wwPDB NMR Structure Validation Summary Report (i)

Feb 9, 2022 – 08:48 AM EST

PDB ID	:	1EES
Title	:	SOLUTION STRUCTURE OF CDC42HS COMPLEXED WITH A PEPTIDE DERIVED FROM P-21 ACTIVATED KINASE, NMR, 20 STRUCTURES
Authors	:	Gizachew, D.; Guo, W.; Chohan, K.C.; Sutcliffe, M.J.; Oswald, R.E.
Deposited on	:	2000-02-02

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 (i)) were used in the production of this report:

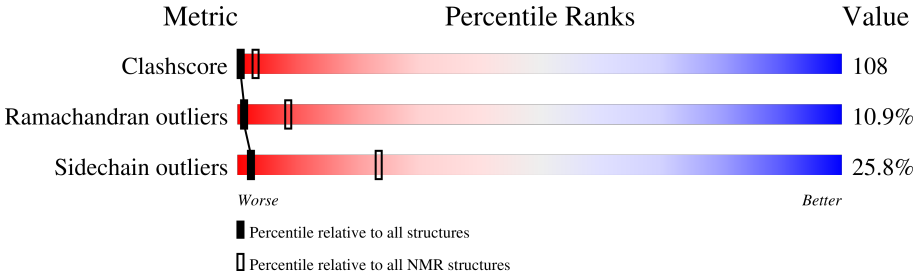
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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 was not calculated.

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	178	13% 62% 20% • •
2	B	46	7% 57% 17% • 17%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 19 is the overall representative, medoid model (most similar to other models). The authors have identified model 9 as representative, based on the following criterion: *closest to the average*.

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:28, A:35-A:178, B:1-B:38 (210)	1.45	19

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. No single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3501 atoms, of which 1750 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called GTP-BINDING PROTEIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	178	2794	893	1405	221	268	7	0

- Molecule 2 is a protein called P21-ACTIVATED KINASE.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	46	707	228	345	61	73	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLY	LYS	see remark 999	UNP Q61036
B	2	SER	GLU	see remark 999	UNP Q61036

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: GTP-BINDING PROTEIN

Chain A: 13% 62% 20% • •

M1 Q2 T3 I4 K5 C6 V7 V8 V9 G10 D11 G12 A13 V14 G15 K16 T17 C18 L19 L20 I21 S22 Y23 T24 T25 N26 K27 F28 P29 S30 E31 Y32 V33 P34 T35 V36 F37 D38 N39 Y40 A41 V42 T43 V44 M45 I46 G47 G48 E49 P50 Y51 T52 L53 G54 L55 F56 D57 T58 A59 G60

Q61 E62 D63 Y64 D65 R66 L67 R68 P69 L70 S71 Y72 P73 Q74 T75 D76 V77 F78 L79 V80 C81 F82 S83 V84 V85 S86 P87 S88 S89 F90 E91 N92 V93 K94 E95 K96 W97 V98 P99 E100 I101 T102 H103 H104 C105 P106 K107 T108 P109 F110 L111 L112 V113 G114 T115 Q116 I117 D118 L119 R120

D121 D122 T125 I126 E127 K128 L129 A130 K131 N132 K133 Q134 K135 P136 I137 T138 P139 K144 L145 A146 R147 D148 L149 K150 A151 V152 K153 Y154 V155 E156 C157 S158 A159 L160 T161 Q162 K163 G164 L165 K166 N167 V168 F169 D170 E171 A172 I173 L174 A175 A176 L177 E178

• Molecule 2: P21-ACTIVATED KINASE

Chain B: 7% 57% 17% • 17%

G1 S2 K3 E4 R5 P6 E7 I8 S9 L10 P11 S12 D13 F14 E15 H16 T17 I18 H19 V20 G21 F22 D23 A24 V25 T26 G27 E28 F29 T30 G31 I32 P33 E34 Q35 W36 A37 R38 L39 L40 Q41 T42 S43 N44 I45 T46

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

The representative model is number 19. Colouring as in section 4.1 above.

