0	0
			1777	1120	309	332	16			

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	cloning artifact	UNP P43890
A	0	SER	-	cloning artifact	UNP P43890
A	1	LEU	-	cloning artifact	UNP P43890
A	238	GLU	-	cloning artifact	UNP P43890
A	239	GLY	-	cloning artifact	UNP P43890
A	240	HIS	-	expression tag	UNP P43890
A	241	HIS	-	expression tag	UNP P43890
A	242	HIS	-	expression tag	UNP P43890
A	243	HIS	-	expression tag	UNP P43890
A	244	HIS	-	expression tag	UNP P43890
A	245	HIS	-	expression tag	UNP P43890
B	-1	MET	-	cloning artifact	UNP P43890
B	0	SER	-	cloning artifact	UNP P43890
B	1	LEU	-	cloning artifact	UNP P43890
B	238	GLU	-	cloning artifact	UNP P43890
B	239	GLY	-	cloning artifact	UNP P43890
B	240	HIS	-	expression tag	UNP P43890

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	241	HIS	-	expression tag	UNP P43890
B	242	HIS	-	expression tag	UNP P43890
B	243	HIS	-	expression tag	UNP P43890
B	244	HIS	-	expression tag	UNP P43890
B	245	HIS	-	expression tag	UNP P43890
C	-1	MET	-	cloning artifact	UNP P43890
C	0	SER	-	cloning artifact	UNP P43890
C	1	LEU	-	cloning artifact	UNP P43890
C	238	GLU	-	cloning artifact	UNP P43890
C	239	GLY	-	cloning artifact	UNP P43890
C	240	HIS	-	expression tag	UNP P43890
C	241	HIS	-	expression tag	UNP P43890
C	242	HIS	-	expression tag	UNP P43890
C	243	HIS	-	expression tag	UNP P43890
C	244	HIS	-	expression tag	UNP P43890
C	245	HIS	-	expression tag	UNP P43890
D	-1	MET	-	cloning artifact	UNP P43890
D	0	SER	-	cloning artifact	UNP P43890
D	1	LEU	-	cloning artifact	UNP P43890
D	238	GLU	-	cloning artifact	UNP P43890
D	239	GLY	-	cloning artifact	UNP P43890
D	240	HIS	-	expression tag	UNP P43890
D	241	HIS	-	expression tag	UNP P43890
D	242	HIS	-	expression tag	UNP P43890
D	243	HIS	-	expression tag	UNP P43890
D	244	HIS	-	expression tag	UNP P43890
D	245	HIS	-	expression tag	UNP P43890
E	-1	MET	-	cloning artifact	UNP P43890
E	0	SER	-	cloning artifact	UNP P43890
E	1	LEU	-	cloning artifact	UNP P43890
E	238	GLU	-	cloning artifact	UNP P43890
E	239	GLY	-	cloning artifact	UNP P43890
E	240	HIS	-	expression tag	UNP P43890
E	241	HIS	-	expression tag	UNP P43890
E	242	HIS	-	expression tag	UNP P43890
E	243	HIS	-	expression tag	UNP P43890
E	244	HIS	-	expression tag	UNP P43890
E	245	HIS	-	expression tag	UNP P43890
F	-1	MET	-	cloning artifact	UNP P43890
F	0	SER	-	cloning artifact	UNP P43890
F	1	LEU	-	cloning artifact	UNP P43890
F	238	GLU	-	cloning artifact	UNP P43890

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	239	GLY	-	cloning artifact	UNP P43890
F	240	HIS	-	expression tag	UNP P43890
F	241	HIS	-	expression tag	UNP P43890
F	242	HIS	-	expression tag	UNP P43890
F	243	HIS	-	expression tag	UNP P43890
F	244	HIS	-	expression tag	UNP P43890
F	245	HIS	-	expression tag	UNP P43890

- Molecule 2 is water.

