wwPDB NMR Structure Validation Summary Report

Jun 3, 2023 – 12:19 PM EDT

PDB ID	:	6P6C
BMRB ID	:	30614
Title	:	CS-Rosetta Model of PEA-15 Death Effector Domain in the Complex with ERK2
Authors	:	Wei, Y.; Crespo Flores, S.; Cabezas, A.; Hassan, S.
Deposited on	:	2019-06-03

This is a wwPDB NMR 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/NMRValidationReportHelp
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 (1)) were used in the production of this report:

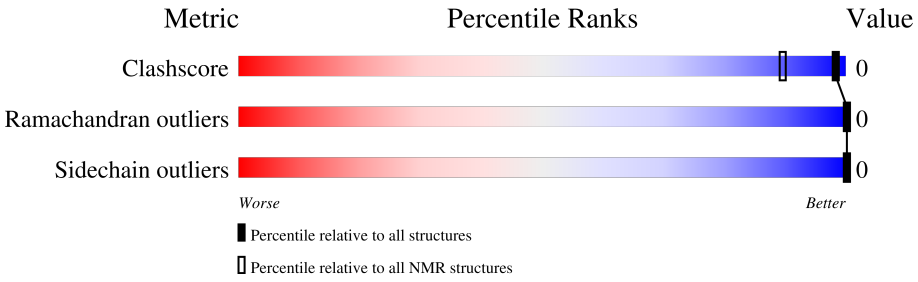
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 26%.

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	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	130	48% 22% 31%

2 Ensemble composition and analysis

This entry contains 10 models. Model 6 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:26-A:87 (62)	2.81	6

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 2 clusters. No single-model clusters were found.

Cluster number	Models
1	6, 7, 8, 9, 10
2	1, 2, 3, 4, 5

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1458 atoms, of which 726 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Astrocytic phosphoprotein PEA-15.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	90	1458	461	726	119	149	3	0

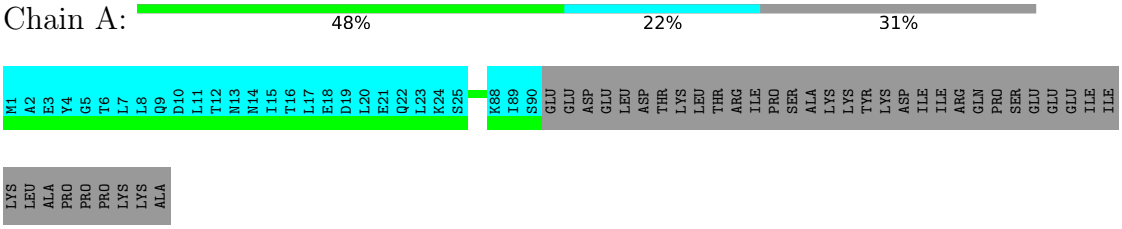
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide 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: Astrocytic phosphoprotein PEA-15

Chain A:

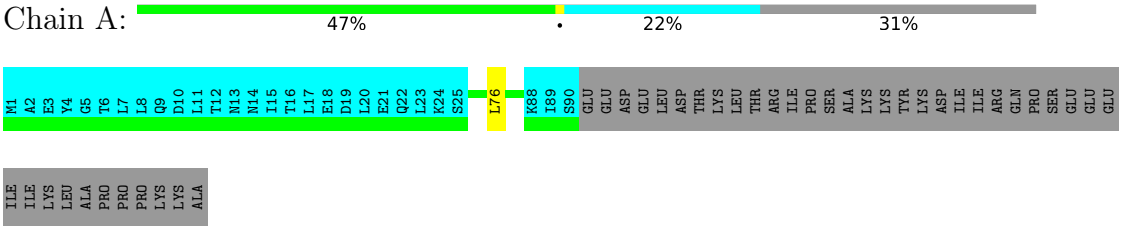


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 6. Colouring as in section 4.1 above.

