Full wwPDB NMR Structure Validation Report (i)

Aug 15, 2023 – 01:16 pm BST

PDB ID	:	8AU4
BMRB ID	:	34754
Title	:	Structural insights reveal a heterotetramer between oncogenic K-Ras4BG12V and Rgl2, a RalA/B activator
Authors	:	Tariq, M.; Ikeya, T.; Togashi, N.; Fairall, L.; Alejo, C.B.; Kamei, S.; Alonso, B.R.; Campillo, M.A.M.; Hudson, A.; Ito, Y.; Schwabe, J.; Dominguez, C.; Tanaka, K.
Deposited on	:	2022-08-25

This is a Full wwPDB NMR Structure Validation 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:

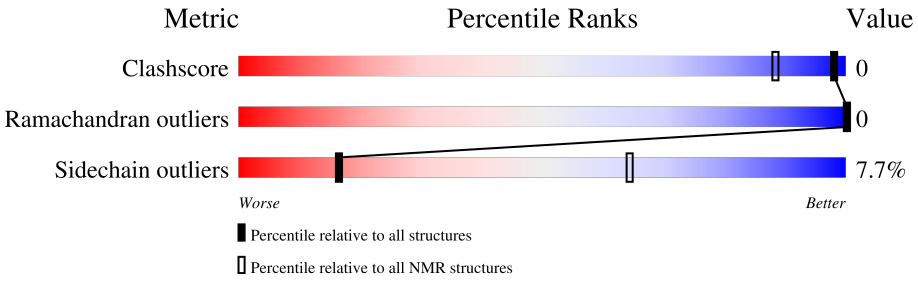
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance

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

The overall completeness of chemical shifts assignment is 91%.

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	100	72% · 26%

2 Ensemble composition and analysis

This entry contains 20 models. Model 4 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:649-A:657, A:664-A:687, A:693-A:733 (74)	0.47	4

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1539 atoms, of which 773 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Ral guanine nucleotide dissociation stimulator-like 2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	100	1539	471	773	140	151	4	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	641	SER	-	expression tag	UNP O15211
A	642	MET	-	expression tag	UNP O15211

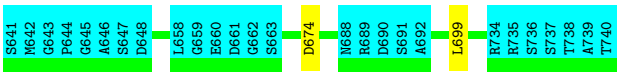
4 Residue-property plots

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: Ral guanine nucleotide dissociation stimulator-like 2

Chain A:



4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

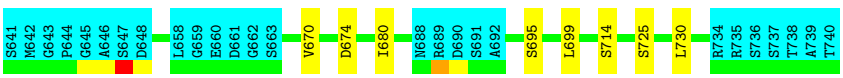
• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

Chain A:

4.2.2 Score per residue for model 2

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

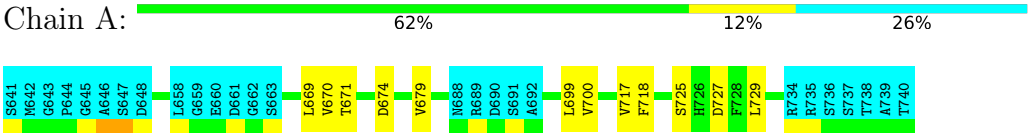
Chain A:



4.2.3 Score per residue for model 3

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

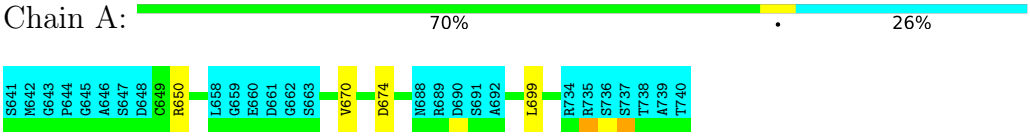
Chain A:



4.2.4 Score per residue for model 4 (medoid)

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

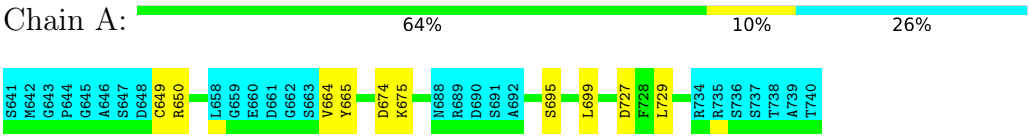
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

