wwPDB X-ray Structure Validation Summary Report (i)

Aug 26, 2025 – 01:05 pm BST

PDB ID	:	9F79 / pdb_00009f79
Title	:	Crystal structure of the S. cerevisiae eIF2beta N-terminal tail bound to the C-terminal domain of eIF5
Authors	:	Kuhle, B.; Marintchev, A.; Ficner, R.
Deposited on	:	2024-05-03
Resolution	:	2.00 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

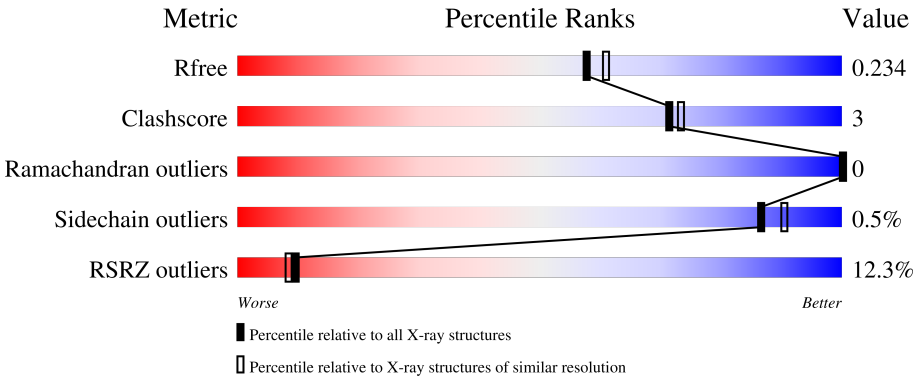
MolProbity	:	4-5-2 with Phenix2.0rc1
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.45.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

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}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (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 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	210	10% 85% 12% .
1	B	210	9% 88% 8% .
2	C	73	23% 49% 5% 45%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3824 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 Eukaryotic translation initiation factor 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	204	Total	C	N	O	S	0	3	0
			1650	1048	266	331	5			
1	B	203	Total	C	N	O	S	0	4	0
			1644	1045	262	332	5			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	196	GLY	-	expression tag	UNP P38431
A	197	PRO	-	expression tag	UNP P38431
A	198	LEU	-	expression tag	UNP P38431
A	199	GLY	-	expression tag	UNP P38431
A	200	SER	-	expression tag	UNP P38431
B	196	GLY	-	expression tag	UNP P38431
B	197	PRO	-	expression tag	UNP P38431
B	198	LEU	-	expression tag	UNP P38431
B	199	GLY	-	expression tag	UNP P38431
B	200	SER	-	expression tag	UNP P38431

- Molecule 2 is a protein called Eukaryotic translation initiation factor 2 subunit beta.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	40	Total	C	N	O	0	0	0
			305	192	50	63			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	34	GLY	-	expression tag	UNP P09064
C	35	PRO	-	expression tag	UNP P09064
C	36	LEU	-	expression tag	UNP P09064
C	37	GLY	-	expression tag	UNP P09064

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	38	SER	-	expression tag	UNP P09064

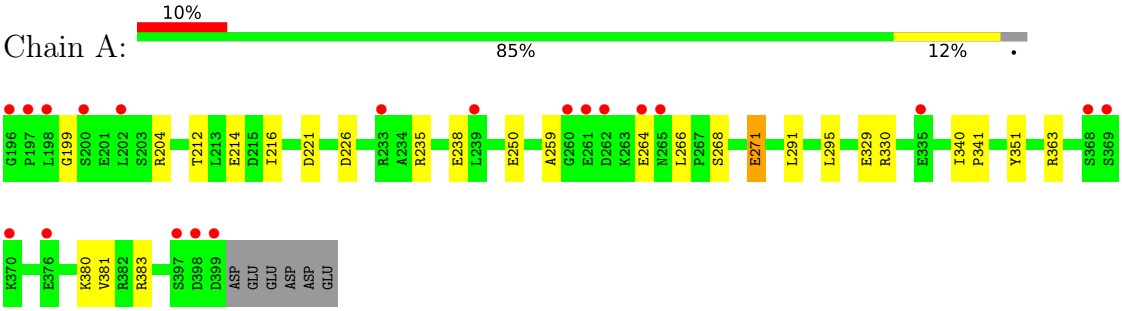
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	109	Total O 109 109	0	0
3	B	108	Total O 108 108	0	0
3	C	8	Total O 8 8	0	0

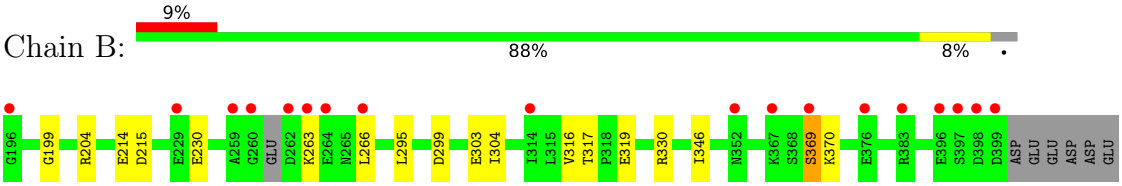
3 Residue-property plots (i)

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.