• Molecule 1: Astrocytic phosphoprotein PEA-15

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing.*

Of the 3000 calculated structures, 10 were deposited, based on the following criterion: *structures with the lowest energy.*

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

Software name	Classification	Version
CS-ROSETTA	refinement	
CS-ROSETTA	structure calculation	

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)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	465
Number of shifts mapped to atoms	333
Number of unparsed shifts	0
Number of shifts with mapping errors	132
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	26%

6 Model quality

6.1 Standard geometry

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

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 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 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	511	499	499	0±0
All	All	5110	4990	4990	2

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 0.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:56:ASP:C	1:A:56:ASP:OD1	0.40	2.60	8	1
1:A:76:LEU:HD23	1:A:76:LEU:C	0.40	2.37	6	1

6.3 Torsion angles

6.3.1 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 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	62/130 (48%)	61±0 (98±1%)	1±0 (2±1%)	0±0 (0±0%)	100	100
All	All	620/1300 (48%)	610 (98%)	10 (2%)	0 (0%)	100	100

There are no Ramachandran outliers.

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	59/122 (48%)	59±0 (100±0%)	0±0 (0±0%)	100	100
All	All	590/1220 (48%)	590 (100%)	0 (0%)	100	100

There are no protein residues with a non-rotameric sidechain to report.

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 monosaccharides 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

The completeness of assignment taking into account all chemical shift lists is 26% for the well-defined parts and 27% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *PEA-15-ERK2.str*