• Molecule 1: GTP-BINDING PROTEIN

Chain A: 18% 52% 20% 6% •

M1 Q2 T3 I4 K5 C6 V7 V8 V9 G10 D11 G12 A13 V14 G15 K16 T17 C18 L19 L20 I21 S22 Y23 T24 T25 N26 K27 F28 P29 S30 E31 Y32 V33 P34 T35 V36 F37 D38 N39 Y40 A41 V42 T43 V44 M45 I46 G47 G48 E49 P50 Y51 T52 L53 G54 L55 F56 D57 T58 A59 G60

Q61 E62 D63 Y64 D65 R66 L67 R68 P69 L70 S71 Y72 D76 V77 F78 L79 V80 C81 V84 V85 S86 P87 S88 S89 F90 E91 N92 V93 K94 E95 K96 W97 V98 P99 E100 I101 T102 H103 H104 K107 F110 L111 L112 V113 G114 T115 Q116 I117 D118 L119 R120 D121 D122 P123 S124 T125

• Molecule 2: P21-ACTIVATED KINASE

Chain B:

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *distance geometry simulated annealing Ramachandran refinement.*

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with acceptable covalent geometry,structures with favorable non-bond energy,structures with the least restraint violations,structures with the lowest energy.*

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

Software name	Classification	Version
X-PLOR	structure solution	3.851
X-PLOR	refinement	3.851

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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 (average) 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	1.02±0.01	0±0/1368 (0.0± 0.0%)	1.25±0.01	0±0/1859 (0.0± 0.0%)
2	B	1.12±0.01	0±0/310 (0.0± 0.0%)	1.22±0.02	0±0/422 (0.0± 0.0%)
All	All	1.04	0/33560 (0.0%)	1.24	4/45620 (0.0%)

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	4.0±0.2
2	B	0.0±0.0	1.9±0.2
All	All	0	118

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	27	LYS	N-CA-CB	-5.34	100.99	110.60	5	1
1	A	28	PHE	N-CA-CB	-5.22	101.20	110.60	14	1
1	A	159	ALA	N-CA-CB	-5.21	102.81	110.10	19	1
1	A	107	LYS	N-CA-CB	-5.04	101.52	110.60	12	1

There are no chirality outliers.

5 of 6 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	66	ARG	Sidechain	20
1	A	120	ARG	Sidechain	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	147	ARG	Sidechain	20
2	B	38	ARG	Sidechain	20
1	A	68	ARG	Sidechain	19

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	1341	1362	1362	319±29
2	B	300	279	279	94±13
All	All	32820	32820	32820	7056

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 107.

5 of 3615 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:23:TYR:CE2	1:A:165:LEU:HD11	1.13	1.78	6	2
1:A:23:TYR:CE2	1:A:165:LEU:HD12	1.06	1.84	7	2
1:A:111:LEU:HD13	1:A:152:VAL:HG11	1.05	1.15	16	14
1:A:24:THR:HG22	1:A:25:THR:HG22	1.05	1.28	14	2
1:A:19:LEU:HD13	1:A:159:ALA:HB2	1.03	1.07	20	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	170/178 (96%)	127±4 (75±2%)	27±4 (16±2%)	16±3 (10±2%)	1 10
2	B	37/46 (80%)	21±2 (56±6%)	10±2 (27±6%)	6±2 (17±5%)	0 3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	4140/4480 (92%)	2954 (71%)	733 (18%)	453 (11%)	1 8

5 of 80 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	118	ASP	20
1	A	119	LEU	20
1	A	101	ILE	18
1	A	27	LYS	17
1	A	36	VAL	17

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	152/158 (96%)	113±5 (75±3%)	39±5 (25±3%)	2 24
2	B	32/40 (80%)	23±2 (73±7%)	9±2 (27±7%)	2 21
All	All	3680/3960 (93%)	2732 (74%)	948 (26%)	2 23

5 of 146 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	92	ASN	20
1	A	117	ILE	20
1	A	165	LEU	20
2	B	5	ARG	20
1	A	177	LEU	19

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 (i)

No chemical shift data were provided