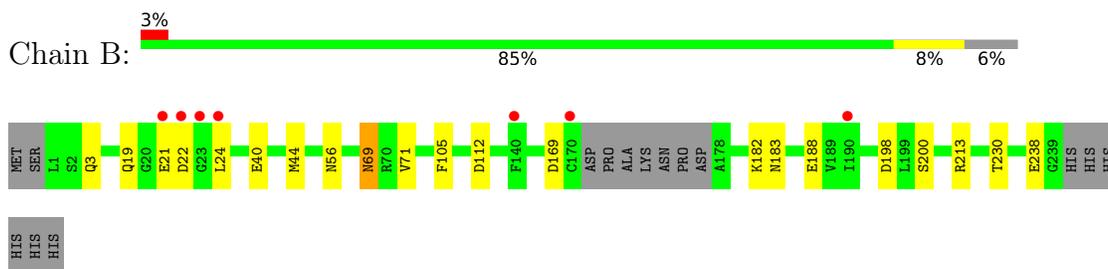| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	154	Total 154	O 154	0	0
2	B	145	Total 145	O 145	0	0
2	C	129	Total 129	O 129	0	0
2	D	167	Total 167	O 167	0	0
2	E	156	Total 156	O 156	0	0
2	F	138	Total 138	O 138	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.
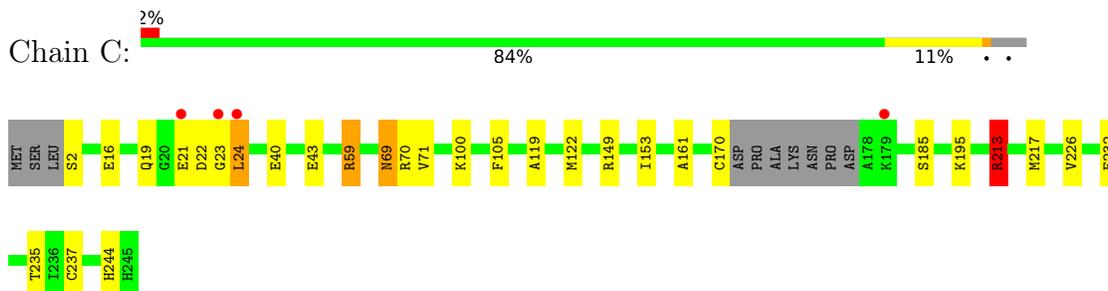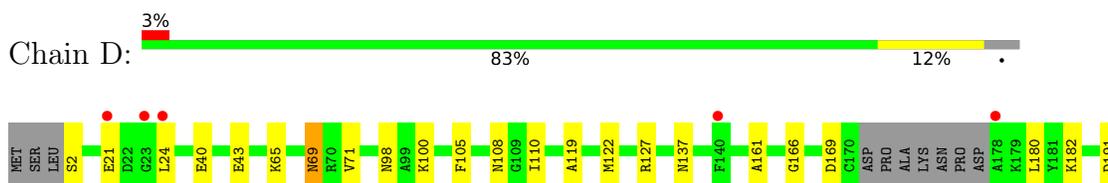
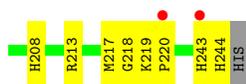• Molecule 1: Uridylate kinase

• Molecule 1: Uridylate kinase

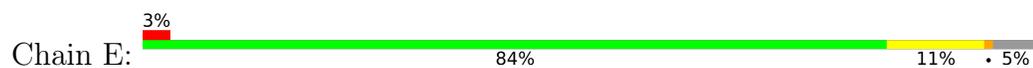• Molecule 1: Uridylate kinase

• Molecule 1: Uridylate kinase

H208 R213 M217 G218 K219 P220 H243 H244 HIS

• Molecule 1: Uridylate kinase

Chain E:

3%

84%   11%   5%

MET SER LEU S2 E21 D22 G23 E40 M44 R59 N69 R70 V71 N98 F105 S117 E118 A119 I120 K121 M122 R127 R149 A161 C170 ASP ALA LYS ASN PRO ASP A178 K182 N183 L184 S185 I190 K195 S200 R206 R213

M217 V226 G229 C237 H243 HIS HIS

• Molecule 1: Uridylate kinase

Chain F:

3%

79%   14%   5%

MET SER LEU S2 Q3 P4 E16 A17 L18 Q19 G20 E21 D22 G23 L24 R33 E47 N56 R59 N69 R70 V71 K100 F105 N108 E118 A119 I120 K121 M122 L123 R124 R127 N137 F140 D143 R149 I153 K160 A161 C170 ASP

