wwPDB NMR Structure Validation Summary Report

Jun 4, 2023 – 12:42 AM EDT

PDB ID	:	7RSC
BMRB ID	:	30734
Title	:	NMR-driven structure of the KRAS4B-G12D "alpha-alpha" dimer on a lipid bilayer nanodisc
Authors	:	Lee, K.; Enomoto, M.; Gebregiworgis, T.; Gasmi-Seabrook, G.M.; Ikura, M.; Marshall, C.B.
Deposited on	:	2021-08-11

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

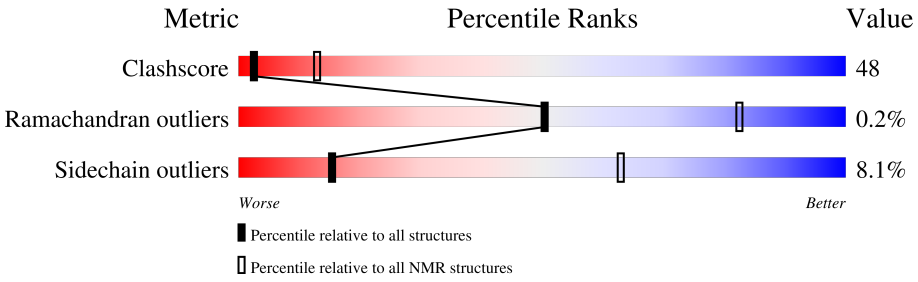
MolProbity	:	4.02b-467
Mogul	:	FAILED
buster-report	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (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 1%.

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	185	77% 13% 9% .
1	B	185	79% 11% 9% .
2	D	200	84% 8% 6%
2	E	200	84% 10% 6%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 5 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:2-A:167, B:2-B:167 (332)	0.98	6
2	D:256-D:441, E:555-E:741 (373)	0.39	5

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

3 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 14322 atoms, of which 1540 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called GTPase KRas.

Mol	Chain	Residues	Atoms						Trace
1	A	183	Total	C	H	N	O	S	0
			1822	917	356	255	287	7	
1	B	183	Total	C	H	N	O	S	0
			1822	917	356	255	287	7	

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP P01116
A	12	ASP	GLY	engineered mutation	UNP P01116
B	1	SER	-	expression tag	UNP P01116
B	12	ASP	GLY	engineered mutation	UNP P01116

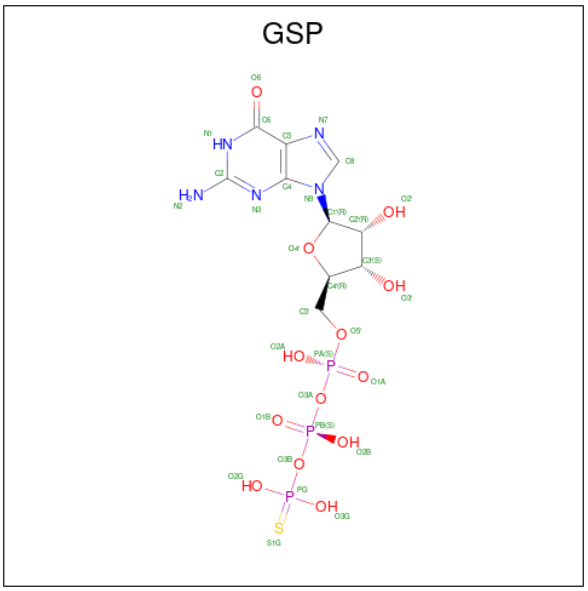
- Molecule 2 is a protein called Apolipoprotein A-I.

Mol	Chain	Residues	Atoms						Trace
2	D	187	Total	C	H	N	O	S	0
			1892	960	360	273	296	3	
2	E	187	Total	C	H	N	O	S	0
			1892	960	360	273	296	3	

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	242	GLY	-	expression tag	UNP P02647
D	243	PRO	-	expression tag	UNP P02647
E	542	GLY	-	expression tag	UNP P02647
E	543	PRO	-	expression tag	UNP P02647

- Molecule 3 is 5'-GUANOSINE-DIPHOSPHATE-MONOTHIOPHOSPHATE (three-letter code: GSP) (formula: $C_{10}H_{16}N_5O_{13}P_3S$) (labeled as "Ligand of Interest" by depositor).



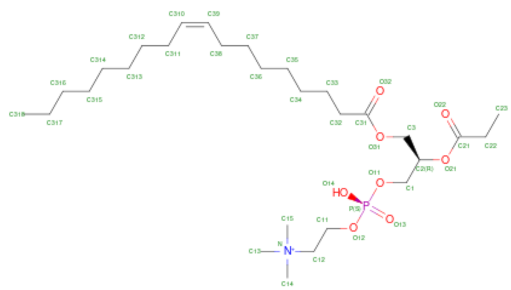
Mol	Chain	Residues	Atoms						
3	A	1	Total	C	H	N	O	P	S
			38	10	6	5	13	3	1
3	B	1	Total	C	H	N	O	P	S
			38	10	6	5	13	3	1

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	
4	A	1	Total	Mg
			1	1
4	B	1	Total	Mg
			1	1

- Molecule 5 is [(2 {R})-3-[oxidanyl-[2-(trimethyl-$l^{4}-azanyl)ethoxy]phosphoryl]oxy-2-propanoyloxy-propyl] ({Z})-octadec-9-enoate (three-letter code: 7Q9) (formula: $C_{29}H_{57}NO_8P$) (labeled as "Ligand of Interest" by depositor).

7Q9



Mol	Chain	Residues	Atoms
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1
5	A	1	Total C N O P 39 29 1 8 1

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				
5	B	1	Total 39	C 29	N 1	O 8	P 1
5	B	1	Total 39	C 29	N 1	O 8	P 1
5	B	1	Total 39	C 29	N 1	O 8	P 1
5	B	1	Total 39	C 29	N 1	O 8	P 1
5	B	1	Total 39	C 29	N 1	O 8	P 1
5	B	1	Total 39	C 29	N 1	O 8	P 1
5	B	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1
5	D	1	Total	C	N	O	P
			39	29	1	8	1

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	D	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1
5	E	1	Total 39	C 29	N 1	O 8	P 1

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1

Continued on next page...

Continued from previous page...

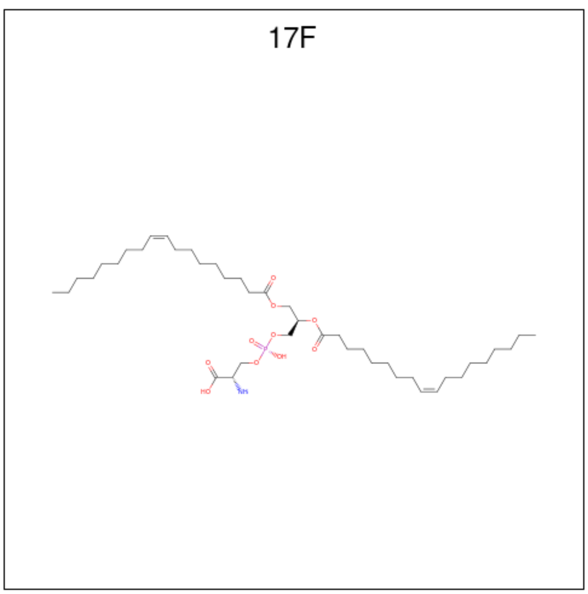
Mol	Chain	Residues	Atoms				
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1
5	E	1	Total	C	N	O	P
			39	29	1	8	1

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms
5	E	1	Total C N O P 39 29 1 8 1
5	E	1	Total C N O P 39 29 1 8 1
5	E	1	Total C N O P 39 29 1 8 1
5	E	1	Total C N O P 39 29 1 8 1
5	E	1	Total C N O P 39 29 1 8 1
5	E	1	Total C N O P 39 29 1 8 1
5	E	1	Total C N O P 39 29 1 8 1
5	E	1	Total C N O P 39 29 1 8 1
5	E	1	Total C N O P 39 29 1 8 1

- Molecule 6 is O-[(S)-({(2R)-2,3-bis[(9Z)-octadec-9-enoyloxy]propyl}oxy)(hydroxy)phosphoryl]-L-serine (three-letter code: 17F) (formula: $C_{42}H_{78}NO_{10}P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms
6	A	1	Total C H N O P 57 42 3 1 10 1

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					
6	A	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	A	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	A	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	B	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	D	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1
6	E	1	Total 57	C 42	H 3	N 1	O 10	P 1

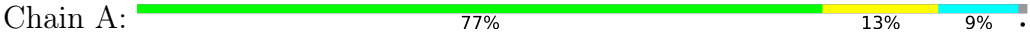
4 Residue-property plots

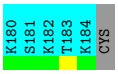
4.1 Average score per residue in the NMR ensemble

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: GTPase KRas

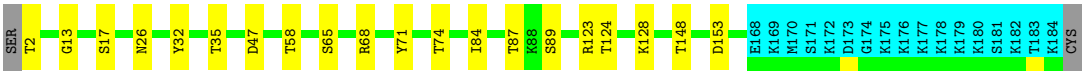
Chain A:





• Molecule 1: GTPase KRas

Chain B:



• Molecule 2: Apolipoprotein A-I

Chain D:

• Molecule 2: Apolipoprotein A-I

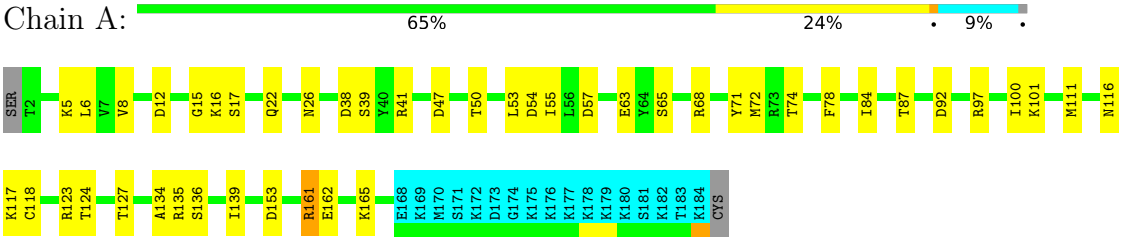
Chain E:

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

The representative model is number 5. Colouring as in section 4.1 above.

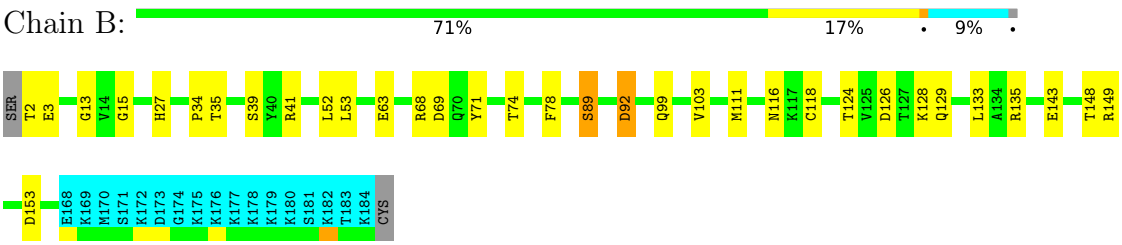
• Molecule 1: GTPase KRas

Chain A:



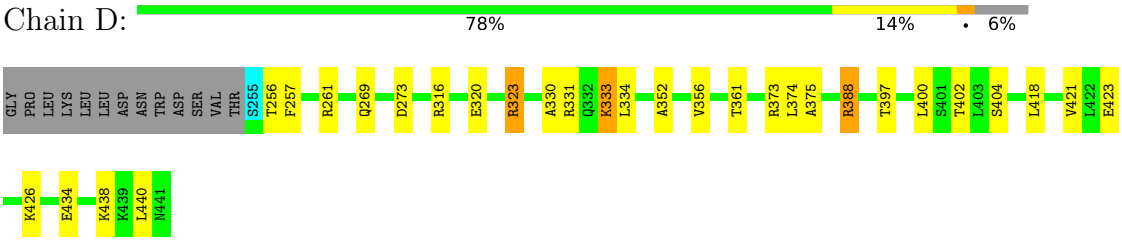
• Molecule 1: GTPase KRas

Chain B:



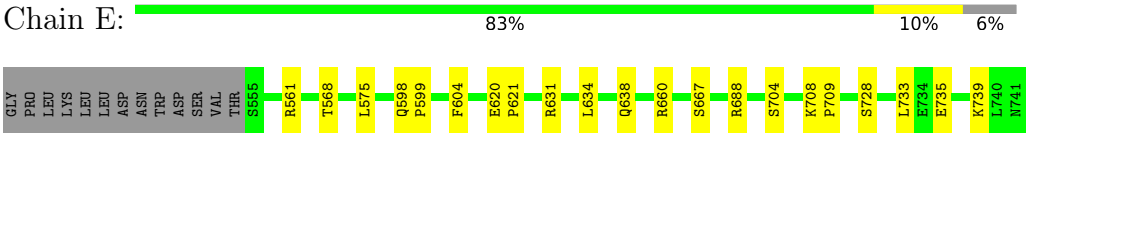
• Molecule 2: Apolipoprotein A-I

Chain D:



• Molecule 2: Apolipoprotein A-I

Chain E:



5 Refinement protocol and experimental data overview

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

Of the 1000 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
HADDOCK	refinement	
HADDOCK	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	154
Number of shifts mapped to atoms	61
Number of unparsed shifts	0
Number of shifts with mapping errors	93
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	1%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 7Q9, GSP, 17F, 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

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	1328	306	1302	17±5
1	B	1328	306	1303	17±4
2	D	1526	358	1531	18±4
2	E	1532	360	1535	21±5
3	A	32	6	12	8±3
3	B	32	6	12	8±2
4	A	1	0	0	1±0
4	B	1	0	0	1±0
5	A	546	0	0	61±7
5	B	273	0	0	35±9
5	D	1755	0	0	187±14
5	E	2418	0	0	275±27
6	A	216	12	304	73±10
6	B	54	3	76	23±7
6	D	594	33	836	199±17
6	E	864	48	1216	213±36
All	All	250000	28760	162529	19712

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 48.

5 of 5532 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
6:D:510:17F:C9	6:D:517:17F:H9	1.24	1.62	8	8
6:D:510:17F:C9	6:D:517:17F:H8A	1.23	1.64	15	12
6:D:510:17F:C12	6:D:517:17F:H8A	1.21	1.66	3	16
6:D:510:17F:C9	6:D:517:17F:H11	1.19	1.67	4	10
6:D:510:17F:H9A	6:D:517:17F:C8	1.18	1.66	18	19

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	165/185 (89%)	156±2 (95±1%)	8±2 (5±1%)	0±1 (0±0%)	54	85
1	B	165/185 (89%)	156±2 (94±1%)	9±2 (5±1%)	0±1 (0±0%)	50	82
2	D	185/200 (92%)	182±2 (98±1%)	3±2 (2±1%)	0±1 (0±0%)	50	82
2	E	185/200 (92%)	180±2 (97±1%)	4±2 (2±1%)	0±1 (0±0%)	50	82
All	All	14000/15400 (91%)	13482 (96%)	493 (4%)	25 (0%)	50	82

5 of 18 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	47	ASP	3
1	B	64	TYR	3
2	E	718	LEU	2
2	E	708	LYS	2
2	D	333	LYS	2

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	147/165 (89%)	133±3 (90±2%)	14±3 (10±2%)	12	58
1	B	147/165 (89%)	133±3 (91±2%)	14±3 (9±2%)	12	58
2	D	162/175 (93%)	151±2 (93±1%)	11±2 (7±1%)	20	68
2	E	163/175 (93%)	152±4 (93±2%)	11±4 (7±2%)	19	68
All	All	12380/13600 (91%)	11379 (92%)	1001 (8%)	15	63

5 of 230 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)
2	D	388	ARG	18
2	D	402	THR	18
1	B	74	THR	17
2	D	316	ARG	16
1	B	2	THR	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)

Mogul failed to run properly - this section will have to be empty.

6.5 Carbohydrates (i)

Mogul failed to run properly - this section will have to be empty.

6.6 Ligand geometry (i)

Mogul failed to run properly - this section will have to be empty.

6.7 Other polymers (i)

Mogul failed to run properly - this section will have to be empty.

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 1% for the well-defined parts and 1% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_0*

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	154
Number of shifts mapped to atoms	61
Number of unparsed shifts	0
Number of shifts with mapping errors	93
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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 93) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	6	LEU	HD11	0.763	.	2
1	A	6	LEU	HD12	0.763	.	2
1	A	6	LEU	HD13	0.763	.	2
1	A	8	VAL	HG21	0.7528	.	2
1	A	8	VAL	HG22	0.7528	.	2
1	A	8	VAL	HG23	0.7528	.	2
1	A	14	VAL	HG21	0.9725	.	2
1	A	14	VAL	HG22	0.9725	.	2
1	A	14	VAL	HG23	0.9725	.	2
1	A	19	LEU	HD11	0.4104	.	2
1	A	19	LEU	HD12	0.4104	.	2
1	A	19	LEU	HD13	0.4104	.	2
1	A	19	LEU	HD21	0.5804	.	2
1	A	19	LEU	HD22	0.5804	.	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	19	LEU	HD23	0.5804	.	2
1	A	21	ILE	HD11	0.4182	.	1
1	A	21	ILE	HD12	0.4182	.	1
1	A	21	ILE	HD13	0.4182	.	1
1	A	23	LEU	HD11	-0.477	.	2
1	A	23	LEU	HD12	-0.477	.	2
1	A	23	LEU	HD13	-0.477	.	2
1	A	24	ILE	HD11	0.2691	.	1
1	A	24	ILE	HD12	0.2691	.	1
1	A	24	ILE	HD13	0.2691	.	1
1	A	36	ILE	HD11	0.5397	.	1
1	A	36	ILE	HD12	0.5397	.	1
1	A	36	ILE	HD13	0.5397	.	1
1	A	44	VAL	HG11	0.5353	.	2
1	A	44	VAL	HG12	0.5353	.	2
1	A	44	VAL	HG13	0.5353	.	2
1	A	45	VAL	HG11	0.4379	.	2
1	A	45	VAL	HG12	0.4379	.	2
1	A	45	VAL	HG13	0.4379	.	2
1	A	46	ILE	HD11	0.2525	.	1
1	A	46	ILE	HD12	0.2525	.	1
1	A	46	ILE	HD13	0.2525	.	1
1	A	55	ILE	HD11	0.3369	.	1
1	A	55	ILE	HD12	0.3369	.	1
1	A	55	ILE	HD13	0.3369	.	1
1	A	56	LEU	HD11	0.5372	.	2
1	A	56	LEU	HD12	0.5372	.	2
1	A	56	LEU	HD13	0.5372	.	2
1	A	79	LEU	HD11	-0.1312	.	2
1	A	79	LEU	HD12	-0.1312	.	2
1	A	79	LEU	HD13	-0.1312	.	2
1	A	79	LEU	HD21	-0.015	.	2
1	A	79	LEU	HD22	-0.015	.	2
1	A	79	LEU	HD23	-0.015	.	2
1	A	81	VAL	HG21	0.4663	.	2
1	A	81	VAL	HG22	0.4663	.	2
1	A	81	VAL	HG23	0.4663	.	2
1	A	84	ILE	HD11	0.5865	.	1
1	A	84	ILE	HD12	0.5865	.	1
1	A	84	ILE	HD13	0.5865	.	1
1	A	93	ILE	HD11	0.6063	.	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	93	ILE	HD12	0.6063	.	1
1	A	93	ILE	HD13	0.6063	.	1
1	A	100	ILE	HD11	0.0964	.	1
1	A	100	ILE	HD12	0.0964	.	1
1	A	100	ILE	HD13	0.0964	.	1
1	A	103	VAL	HG11	0.8243	.	2
1	A	103	VAL	HG12	0.8243	.	2
1	A	103	VAL	HG13	0.8243	.	2
1	A	112	VAL	HG11	0.5953	.	2
1	A	112	VAL	HG12	0.5953	.	2
1	A	112	VAL	HG13	0.5953	.	2
1	A	113	LEU	HD11	0.8748	.	2
1	A	113	LEU	HD12	0.8748	.	2
1	A	113	LEU	HD13	0.8748	.	2
1	A	120	LEU	HD11	0.6958	.	2
1	A	120	LEU	HD12	0.6958	.	2
1	A	120	LEU	HD13	0.6958	.	2
1	A	125	VAL	HG11	-0.3248	.	2
1	A	125	VAL	HG12	-0.3248	.	2
1	A	125	VAL	HG13	-0.3248	.	2
1	A	133	LEU	HD11	0.2014	.	2
1	A	133	LEU	HD12	0.2014	.	2
1	A	133	LEU	HD13	0.2014	.	2
1	A	139	ILE	HD11	0.7118	.	1
1	A	139	ILE	HD12	0.7118	.	1
1	A	139	ILE	HD13	0.7118	.	1
1	A	142	ILE	HD11	0.5035	.	1
1	A	142	ILE	HD12	0.5035	.	1
1	A	142	ILE	HD13	0.5035	.	1
1	A	159	LEU	HD11	0.51	.	2
1	A	159	LEU	HD12	0.51	.	2
1	A	159	LEU	HD13	0.51	.	2
1	A	160	VAL	HG11	-0.2062	.	2
1	A	160	VAL	HG12	-0.2062	.	2
1	A	160	VAL	HG13	-0.2062	.	2
1	A	163	ILE	HD11	0.5261	.	1
1	A	163	ILE	HD12	0.5261	.	1
1	A	163	ILE	HD13	0.5261	.	1

7.1.2 Chemical shift referencing (i)

No chemical shift referencing corrections were calculated (not enough data).

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 1%, i.e. 140 atoms were assigned a chemical shift out of a possible 9884. 0 out of 126 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	16/3515 (0%)	8/1422 (1%)	0/1410 (0%)	8/683 (1%)
Sidechain	124/5741 (2%)	93/3690 (3%)	31/1795 (2%)	0/256 (0%)
Aromatic	0/628 (0%)	0/310 (0%)	0/296 (0%)	0/22 (0%)
Overall	140/9884 (1%)	101/5422 (2%)	31/3501 (1%)	8/961 (1%)

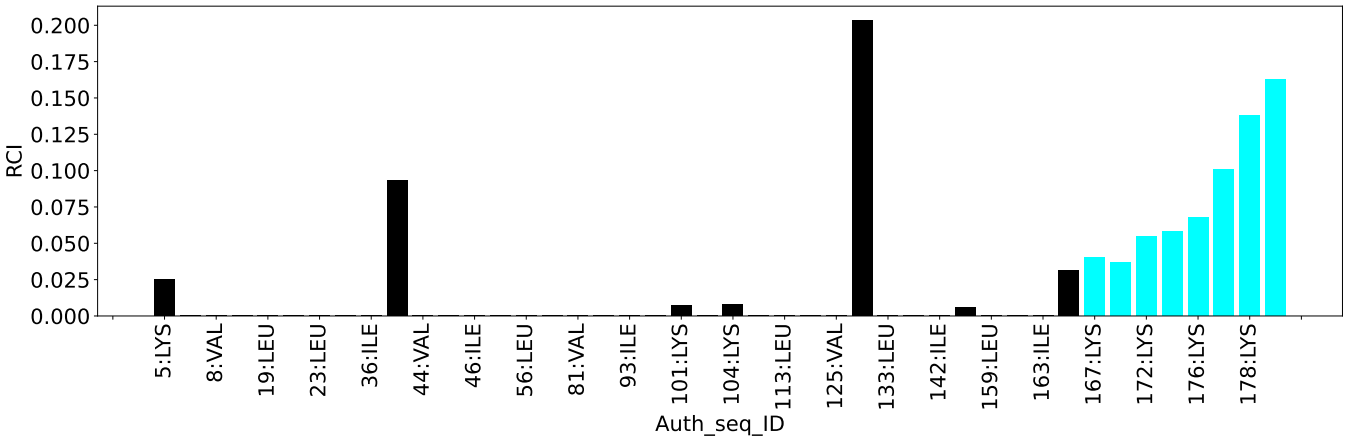
7.1.4 Statistically unusual chemical shifts (i)

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.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:



8 NMR restraints analysis

8.1 Conformationally restricting restraints

The following table provides the summary of experimentally observed NMR restraints in different categories. Restraints are classified into different categories based on the sequence separation of the atoms involved.

Description	Value
Total distance restraints	75
Intra-residue (|i-j|=0)	1
Sequential (|i-j|=1)	0
Medium range (|i-j|>1 and |i-j|<5)	0
Long range (|i-j|≥5)	21
Inter-chain	43
Hydrogen bond restraints	0
Disulfide bond restraints	10
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	0.1
Number of long range restraints per residue[1]	0.0

[1]Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations

This section provides the overview of the restraint violations analysis. The violations are binned as small, medium and large violations based on its absolute value. Average number of violations per model is calculated by dividing the total number of violations in each bin by the size of the ensemble.

8.2.1 Average number of distance violations per model

Distance violations less than 0.1 Å are not included in the calculation.

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	0.2	0.2
0.2-0.5 (Medium)	2.1	0.49
>0.5 (Large)	60.2	25.4

8.2.2 Average number of dihedral-angle violations per model (i)

Dihedral-angle violations less than 1° are not included in the calculation.There are no dihedral-angle violations

9 Distance violation analysis (i)

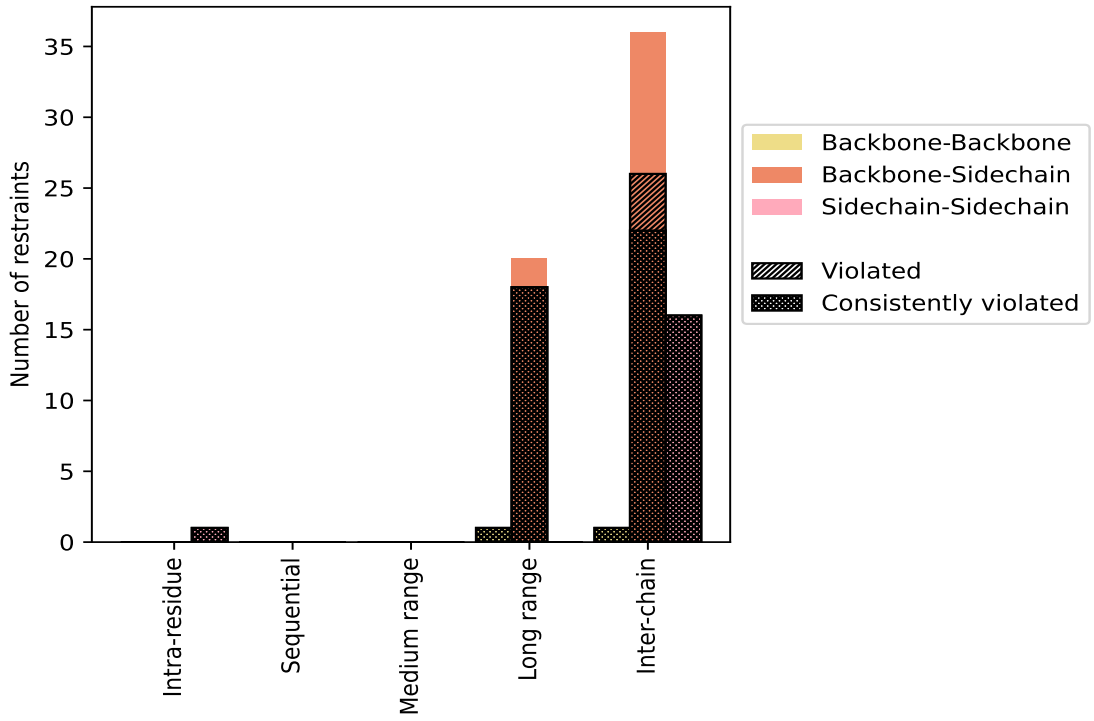
9.1 Summary of distance violations (i)

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	1	1.3	1	100.0	1.3	1	100.0	1.3
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	1	1.3	1	100.0	1.3	1	100.0	1.3
Sequential (|i-j|=1)	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	21	28.0	19	90.5	25.3	19	90.5	25.3
Backbone-Backbone	1	1.3	1	100.0	1.3	1	100.0	1.3
Backbone-Sidechain	20	26.7	18	90.0	24.0	18	90.0	24.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Inter-chain	43	57.3	33	76.7	44.0	29	67.4	38.7
Backbone-Backbone	1	1.3	1	100.0	1.3	1	100.0	1.3
Backbone-Sidechain	36	48.0	26	72.2	34.7	22	61.1	29.3
Sidechain-Sidechain	6	8.0	6	100.0	8.0	6	100.0	8.0
Hydrogen bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Disulfide bond	10	13.3	10	100.0	13.3	10	100.0	13.3
Total	75	100.0	63	84.0	84.0	59	78.7	78.7
Backbone-Backbone	2	2.7	2	100.0	2.7	2	100.0	2.7
Backbone-Sidechain	56	74.7	44	78.6	58.7	40	71.4	53.3
Sidechain-Sidechain	17	22.7	17	100.0	22.7	17	100.0	22.7

[1] percentage calculated with respect to the total number of distance restraints, [2] percentage calculated with respect to the number of restraints in a particular restraint category, [3] violated in at least one model, [4] violated in all the models

9.1.1 Bar chart : Distribution of distance restraints and violations



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

9.2 Distance violation statistics for each model

The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	1	0	0	19	43	63	10.11	17.19	4.59	11.28
2	1	0	0	19	42	62	10.12	17.3	4.28	11.38
3	1	0	0	19	43	63	10.09	18.59	4.63	11.69
4	1	0	0	19	43	63	10.27	18.82	4.72	12.07
5	1	0	0	19	43	63	9.98	16.89	4.59	10.83
6	1	0	0	19	42	62	10.44	18.3	4.39	11.78
7	1	0	0	19	43	63	10.2	17.25	4.85	12.06
8	1	0	0	19	43	63	10.1	20.35	4.46	11.08
9	1	0	0	19	42	62	10.36	17.07	4.17	11.2
10	1	0	0	19	43	63	10.08	19.49	4.48	10.86
11	1	0	0	19	43	63	10.41	22.7	4.93	11.08

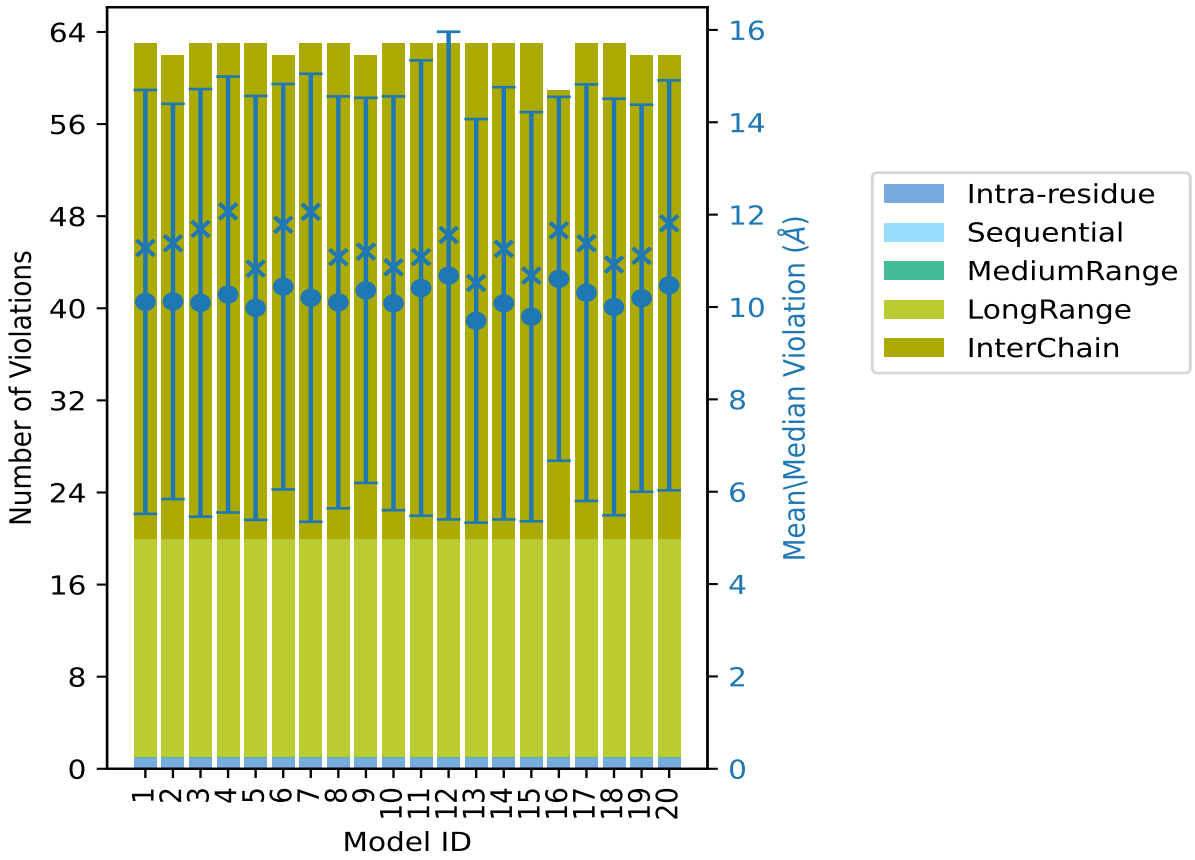
Continued on next page...

Continued from previous page...

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
12	1	0	0	19	43	63	10.68	25.4	5.28	11.56
13	1	0	0	19	43	63	9.7	19.54	4.37	10.52
14	1	0	0	19	43	63	10.08	17.85	4.68	11.26
15	1	0	0	19	43	63	9.79	16.86	4.43	10.68
16	1	0	0	19	39	59	10.61	16.27	3.94	11.66
17	1	0	0	19	43	63	10.31	18.87	4.51	11.38
18	1	0	0	19	43	63	10.0	19.04	4.51	10.92
19	1	0	0	19	42	62	10.19	17.46	4.19	11.11
20	1	0	0	19	42	62	10.47	17.21	4.44	11.81

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6]Standard deviation

9.2.1 Bar graph : Distance Violation statistics for each model



The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

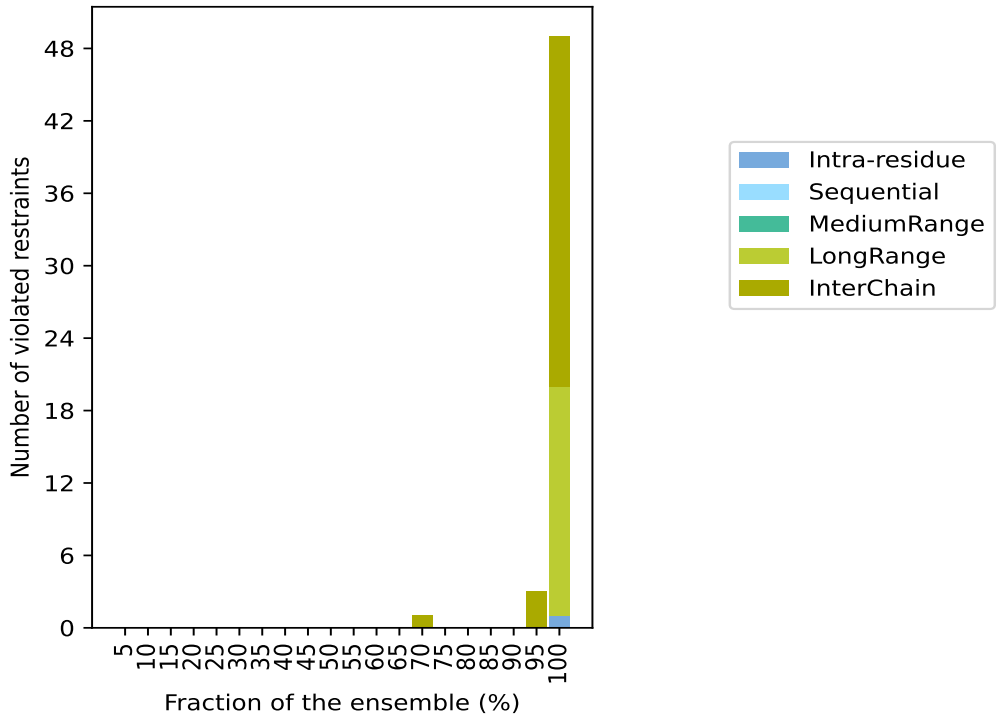
9.3 Distance violation statistics for the ensemble (i)

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 12(IR:0, SQ:0, MR:0, LR:2, IC:10) restraints are not violated in the ensemble.

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
0	0	0	0	0	0	1	5.0
0	0	0	0	0	0	2	10.0
0	0	0	0	0	0	3	15.0
0	0	0	0	0	0	4	20.0
0	0	0	0	0	0	5	25.0
0	0	0	0	0	0	6	30.0
0	0	0	0	0	0	7	35.0
0	0	0	0	0	0	8	40.0
0	0	0	0	0	0	9	45.0
0	0	0	0	0	0	10	50.0
0	0	0	0	0	0	11	55.0
0	0	0	0	0	0	12	60.0
0	0	0	0	0	0	13	65.0
0	0	0	0	1	1	14	70.0
0	0	0	0	0	0	15	75.0
0	0	0	0	0	0	16	80.0
0	0	0	0	0	0	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	3	3	19	95.0
1	0	0	19	29	49	20	100.0

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6] Number of models with violations

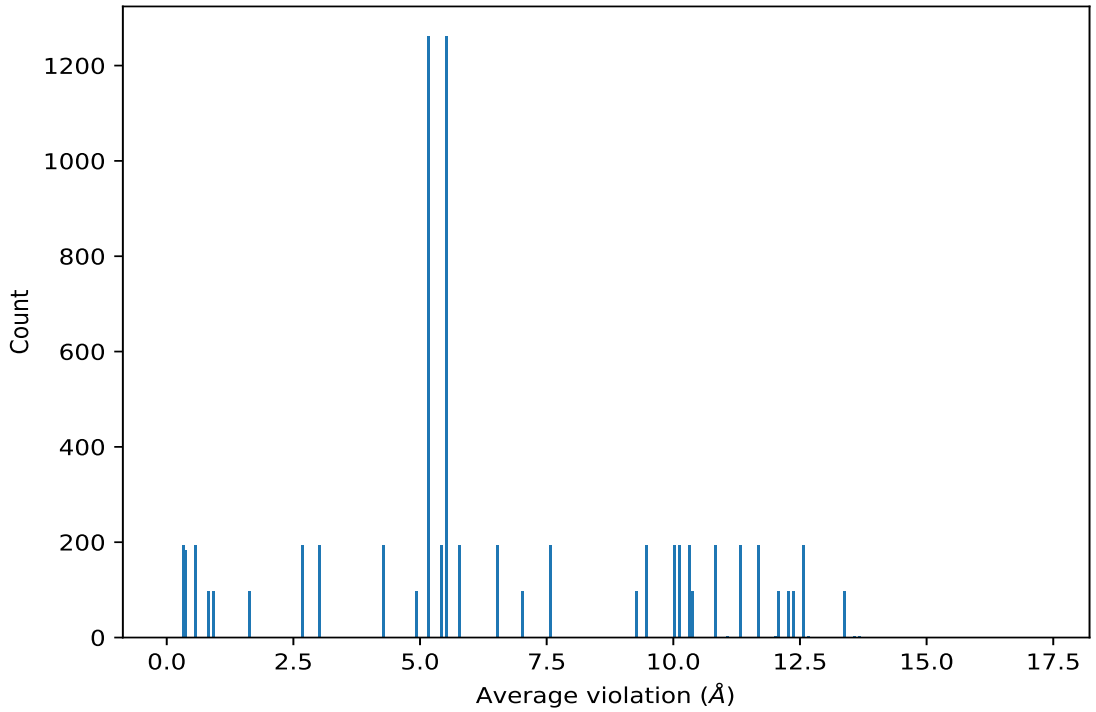
9.3.1 Bar graph : Distance violation statistics for the ensemble



9.4 Most violated distance restraints in the ensemble

9.4.1 Histogram : Distribution of mean distance violations

The following histogram shows the distribution of the average value of the violation. The average is calculated for each restraint that is violated in more than one model over all the violated models in the ensemble



9.4.2 Table: Most violated distance restraints (i)

The following table provides the mean and the standard deviation of the violations for the 10 worst performing restraints, sorted by number of violated models and the mean violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

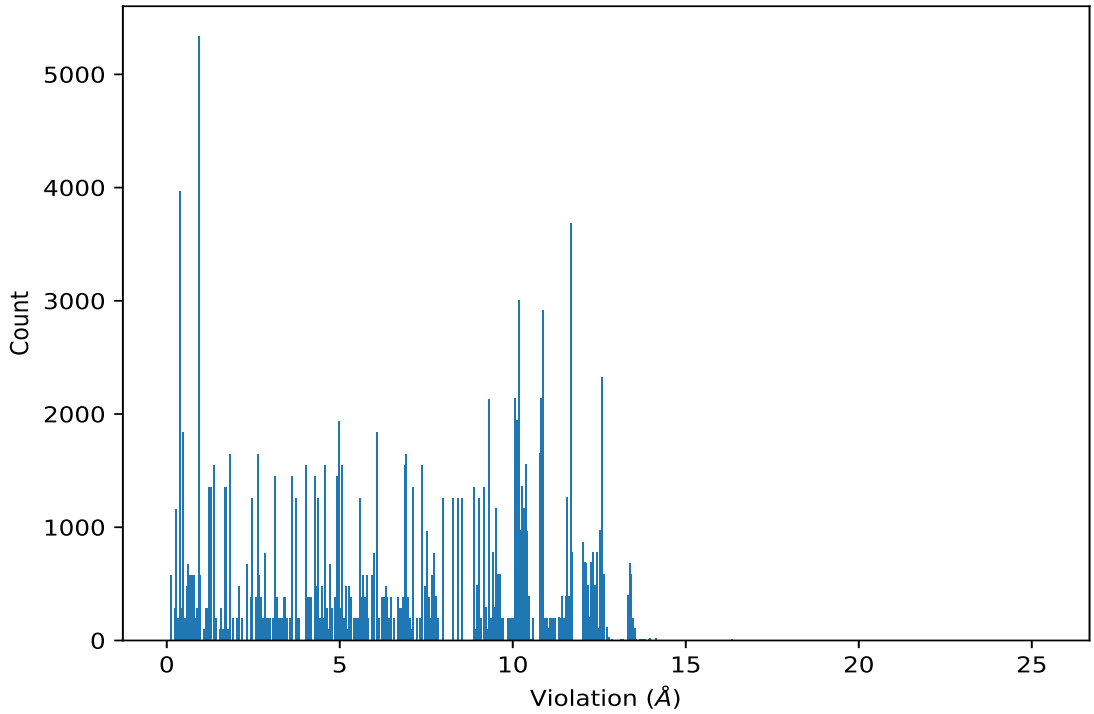
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,32)	1:B:169:LYS:CG	1:A:179:LYS:N	20	17.26	2.48	17.1
(1,9)	1:A:169:LYS:CG	1:B:128:LYS:N	20	16.67	0.99	16.31
(1,25)	1:B:169:LYS:CG	1:A:128:LYS:N	20	16.65	0.81	16.32
(1,16)	1:A:169:LYS:CG	1:B:179:LYS:N	20	16.54	2.4	17.23
(1,31)	1:B:169:LYS:CG	1:A:178:LYS:N	20	15.53	2.09	15.84
(1,30)	1:B:169:LYS:CG	1:A:177:LYS:N	20	15.1	1.57	15.08
(1,15)	1:A:169:LYS:CG	1:B:178:LYS:N	20	14.85	1.82	15.0
(1,29)	1:B:169:LYS:CG	1:A:176:LYS:N	20	13.95	1.66	13.64
(1,19)	1:B:118:CYS:SG	1:A:46:ILE:CD1	20	13.85	0.17	13.92
(1,14)	1:A:169:LYS:CG	1:B:177:LYS:N	20	13.81	1.68	13.73

[1]Number of violated models, [2]Standard deviation

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.



9.5.2 Table : All distance violations (i)

The following table provides the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,32)	1:B:169:LYS:CG	1:A:179:LYS:N	12	25.4
(1,16)	1:A:169:LYS:CG	1:B:179:LYS:N	11	22.7
(1,31)	1:B:169:LYS:CG	1:A:178:LYS:N	12	22.4
(1,16)	1:A:169:LYS:CG	1:B:179:LYS:N	8	20.35
(1,30)	1:B:169:LYS:CG	1:A:177:LYS:N	12	19.58
(1,32)	1:B:169:LYS:CG	1:A:179:LYS:N	13	19.54
(1,32)	1:B:169:LYS:CG	1:A:179:LYS:N	10	19.49
(1,32)	1:B:169:LYS:CG	1:A:179:LYS:N	11	19.27
(1,15)	1:A:169:LYS:CG	1:B:178:LYS:N	11	19.17
(1,32)	1:B:169:LYS:CG	1:A:179:LYS:N	18	19.04

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found