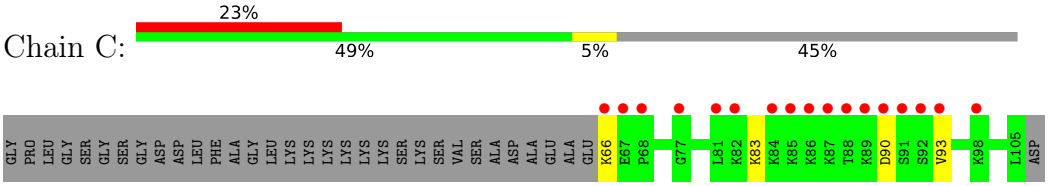
• Molecule 1: Eukaryotic translation initiation factor 5



• Molecule 1: Eukaryotic translation initiation factor 5



• Molecule 2: Eukaryotic translation initiation factor 2 subunit beta



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, α, β, γ	74.18Å 74.18Å 224.42Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	37.40 – 2.00 37.40 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (37.40-2.00) 99.9 (37.40-2.00)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.73 (at 2.00Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.207 , 0.228 0.224 , 0.234	Depositor DCC
R_{free} test set	2411 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	30.4	Xtriage
Anisotropy	0.402	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 50.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.048 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3824	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.23% 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	1.05	6/1686 (0.4%)	1.33	9/2275 (0.4%)
1	B	1.06	8/1682 (0.5%)	1.38	10/2269 (0.4%)
2	C	0.94	0/306	1.25	0/404
All	All	1.05	14/3674 (0.4%)	1.35	19/4948 (0.4%)

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	A	0	3
1	B	0	2
All	All	0	5

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	271[A]	GLU	C-O	7.03	1.32	1.24
1	A	271[B]	GLU	C-O	7.03	1.32	1.24
1	B	214[A]	GLU	C-O	6.82	1.32	1.24
1	B	214[B]	GLU	C-O	6.82	1.32	1.24
1	B	230[A]	GLU	C-O	6.72	1.31	1.24

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	370	LYS	N-CA-C	-9.95	101.47	114.31
1	A	383[A]	ARG	CA-C-O	8.19	129.23	120.55
1	A	383[B]	ARG	CA-C-O	8.19	129.23	120.55
1	B	230[A]	GLU	CA-C-O	8.05	128.96	120.42
1	B	230[B]	GLU	CA-C-O	8.05	128.96	120.42

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	204	ARG	Sidechain
1	A	330	ARG	Sidechain
1	A	363	ARG	Sidechain
1	B	204	ARG	Sidechain
1	B	330	ARG	Sidechain

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	1650	0	1642	11	0
1	B	1644	0	1632	10	0
2	C	305	0	326	3	0
3	A	109	0	0	1	0
3	B	108	0	0	2	0
3	C	8	0	0	0	0
All	All	3824	0	3600	23	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 3.

The worst 5 of 23 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:266:LEU:HD21	1:B:295:LEU:HD22	1.38	1.04
1:B:266:LEU:HD21	1:B:295:LEU:CD2	1.94	0.98
1:B:369:SER:HB3	3:B:560:HOH:O	1.81	0.79
1:B:266:LEU:CD2	1:B:295:LEU:HD22	2.21	0.65
1:B:263:LYS:O	1:B:266:LEU:HB2	1.99	0.63

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	205/210 (98%)	202 (98%)	3 (2%)	0	100	100
1	B	203/210 (97%)	201 (99%)	2 (1%)	0	100	100
2	C	38/73 (52%)	38 (100%)	0	0	100	100
All	All	446/493 (90%)	441 (99%)	5 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	182/185 (98%)	181 (100%)	1 (0%)	86	90
1	B	182/185 (98%)	181 (100%)	1 (0%)	86	90
2	C	33/57 (58%)	33 (100%)	0	100	100
All	All	397/427 (93%)	395 (100%)	2 (0%)	86	90

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	264	GLU
1	B	369	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	323	ASN
1	B	265	ASN
1	B	283	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 oligosaccharides 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

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^{th} 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	204/210 (97%)	0.48	20 (9%)	14	13	19, 36, 57, 70	7 (3%)
1	B	203/210 (96%)	0.63	18 (8%)	17	16	18, 38, 56, 76	8 (3%)
2	C	40/73 (54%)	2.09	17 (42%)	1	1	37, 58, 84, 90	0
All	All	447/493 (90%)	0.69	55 (12%)	9	8	18, 38, 64, 90	15 (3%)

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	399	ASP	7.3
2	C	90	ASP	6.1
1	B	399	ASP	4.8
2	C	81	LEU	4.6
2	C	87	LYS	4.2

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 oligosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.