wwPDB NMR Structure Validation Summary Report (i)

Jan 5, 2021 – 08:46 AM GMT

PDB ID	:	6ZPR
Title	:	Solution structure of MLKL executioner domain in complex with a covalent inhibitor
Authors	:	Ruebbelke, M.; Bauer, M.; Hamilton, J.; Binder, F.; Nar, H.; Zeeb, M.
Deposited on	:	2020-07-09

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 following versions of software and data (see references (1)) were used in the production of this report:

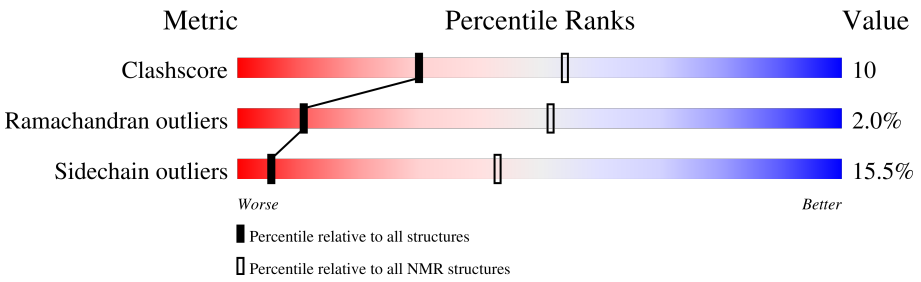
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.16
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.16

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 79%.

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	157	

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *target function*.

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:1-A:47, A:55-A:129, A:133-A:151 (141)	0.35	3

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 4 clusters and 3 single-model clusters were found.

Cluster number	Models
1	2, 3, 4, 12, 14, 15, 16, 17
2	8, 10, 18, 19
3	6, 7, 11
4	1, 9
Single-model clusters	5; 13; 20

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2597 atoms, of which 1309 are hydrogens and 0 are deuteriums.

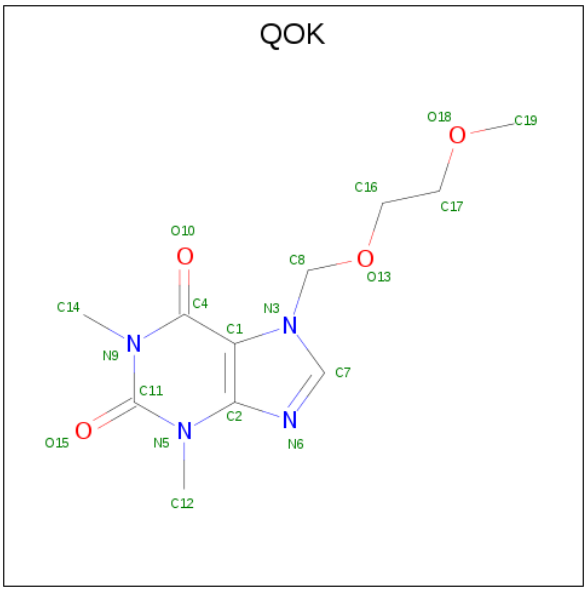
- Molecule 1 is a protein called Mixed lineage kinase domain-like protein,Mixed lineage kinase domain-like protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	157	2563	786	1294	239	235	9	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q8NB16
A	-1	SER	-	expression tag	UNP Q8NB16
A	0	PRO	-	expression tag	UNP Q8NB16
A	1	GLY	-	expression tag	UNP Q8NB16

- Molecule 2 is 7-(2-methoxyethoxymethyl)-1,3-dimethyl-purine-2,6-dione (three-letter code: QOK) (formula: $C_{11}H_{16}N_4O_4$).



Mol	Chain	Residues	Atoms				
			Total	C	H	N	O
2	A	1	34	11	15	4	4

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: Mixed lineage kinase domain-like protein,Mixed lineage kinase domain-like protein

Chain A: 64% 24% •• 10%

G-2 S-1 P0 L4 K5 H6 I7 I8 T9 L10 G11 Q12 V13 K22 K25 L31 L36 G37 L38 I39 K40 P41 L42 E43 Q48 G49 K50 R51 S52 V53 P54 L58 N62 N63 R64 F65 K66 A67 A68 L69 E70 L97 K104 L105 S106 D107 V108 W109 L112

Q120 P123 V124 S125 P126 I127 S128 Q129 G130 A131 S132 D140 R152 R153 D154

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

The representative model is number 3. Colouring as in section 4.1 above.

• Molecule 1: Mixed lineage kinase domain-like protein,Mixed lineage kinase domain-like protein

Chain A: 61% 25% • • 10%

G-2 S-1 P0 G1 E2 N3 L4 K5 H6 I7 I8 T9 L10 G11 Q12 V13 I14 H15 K16 K25 L31 L36 G37 L38 I39 K40 P41 Q48 G49 K50 R51 S52 V53 P54 L58 N62 N63 R64 F65 A68 L69 E70 E71 A72 N73 K78 L89 D94 K95 I96

L97 F98 K99 L105 W109 L112 S113 L114 L115 V118 E119 Q120 S125 P126 I127 S128 Q129 G130 A131 S132 W133 D140 R145 R152 R153 D154

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics*.