7.1.1 Bookkeeping

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	465
Number of shifts mapped to atoms	333
Number of unparsed shifts	0
Number of shifts with mapping errors	132
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 132) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	91	GLU	H	8.499	.	.
1	A	91	GLU	N	123.003	.	.
1	A	91	GLU	CA	56.604	.	.
1	A	91	GLU	CB	29.691	.	.
1	A	92	GLU	H	8.324	.	.
1	A	92	GLU	N	120.053	.	.
1	A	92	GLU	CA	56.756	.	.
1	A	92	GLU	CB	29.605	.	.
1	A	93	ASP	H	8.244	.	.
1	A	93	ASP	N	120.797	.	.
1	A	93	ASP	CA	54.599	.	.
1	A	93	ASP	CB	40.921	.	.
1	A	94	GLU	H	8.265	.	.
1	A	94	GLU	N	121.029	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	94	GLU	CA	56.507	.	.
1	A	94	GLU	CB	29.637	.	.
1	A	95	LEU	H	8.185	.	.
1	A	95	LEU	N	122.061	.	.
1	A	95	LEU	CA	55.16	.	.
1	A	95	LEU	CB	41.538	.	.
1	A	96	ASP	H	8.303	.	.
1	A	96	ASP	N	121.048	.	.
1	A	96	ASP	CA	54.474	.	.
1	A	96	ASP	CB	40.723	.	.
1	A	97	THR	H	8.056	.	.
1	A	97	THR	N	114.55	.	.
1	A	97	THR	CA	61.96	.	.
1	A	97	THR	CB	69.259	.	.
1	A	98	LYS	H	8.254	.	.
1	A	98	LYS	N	122.293	.	.
1	A	98	LYS	CA	56.561	.	.
1	A	98	LYS	CB	31.914	.	.
1	A	99	LEU	H	8.054	.	.
1	A	99	LEU	N	121.1	.	.
1	A	99	LEU	CA	55.112	.	.
1	A	99	LEU	CB	41.514	.	.
1	A	100	THR	H	8.002	.	.
1	A	100	THR	N	114.399	.	.
1	A	100	THR	CA	62.024	.	.
1	A	100	THR	CB	69.231	.	.
1	A	101	ARG	H	8.198	.	.
1	A	101	ARG	N	123.816	.	.
1	A	101	ARG	CA	56.281	.	.
1	A	101	ARG	CB	31.972	.	.
1	A	102	ILE	H	8.124	.	.
1	A	102	ILE	N	123.335	.	.
1	A	102	ILE	CA	54.922	.	.
1	A	102	ILE	CB	41.499	.	.
1	A	104	SER	H	7.652	.	.
1	A	104	SER	N	119.365	.	.
1	A	104	SER	CA	58.387	.	.
1	A	104	SER	CB	63.413	.	.
1	A	105	ALA	CA	52.404	.	.
1	A	105	ALA	CB	18.637	.	.
1	A	106	LYS	H	8.423	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	106	LYS	N	120.606	.	.
1	A	106	LYS	CA	56.646	.	.
1	A	106	LYS	CB	29.62	.	.
1	A	107	LYS	H	8.203	.	.
1	A	107	LYS	N	122.371	.	.
1	A	107	LYS	CA	56.607	.	.
1	A	107	LYS	CB	29.618	.	.
1	A	108	TYR	H	8.18	.	.
1	A	108	TYR	N	120.745	.	.
1	A	108	TYR	CA	54.634	.	.
1	A	108	TYR	CB	40.784	.	.
1	A	109	LYS	H	8.123	.	.
1	A	109	LYS	N	122.102	.	.
1	A	109	LYS	CA	56.312	.	.
1	A	109	LYS	CB	32.207	.	.
1	A	110	ASP	H	8.196	.	.
1	A	110	ASP	N	121.0	.	.
1	A	110	ASP	CA	54.503	.	.
1	A	110	ASP	CB	40.905	.	.
1	A	111	ILE	H	8.006	.	.
1	A	111	ILE	N	120.587	.	.
1	A	111	ILE	CA	60.701	.	.
1	A	111	ILE	CB	38.228	.	.
1	A	112	ILE	H	8.173	.	.
1	A	112	ILE	N	124.965	.	.
1	A	112	ILE	CA	60.588	.	.
1	A	112	ILE	CB	37.679	.	.
1	A	113	ARG	H	8.353	.	.
1	A	113	ARG	N	125.7	.	.
1	A	113	ARG	CA	55.158	.	.
1	A	113	ARG	CB	30.196	.	.
1	A	114	GLN	H	8.514	.	.
1	A	114	GLN	N	123.358	.	.
1	A	114	GLN	CA	53.209	.	.
1	A	114	GLN	CB	28.133	.	.
1	A	115	PRO	CA	63.202	.	.
1	A	115	PRO	CB	31.402	.	.
1	A	116	SER	H	8.444	.	.
1	A	116	SER	N	115.98	.	.
1	A	116	SER	CA	58.098	.	.
1	A	116	SER	CB	63.608	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	117	GLU	H	8.527	.	.
1	A	117	GLU	N	122.274	.	.
1	A	117	GLU	CA	57.074	.	.
1	A	117	GLU	CB	29.539	.	.
1	A	118	GLU	CA	56.646	.	.
1	A	118	GLU	CB	29.651	.	.
1	A	119	GLU	H	8.174	.	.
1	A	119	GLU	N	121.331	.	.
1	A	119	GLU	CA	56.612	.	.
1	A	119	GLU	CB	29.627	.	.
1	A	120	ILE	H	8.071	.	.
1	A	120	ILE	N	122.113	.	.
1	A	120	ILE	CA	60.73	.	.
1	A	120	ILE	CB	38.049	.	.
1	A	121	ILE	CA	59.916	.	.
1	A	121	ILE	CB	38.093	.	.
1	A	122	LYS	H	8.116	.	.
1	A	122	LYS	N	124.515	.	.
1	A	122	LYS	CA	56.035	.	.
1	A	122	LYS	CB	32.309	.	.
1	A	123	LEU	CA	54.504	.	.
1	A	123	LEU	CB	41.845	.	.
1	A	124	ALA	H	8.184	.	.
1	A	124	ALA	N	126.244	.	.
1	A	124	ALA	CA	50.063	.	.
1	A	124	ALA	CB	17.682	.	.
1	A	127	PRO	CA	62.058	.	.
1	A	127	PRO	CB	31.915	.	.
1	A	128	LYS	H	8.345	.	.
1	A	128	LYS	N	121.772	.	.
1	A	128	LYS	CA	56.299	.	.
1	A	128	LYS	CB	31.96	.	.
1	A	129	LYS	H	8.25	.	.
1	A	129	LYS	N	123.894	.	.
1	A	130	ALA	H	7.94	.	.
1	A	130	ALA	N	125.549	.	.

