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 (i)

May 25, 2020 – 12:45 pm BST

PDB ID	:	1FUU
Title	:	YEAST INITIATION FACTOR 4A
Authors	:	Caruthers, J.M.; Johnson, E.R.; McKay, D.B.
Deposited on	:	2000-09-15
Resolution	:	2.50 Å(reported)

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

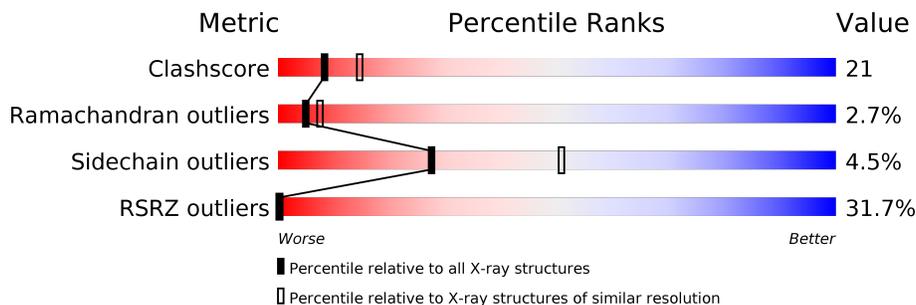|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.50 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
Clashscore		21
Ramachandran outliers		2.7%
Sidechain outliers		4.5%
RSRZ outliers		31.7%

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))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 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	394	5% 36% 16% ·· 45%
1	B	394	41% 56% 38% · ·

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4800 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 YEAST INITIATION FACTOR 4A.

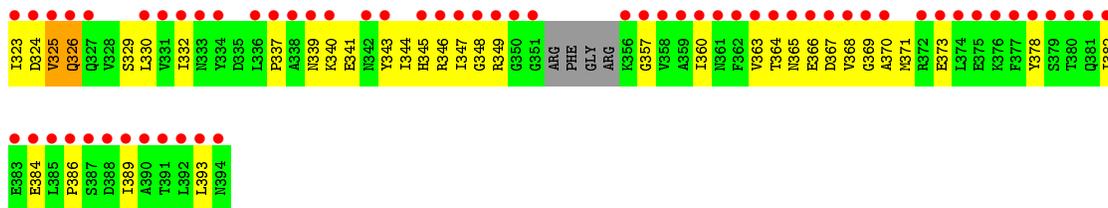| Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	215	Total 1660	C 1060	N 277	O 313	S 1	Se 9	0	0	0
1	B	380	Total 2994	C 1898	N 500	O 582	S 3	Se 11	64	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	MSE	MET	modified residue	UNP P10081
A	53	MSE	MET	modified residue	UNP P10081
A	94	MSE	MET	modified residue	UNP P10081
A	110	MSE	MET	modified residue	UNP P10081
A	116	MSE	MET	modified residue	UNP P10081
A	165	MSE	MET	modified residue	UNP P10081
A	174	MSE	MET	modified residue	UNP P10081
A	203	MSE	MET	modified residue	UNP P10081
A	215	MSE	MET	modified residue	UNP P10081
A	302	MSE	MET	modified residue	UNP P10081
A	371	MSE	MET	modified residue	UNP P10081
B	26	MSE	MET	modified residue	UNP P10081
B	53	MSE	MET	modified residue	UNP P10081
B	94	MSE	MET	modified residue	UNP P10081
B	110	MSE	MET	modified residue	UNP P10081
B	116	MSE	MET	modified residue	UNP P10081
B	165	MSE	MET	modified residue	UNP P10081
B	174	MSE	MET	modified residue	UNP P10081
B	203	MSE	MET	modified residue	UNP P10081
B	215	MSE	MET	modified residue	UNP P10081
B	302	MSE	MET	modified residue	UNP P10081
B	371	MSE	MET	modified residue	UNP P10081

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	76	Total 76	O 76	0	0
2	B	70	Total 70	O 70	0	0

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 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: YEAST INITIATION FACTOR 4A

- Molecule 1: YEAST INITIATION FACTOR 4A

I323 D324 V325 Q326 Q327 V328 S329 L330 V331 I332 N333 Y334 D335 L336 P337 A338 N339 K340 E341 N342 Y343 I344 H345 R346 I347 G348 R349 G350 G351 ARG PHE GLY ARG K356 G357 V358 A359 I360 N361 F362 V363 T364 N365 E366 D367 V368 G369 A370 M371 R372 E373 L374 E375 K376 F377 Y378 S379 T380 Q381 I382

E383 E384 L385 P386 S387 D388 I389 A390 T391 L392 L393 N394

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	38.80Å 71.20Å 73.20Å 94.00° 89.60° 101.00°	Depositor
Resolution (Å)	40.00 – 2.50 27.09 – 2.51	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-2.50) 84.9 (27.09-2.51)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	7.97 (at 2.50Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.243 , 0.273 0.251 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	27.7	Xtriage
Anisotropy	0.147	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 62.5	EDS
L-test for twinning[2]	$<|L|> = 0.48, <L^2> = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	4800	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.81% 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 ⓘ

5.1 Standard geometry ⓘ

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	0.41	0/1678	0.66	1/2258 (0.0%)
1	B	0.35	0/3027	0.57	0/4074
All	All	0.37	0/4705	0.61	1/6332 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	18	LYS	N-CA-C	-5.51	96.12	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	1660	0	1669	64	0
1	B	2994	0	3000	134	0
2	A	76	0	0	2	0
2	B	70	0	0	2	0
All	All	4800	0	4669	193	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 21.

The worst 5 of 193 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:211:THR:HA	1:B:215:MSE:HE3	1.46	0.98
1:A:42:GLU:HG3	1:B:127:THR:H	1.36	0.91
1:B:272:VAL:HG13	1:B:314:ILE:HG22	1.58	0.83
1:B:261:GLN:HG2	1:B:311:ARG:HA	1.63	0.80
1:B:118:ILE:HD12	1:B:118:ILE:O	1.85	0.77

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	213/394 (54%)	203 (95%)	6 (3%)	4 (2%)	8	13
1	B	376/394 (95%)	328 (87%)	36 (10%)	12 (3%)	4	5
All	All	589/788 (75%)	531 (90%)	42 (7%)	16 (3%)	5	7

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	224	LYS
1	B	228	LEU
1	B	325	VAL
1	B	326	GLN
1	A	17	ASP

5.3.2 Protein sidechains (i)

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	179/336 (53%)	163 (91%)	16 (9%)	9	19
1	B	330/336 (98%)	323 (98%)	7 (2%)	53	78
All	All	509/672 (76%)	486 (96%)	23 (4%)	27	51

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	112	LEU
1	A	158	PHE
1	B	175	LEU
1	A	129	PHE
1	A	175	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	195	GLN
1	B	205	ASN
1	B	296	GLN
1	B	121	HIS
1	B	152	ASN

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)

There are no ligands in this entry.

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	206/394 (52%)	0.80	21 (10%)	6	6	8, 26, 48, 77	0
1	B	369/394 (93%)	2.18	161 (43%)	0	0	9, 38, 110, 114	18 (4%)
All	All	575/788 (72%)	1.69	182 (31%)	0	0	8, 30, 108, 114	18 (3%)

The worst 5 of 182 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	325	VAL	17.5
1	B	257	ILE	12.6
1	B	318	LEU	9.5
1	B	285	THR	8.8
1	B	381	GLN	8.6

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.