Of the 100 calculated structures, 20 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
CYANA	structure calculation	3.98.9

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	1807
Number of shifts mapped to atoms	1802
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	5
Assignment completeness (well-defined parts)	79%

6 Model quality (i)

6.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: QOK

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 (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	1154	1178	1178	23±4
2	A	19	15	0	1±1
All	All	23460	23860	23560	468

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 10.

5 of 113 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:68:ALA:HB1	1:A:105:LEU:HD12	0.85	1.47	1	15
1:A:36:LEU:HD13	1:A:39:ILE:HD11	0.83	1.49	12	10
1:A:68:ALA:HB1	1:A:105:LEU:HD22	0.77	1.54	4	1
1:A:45:LEU:HD11	1:A:116:LEU:HD21	0.69	1.64	17	2
1:A:38:LEU:HD22	1:A:109:TRP:CE3	0.65	2.27	5	18

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	141/157 (90%)	128±2 (91±1%)	10±2 (7±1%)	3±1 (2±1%)	11 52
All	All	2820/3140 (90%)	2569 (91%)	194 (7%)	57 (2%)	11 52

5 of 6 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	125	SER	20
1	A	124	VAL	11
1	A	129	GLN	10
1	A	133	TRP	7
1	A	1	GLY	7

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	128/140 (91%)	108±4 (84±3%)	20±4 (16±3%)	5 43
All	All	2560/2800 (91%)	2162 (84%)	398 (16%)	5 43

5 of 67 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	125	SER	20
1	A	58	LEU	20
1	A	63	ASN	20
1	A	31	LEU	16
1	A	66	LYS	16

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)

1 ligand is modelled in this entry.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 79% for the well-defined parts and 79% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *molecule_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	1807
Number of shifts mapped to atoms	1802
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	5
Number of shift outliers (ShiftChecker)	1

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

- No matching atoms found in structure. All 5 occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	201	QOK	Q16	3.566	0.0	1
A	201	QOK	Q12	3.165	0.0	1
A	201	QOK	Q17	3.67	0.0	1
A	201	QOK	Q19	3.459	0.0	1
A	201	QOK	Q14	3.141	0.0	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	156	-0.40 ± 0.07	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	142	0.17 ± 0.06	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
^{15}N	144	-0.00 ± 0.23	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 79%, i.e. 1478 atoms were assigned a chemical shift out of a possible 1868. 23 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	544/699 (78%)	272/279 (97%)	140/282 (50%)	132/138 (96%)
Sidechain	856/1071 (80%)	534/631 (85%)	306/377 (81%)	16/63 (25%)
Aromatic	78/98 (80%)	40/53 (75%)	36/40 (90%)	2/5 (40%)
Overall	1478/1868 (79%)	846/963 (88%)	482/699 (69%)	150/206 (73%)

7.1.4 Statistically unusual chemical shifts (i)

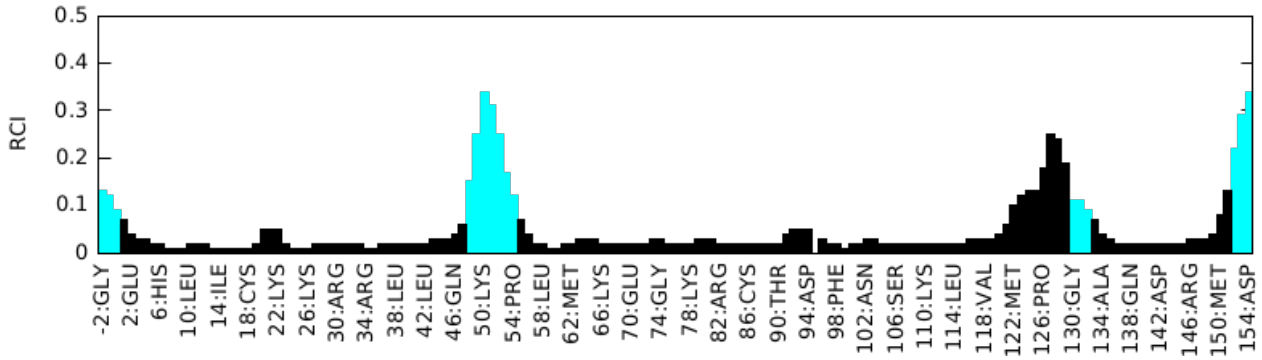
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	113	SER	HB2	2.53	5.18 – 2.58	-5.2

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:


RCI
0.5
0.4
0.3
0.2
0.1
0
-2:GLY
2:GLU
6:HIS
10:LEU
14:ILE
18:CYS
22:LYS
26:LYS
30:ARG
34:ARG
38:LEU
42:LEU
46:GLN
50:LYS
54:PRO
58:LEU
62:MET
66:LYS
70:GLU
74:GLY
78:LYS
82:ARG
86:CYS
90:THR
94:ASP
98:PHE
102:ASN
106:SER
110:LYS
114:LEU
118:VAL
122:MET
126:PRO
130:GLY
134:ALA
138:GLN
142:ASP
146:ARG
150:MET
154:ASP