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	120	-0.40 ± 0.18	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	115	1.06 ± 0.14	Should be checked
$^{13}C'$	0	—	None (insufficient data)
^{15}N	115	0.97 ± 0.19	Should be applied

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 26%, i.e. 228 atoms were assigned a chemical shift out of a possible 869. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	173/307 (56%)	57/123 (46%)	59/124 (48%)	57/60 (95%)
Sidechain	55/488 (11%)	0/315 (0%)	55/155 (35%)	0/18 (0%)
Aromatic	0/74 (0%)	0/37 (0%)	0/34 (0%)	0/3 (0%)
Overall	228/869 (26%)	57/475 (12%)	114/313 (36%)	57/81 (70%)

7.1.4 Statistically unusual chemical shifts

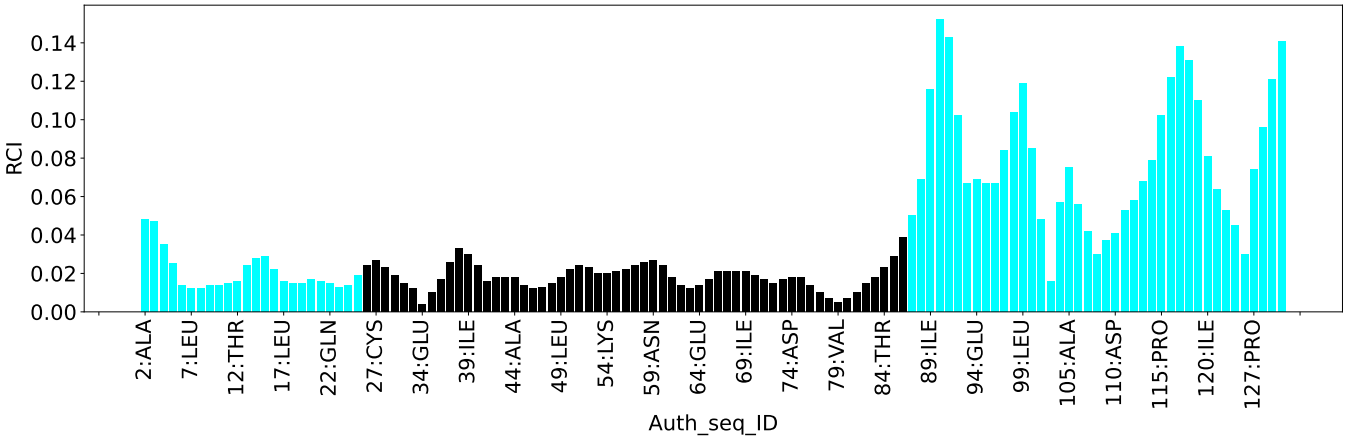
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	34	GLU	CA	71.16	47.03 – 67.62	6.7
1	A	79	VAL	CA	77.14	48.38 – 76.73	5.2

7.1.5 Random Coil Index (RCI) plots

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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:


RCI
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
2:ALA
7:LEU
12:THR
17:LEU
22:GLN
27:CYS
34:GLU
39:ILE
44:ALA
49:LEU
54:LYS
59:ASN
64:GLU
69:ILE
74:ASP
79:VAL
84:THR
89:ILE
94:GLU
99:LEU
105:ALA
110:ASP
115:PRO
120:ILE
127:PRO
Auth_seq_ID

8 NMR restraints analysis (i)

No restraints data found

9 Distance violation analysis (i)

No distance restraints data found

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found