wwPDB NMR Structure Validation Summary Report (i)

Feb 8, 2022 – 05:17 PM EST

PDB ID	:	1CRR
Title	:	THE SOLUTION STRUCTURE AND DYNAMICS OF RAS P21. GDP DETERMINED BY HETERONUCLEAR THREE AND FOUR DIMENSIONAL NMR SPECTROSCOPY
Authors	:	Kraulis, P.J.; Domaille, P.J.; Campbell-Burk, S.L.; Van Aken, T.; Laue, E.D.
Deposited on	:	1993-11-24

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:

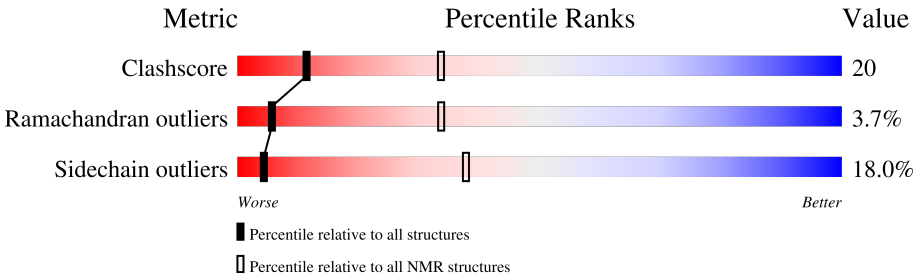
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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	166	54% 37% • 6%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 37 is the overall representative, medoid model (most similar to other models).

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:59, A:67-A:119, A:123-A:166 (156)	0.44	37

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 5 single-model clusters were found.

Cluster number	Models
1	22, 24, 25, 26, 27, 36, 37, 38, 39
2	32, 40
3	23, 30
4	31, 33
Single-model clusters	21; 28; 29; 34; 35

3 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2672 atoms, of which 1316 are hydrogens and 0 are deuteriums.

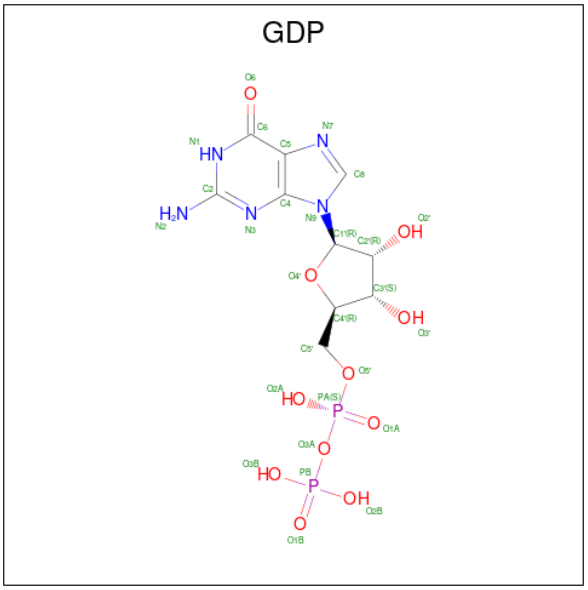
- Molecule 1 is a protein called C-H-RAS P21 PROTEIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	166	2619	824	1296	228	264	7	0

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	
			Total	Mg
2	A	1	1	1

- Molecule 3 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).



Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	P
3	A	1	40	10	12	5	11	2

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		
			Total	H	O
4	A	4	12	8	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: C-H-RAS P21 PROTEIN

Chain A: 54% 37% • 6%

M1 T2 L6 V7 V8 V9 G10 A11 V14 G15 K16 S17 A18 L19 T20 I21 I24 Q25 N26 H27 F28 V29 D30 E31 Y32 D33 I36 E37 D38 K42 E49 L52 I55 L56 D57 T58 A59 G60 Q61 E62 E63 Y64 S65 A66 M67 R68 Y71 M72 R73 F78

L79 C80 V81 F82 A83 I84 N85 N86 T87 D92 I93 Y96 R97 E98 Q99 I100 V103 V109 P110 L113 V114 G115 N116 K117 L120 A121 A122 V125 R128 Q129 A130 L133 Y137 Y141 I142 E143 T144 K147 T148 R149 L159 R164 Q165 H166

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

The representative model is number 37. Colouring as in section 4.1 above.

• Molecule 1: C-H-RAS P21 PROTEIN

Chain A: 57% 31% 7% 6%

M1 T2 E3 Y4 V7 V8 V9 G10 A11 V14 G15 K16 L19 T20 I21 I24 H27 F28 V29 D30 E31 Y32 D33 D38 S39 Y40 R41 K42 Q43 I46 E49 L52 I55 L56 G60 Q61 E62 E63 Y64 S65 A66 M67 R68 Y71 M72 G75 E76

G77 F78 V81 F82 A83 I84 N85 N86 T87 D92 I93 Y96 Q99 I100 V103 S106 D107 D108 V109 P110 L113 V114 G115 N116 K117 L120 A121 A122 R123 R128 L133 A134 Y137 G138 I139 P140 Y141 T144 T148 R149 Y157 T158 L159 R164 Q165

H166

5 Refinement protocol and experimental data overview (i)

Of the ? calculated structures, 20 were deposited, based on the following criterion: *?*.

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

Software name	Classification	Version
X-PLOR	refinement	

No chemical shift data was provided.

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: GDP, MG

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	1251	1233	1233	50±6
3	A	28	12	12	0±1
4	A	4	8	0	0±1
All	All	25680	25060	24900	1010

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 20.

5 of 378 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:93:ILE:HG21	1:A:113:LEU:HD12	0.91	1.40	22	6
1:A:56:LEU:HD13	1:A:71:TYR:CE2	0.89	2.02	29	6
1:A:78:PHE:CE2	1:A:109:VAL:HG11	0.88	2.03	32	2
1:A:56:LEU:HD22	1:A:71:TYR:CG	0.88	2.04	29	10
1:A:93:ILE:HG21	1:A:113:LEU:HD13	0.88	1.42	30	15

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	154/166 (93%)	119±3 (77±2%)	29±3 (19±2%)	6±2 (4±2%)	6 34
All	All	3080/3320 (93%)	2379 (77%)	587 (19%)	114 (4%)	6 34

5 of 37 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	31	GLU	15
1	A	7	VAL	9
1	A	148	THR	7
1	A	26	ASN	7
1	A	33	ASP	6

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	138/144 (96%)	113±3 (82±2%)	25±3 (18±2%)	4 38
All	All	2760/2880 (96%)	2264 (82%)	496 (18%)	4 38

5 of 88 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	2	THR	20
1	A	109	VAL	20
1	A	87	THR	19
1	A	27	HIS	18
1	A	148	THR	15

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)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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)

No chemical shift data were provided