PRO ALA LYS ASN PRO ASP A178 K179 K182 N183 L184 S185 V196 M197 R213 M217 T230 E231 E232 C237 H242 HIS HIS HIS

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	77.37Å 79.89Å 79.90Å 94.85° 96.68° 96.88°	Depositor
Resolution (Å)	20.00 – 2.10 19.85 – 2.10	Depositor EDS
% Data completeness (in resolution range)	96.5 (20.00-2.10) 95.7 (19.85-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.79 (at 2.09Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.210 , 0.263 0.216 , 0.265	Depositor DCC
R_{free} test set	5265 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	24.6	Xtriage
Anisotropy	0.224	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 41.4	EDS
L-test for twinning[2]	$< |L| >$ = 0.46, $< L^2 >$ = 0.29	Xtriage
Estimated twinning fraction	0.038 for k,l,h 0.038 for l,h,k 0.037 for -k,-h,-l 0.038 for -l,-k,-h 0.438 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	11567	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.33% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	0.57	0/1775	0.73	2/2386 (0.1%)
1	B	0.58	0/1775	0.73	2/2386 (0.1%)
1	C	0.57	0/1833	0.68	2/2465 (0.1%)
1	D	0.58	0/1822	0.72	3/2450 (0.1%)
1	E	0.58	0/1811	0.76	4/2435 (0.2%)
1	F	0.58	0/1800	0.74	3/2420 (0.1%)
All	All	0.58	0/10816	0.73	16/14542 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	213	ARG	NE-CZ-NH2	-11.46	114.57	120.30
1	F	213	ARG	NE-CZ-NH2	-11.26	114.67	120.30
1	A	213	ARG	NE-CZ-NH2	-10.11	115.25	120.30
1	C	213	ARG	NE-CZ-NH2	-9.68	115.46	120.30
1	D	213	ARG	NE-CZ-NH2	-9.37	115.62	120.30

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	1755	0	1806	16	0
1	B	1755	0	1806	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1807	0	1834	17	0
1	D	1797	0	1827	19	0
1	E	1787	0	1820	15	0
1	F	1777	0	1813	23	0
2	A	154	0	0	4	0
2	B	145	0	0	1	0
2	C	129	0	0	2	0
2	D	167	0	0	6	0
2	E	156	0	0	3	0
2	F	138	0	0	5	0
All	All	11567	0	10906	97	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 4.

The worst 5 of 97 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:C:19:GLN:HE21	1:C:24:LEU:HA	1.41	0.86
1:E:182:LYS:HD3	2:E:377:HOH:O	1.90	0.71
1:A:108:ASN:ND2	2:A:383:HOH:O	2.27	0.67
1:F:108:ASN:ND2	2:F:367:HOH:O	2.29	0.65
1:D:182:LYS:HD3	2:D:412:HOH:O	1.97	0.65

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	228/247 (92%)	226 (99%)	2 (1%)	0	100	100
1	B	228/247 (92%)	223 (98%)	5 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	233/247 (94%)	229 (98%)	4 (2%)	0	100	100
1	D	232/247 (94%)	226 (97%)	6 (3%)	0	100	100
1	E	231/247 (94%)	227 (98%)	4 (2%)	0	100	100
1	F	230/247 (93%)	227 (99%)	3 (1%)	0	100	100
All	All	1382/1482 (93%)	1358 (98%)	24 (2%)	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	187/201 (93%)	174 (93%)	13 (7%)	15	12
1	B	187/201 (93%)	177 (95%)	10 (5%)	22	20
1	C	192/201 (96%)	182 (95%)	10 (5%)	23	21
1	D	191/201 (95%)	184 (96%)	7 (4%)	34	35
1	E	190/201 (94%)	185 (97%)	5 (3%)	46	50
1	F	189/201 (94%)	178 (94%)	11 (6%)	20	17
All	All	1136/1206 (94%)	1080 (95%)	56 (5%)	25	23

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	170	CYS
1	F	232	GLU
1	D	65	LYS
1	F	197	MET
1	F	69	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	108	ASN
1	F	69	ASN
1	D	208	HIS
1	F	137	ASN
1	E	208	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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	232/247 (93%)	0.07	7 (3%)	50	56	13, 25, 43, 51	1 (0%)
1	B	232/247 (93%)	0.04	7 (3%)	50	56	13, 25, 43, 48	1 (0%)
1	C	237/247 (95%)	-0.03	4 (1%)	70	74	12, 26, 45, 53	1 (0%)
1	D	236/247 (95%)	-0.01	7 (2%)	50	56	13, 25, 42, 51	1 (0%)
1	E	235/247 (95%)	0.10	8 (3%)	45	51	12, 25, 42, 51	1 (0%)
1	F	234/247 (94%)	0.10	8 (3%)	45	51	13, 26, 42, 53	1 (0%)
All	All	1406/1482 (94%)	0.04	41 (2%)	51	57	12, 25, 43, 53	6 (0%)

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	23	GLY	5.3
1	D	23	GLY	4.8
1	E	23	GLY	4.6
1	A	170	CYS	4.4
1	E	178	ALA	4.3

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 monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.