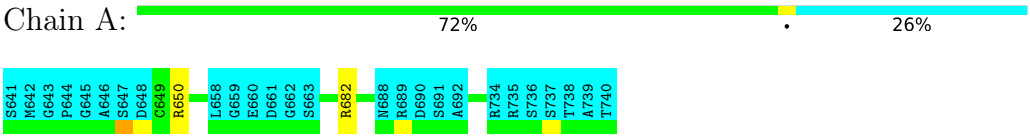
Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

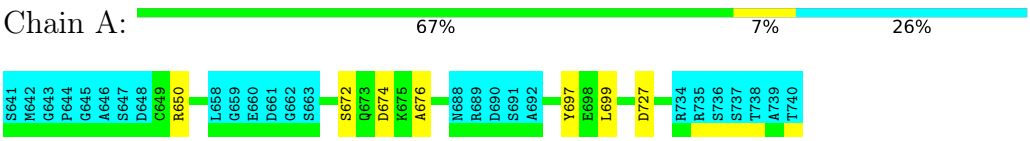
Chain A:



4.2.7 Score per residue for model 7

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

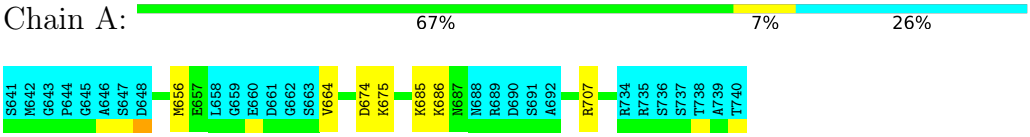
Chain A:



4.2.8 Score per residue for model 8

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

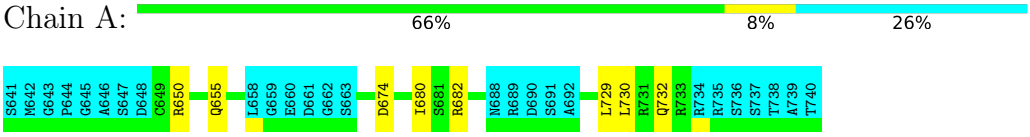
Chain A:



4.2.9 Score per residue for model 9

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

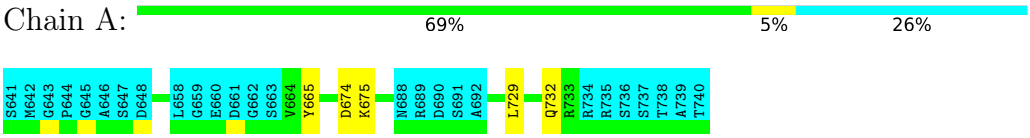
Chain A:



4.2.10 Score per residue for model 10

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

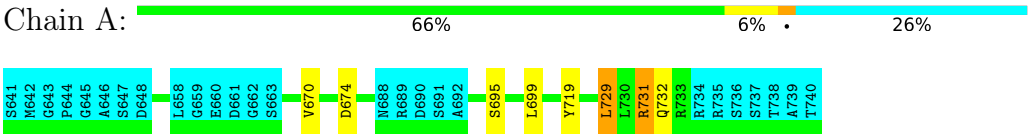
Chain A:



4.2.11 Score per residue for model 11

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

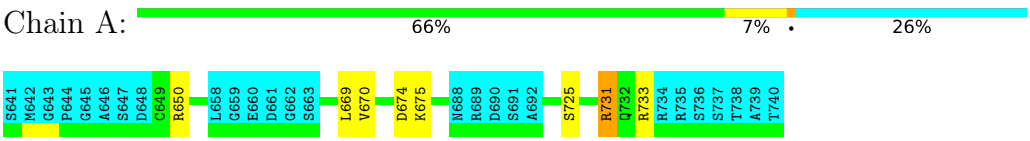
Chain A:



4.2.12 Score per residue for model 12

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

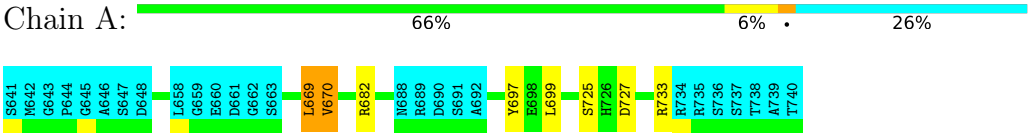
Chain A:



4.2.13 Score per residue for model 13

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

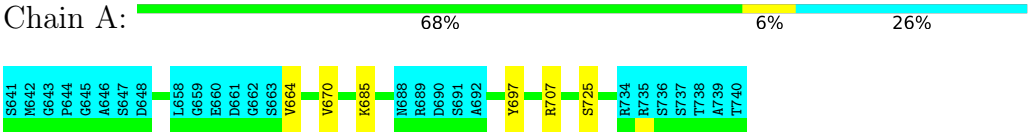
Chain A:



4.2.14 Score per residue for model 14

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

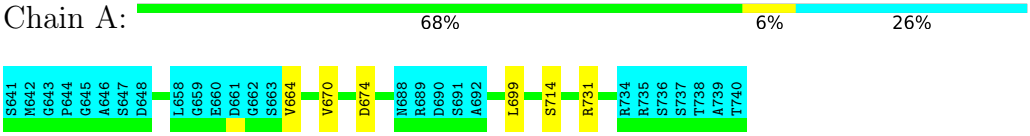
Chain A:



4.2.15 Score per residue for model 15

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

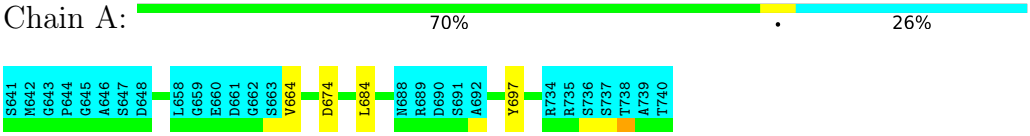
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

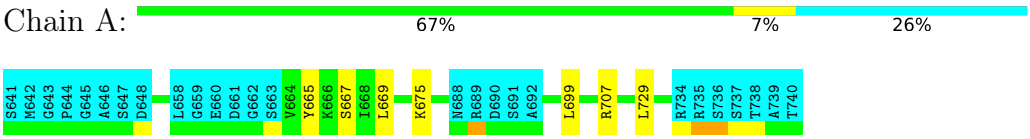
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

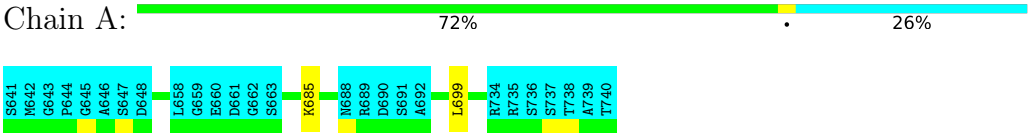
Chain A:



4.2.18 Score per residue for model 18

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

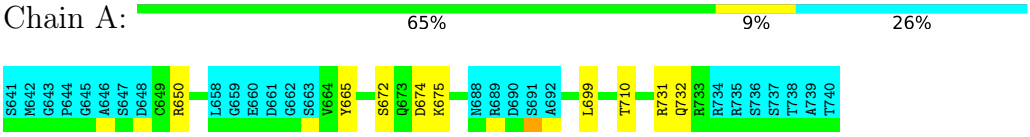
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

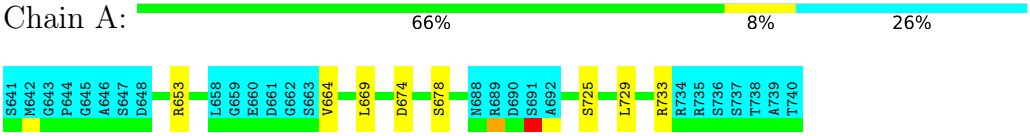
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: Ral guanine nucleotide dissociation stimulator-like 2

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *energy minimization.*

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function.*

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

Software name	Classification	Version
CYANA	structure calculation	3.99.0
OPALp	refinement	1.4

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	1188
Number of shifts mapped to atoms	1188
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	91%

6 Model quality

6.1 Standard geometry

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	0.56±0.01	0±0/597 (0.0± 0.0%)	1.03±0.04	0±1/807 (0.1± 0.1%)
All	All	0.56	0/11940 (0.0%)	1.03	9/16140 (0.1%)

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	0.8±0.8
All	All	0	15

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	674	ASP	CB-CG-OD2	-7.72	111.35	118.30	10	2
1	A	674	ASP	CB-CG-OD1	7.46	125.02	118.30	10	2
1	A	665	TYR	CB-CG-CD1	-7.46	116.53	121.00	10	4
1	A	718	PHE	CB-CG-CD1	-5.09	117.24	120.80	3	1

There are no chirality outliers.

All 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	682	ARG	Sidechain	3
1	A	731	ARG	Sidechain	3
1	A	697	TYR	Sidechain	3
1	A	707	ARG	Sidechain	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	650	ARG	Sidechain	1
1	A	719	TYR	Sidechain	1
1	A	733	ARG	Sidechain	1
1	A	653	ARG	Sidechain	1

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	588	612	612	0±1
All	All	11760	12240	12240	7

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:669:LEU:HD13	1:A:670:VAL:H	0.49	1.67	13	1
1:A:699:LEU:HD13	1:A:700:VAL:N	0.42	2.28	3	1
1:A:679:VAL:HG21	1:A:717:VAL:HG21	0.41	1.92	3	1
1:A:729:LEU:HD11	1:A:731:ARG:HD3	0.41	1.93	11	1
1:A:680:ILE:HG12	1:A:730:LEU:HD13	0.40	1.93	2	2
1:A:676:ALA:HB1	1:A:699:LEU:HB2	0.40	1.91	7	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	74/100 (74%)	71±1 (96±2%)	3±1 (4±2%)	0±0 (0±0%)	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1480/2000 (74%)	1418 (96%)	62 (4%)	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	66/85 (78%)	61±2 (92±3%)	5±2 (8±3%)	16 64
All	All	1320/1700 (78%)	1218 (92%)	102 (8%)	16 64

All 28 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	674	ASP	13
1	A	699	LEU	10
1	A	670	VAL	8
1	A	650	ARG	7
1	A	725	SER	7
1	A	729	LEU	7
1	A	664	VAL	6
1	A	675	LYS	6
1	A	669	LEU	5
1	A	727	ASP	4
1	A	732	GLN	4
1	A	695	SER	3
1	A	685	LYS	3
1	A	714	SER	2
1	A	672	SER	2
1	A	731	ARG	2
1	A	733	ARG	2
1	A	706	GLU	1
1	A	671	THR	1
1	A	649	CYS	1
1	A	697	TYR	1
1	A	656	MET	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	686	LYS	1
1	A	707	ARG	1
1	A	655	GLN	1
1	A	667	SER	1
1	A	710	THR	1
1	A	678	SER	1

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates

There are no monosaccharides in this entry.

6.6 Ligand geometry

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 91% for the well-defined parts and 89% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *starch_output*

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	1188
Number of shifts mapped to atoms	1188
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	97	-1.57 ± 0.20	Should be applied
$^{13}C_{\beta}$	93	-1.39 ± 0.23	Should be applied
$^{13}C'$	97	-0.97 ± 0.19	Should be applied
^{15}N	94	0.03 ± 0.30	None needed (< 0.5 ppm)

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 91%, i.e. 955 atoms were assigned a chemical shift out of a possible 1055. 0 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	366/366 (100%)	147/147 (100%)	148/148 (100%)	71/71 (100%)
Sidechain	545/635 (86%)	375/415 (90%)	165/192 (86%)	5/28 (18%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	44/54 (81%)	23/26 (88%)	21/27 (78%)	0/1 (0%)
Overall	955/1055 (91%)	545/588 (93%)	334/367 (91%)	76/100 (76%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 89%, i.e. 1188 atoms were assigned a chemical shift out of a possible 1341. 0 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	487/498 (98%)	199/202 (99%)	194/200 (97%)	94/96 (98%)
Sidechain	657/789 (83%)	450/513 (88%)	201/238 (84%)	6/38 (16%)
Aromatic	44/54 (81%)	23/26 (88%)	21/27 (78%)	0/1 (0%)
Overall	1188/1341 (89%)	672/741 (91%)	416/465 (89%)	100/135 (74%)

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: