



# wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 04:24 PM BST

PDB ID : 1Q8K  
Title : Solution structure of alpha subunit of human eIF2  
Authors : Ito, T.; Marintchev, A.; Wagner, G.  
Deposited on : 2003-08-21

This is a wwPDB NMR 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/NMRValidationReportHelp>

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:

Cyrange : Kirchner and Güntert (2011)  
NmrClust : Kelley et al. (1996)  
MolProbity : 4.02b-467  
Mogul : unknown  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
ShiftChecker : rb-20027457  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027457

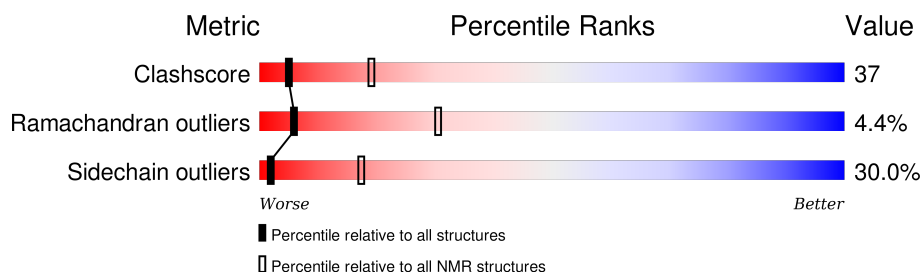
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 30%.

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)	NMR archive (#Entries)
Clashscore	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	308	

## 2 Ensemble composition and analysis

This entry contains 15 models. Model 7 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:11-A:35, A:42-A:183 (167)	0.83	7
2	A:187-A:198, A:206-A:219, A:223-A:228, A:233-A:276 (76)	0.27	6
3	A:278-A:292 (15)	1.30	4

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 2, 4, 6, 11, 14, 15
2	3, 9, 10, 12, 13
3	5, 7, 8

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 4874 atoms, of which 2449 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Eukaryotic translation initiation factor 2 subunit 1.

Mol	Chain	Residues	Atoms						Trace
1	A	300	Total	C	H	N	O	S	0
			4874	1515	2449	427	469	14	

There are 12 discrepancies between the modelled and reference sequences:

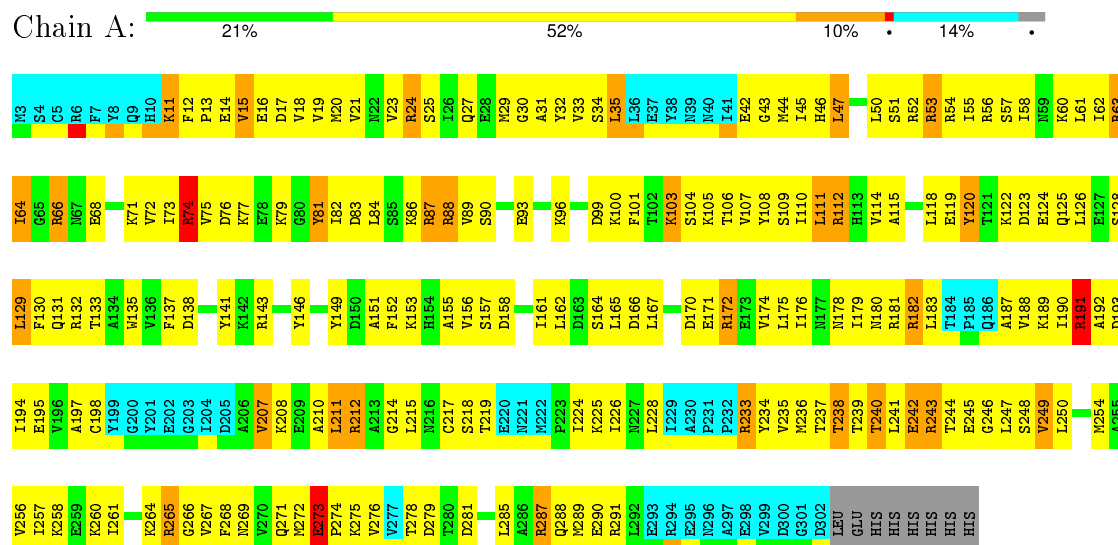
Chain	Residue	Modelled	Actual	Comment	Reference
A	3	MET	-	CLONING ARTIFACT	UNP P05198
A	27	GLN	ALA	ENGINEERED	UNP P05198
A	46	HIS	LEU	ENGINEERED	UNP P05198
A	71	LYS	VAL	ENGINEERED	UNP P05198
A	303	LEU	-	CLONING ARTIFACT	UNP P05198
A	304	GLU	-	CLONING ARTIFACT	UNP P05198
A	305	HIS	-	EXPRESSION TAG	UNP P05198
A	306	HIS	-	EXPRESSION TAG	UNP P05198
A	307	HIS	-	EXPRESSION TAG	UNP P05198
A	308	HIS	-	EXPRESSION TAG	UNP P05198
A	309	HIS	-	EXPRESSION TAG	UNP P05198
A	310	HIS	-	EXPRESSION TAG	UNP P05198

## 4 Residue-property plots

### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

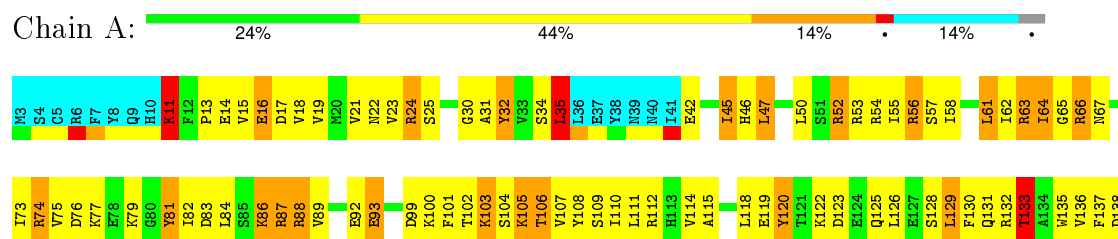
- Molecule 1: Eukaryotic translation initiation factor 2 subunit 1



### 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 7. Colouring as in section 4.1 above.

- Molecule 1: Eukaryotic translation initiation factor 2 subunit 1



N269	V270	Q271	N272	E273	P274	K275	V276	V277	T278	D279	T280	D281	E282	T283	E284	L285	A286	R287	Q288	N289	E290	R291	L292	E293	R294	E295	N296	A297	E298	V299	D300	G301	D302	LEU	GLU	HIS	HIS	HIS	HIS	HIS	HIS																		
A206	V207	R208	E209	A210	L211	R212	A213	G214	L215	N216	G217	S218	T219	E220	N221	M222	P223	I224	K225	I226	N227	L228	I229	A230	P231	P232	R233	V234	V235	M236	T237	T238	T239	T240	L241	E242	R243	T244	E245	G246	L247	S248	V249	L250	M254	A255	V256	I257	K258	E259	K260	I261	K264	R265	G266	V267	F268		
Y141	K142	R143	P144	G145	Y146		Y149		D150	A151	F152		D158		I161	L162	D163	S164	L165	D166	L167	N168	E169	D170	E171	R172	E173	V174	L175	I176	N177	N178	I179	N180	R181	R182	L183	T184	P185	Q186	A187	V188	K189	I190	R191	A192	D193	I194	E195	V196	A197	C198	Y199	G200	Y201	E202	G203	I204	D205

## 5 Refinement protocol and experimental data overview

Of the 51 calculated structures, 15 were deposited, based on the following criterion: ?.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
NIH-XPLOR1	refinement	2.1
NIH-XPLOR1	structure solution	2.1

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	BMRB entry 5917
Number of chemical shift lists	1
Total number of shifts	1122
Number of shifts mapped to atoms	1122
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	30%

No validations of the models with respect to experimental NMR restraints is performed at this time.

## 6 Model quality [i](#)

### 6.1 Standard geometry [i](#)

There are no covalent bond-length or bond-angle outliers.

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	Planarity
1	A	0.0±0.0	22.1±0.7
All	All	0	332

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

5 of 23 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	287	ARG	Sidechain	15
1	A	182	ARG	Sidechain	15
1	A	24	ARG	Sidechain	15
1	A	52	ARG	Sidechain	15
1	A	191	ARG	Sidechain	15

### 6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	2082	2144	2144	158±12
All	All	31230	32160	32160	2374

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 37.

5 of 992 unique clashes are listed below, sorted by their clash magnitude.



Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:190:ILE:HG21	1:A:247:LEU:HD21	1.10	1.18	7	2
1:A:190:ILE:CG2	1:A:247:LEU:HD21	1.02	1.83	7	2
1:A:194:ILE:HD11	1:A:196:VAL:HG23	0.99	1.31	9	1
1:A:247:LEU:HD11	1:A:276:VAL:HG13	0.97	1.36	14	9
1:A:31:ALA:HB3	1:A:45:ILE:HG23	0.93	1.40	10	6

## 6.3 Torsion angles [i](#)

### 6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	258/308 (84%)	219±4 (85±1%)	27±3 (11±1%)	11±2 (4±1%)	6	30
All	All	3870/4620 (84%)	3288 (85%)	412 (11%)	170 (4%)	6	30

5 of 35 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	273	GLU	15
1	A	64	ILE	15
1	A	242	GLU	13
1	A	11	LYS	12
1	A	289	MET	9

### 6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	233/278 (84%)	163±7 (70±3%)	70±7 (30±3%)	2	17
All	All	3495/4170 (84%)	2446 (70%)	1049 (30%)	2	17

5 of 184 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	129	LEU	15
1	A	248	SER	14
1	A	81	TYR	14
1	A	11	LYS	13
1	A	211	LEU	13

### 6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 6.7 Other polymers [i](#)

There are no such molecules in this entry.

### 6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 7 Chemical shift validation [i](#)

The completeness of assignment taking into account all chemical shift lists is 30% for the well-defined parts and 29% for the entire structure.

### 7.1 Chemical shift list 1

File name: BMRB entry 5917

Chemical shift list name: *assigned\_chem\_shift\_list\_1*

#### 7.1.1 Bookkeeping [i](#)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1122
Number of shifts mapped to atoms	1122
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

#### 7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	288	$0.09 \pm 0.14$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	0	—	—
$^{13}\text{C}'$	278	$-0.14 \pm 0.07$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	278	$0.36 \pm 0.17$	None needed ( $< 0.5$ ppm)

#### 7.1.3 Completeness of resonance assignments [i](#)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 30%, i.e. 991 atoms were assigned a chemical shift out of a possible 3329. 0 out of 47 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	991/1278 (78%)	246/510 (48%)	499/516 (97%)	246/252 (98%)
Sidechain	0/1900 (0%)	0/1108 (0%)	0/686 (0%)	0/106 (0%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	0/151 (0%)	0/80 (0%)	0/67 (0%)	0/4 (0%)
Overall	991/3329 (30%)	246/1698 (14%)	499/1269 (39%)	246/362 (68%)

#### 7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

#### 7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:

