



wwPDB NMR Structure Validation Summary Report ⓘ

Jun 12, 2024 – 02:55 PM EDT

PDB ID : 2L2P
BMRB ID : 17149
Title : Folding Intermediate of the Fyn SH3 A39V/N53P/V55L from NMR Relaxation Dispersion Experiments
Authors : Neudecker, P.; Robustelli, P.; Cavalli, A.; Vendruscolo, M.; Kay, L.E.
Deposited on : 2010-08-25

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 ⓘ 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 ⓘ](#)) 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) : 2.36.2

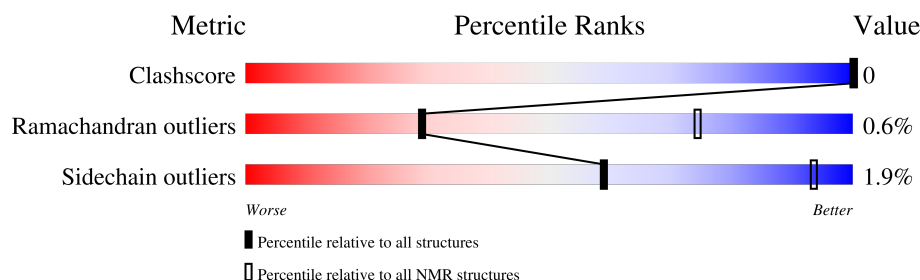
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 76%.

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	66	

2 Ensemble composition and analysis

This entry contains 10 models. Model 2 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:1-A:53 (53)	0.65	2

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	2, 3, 4, 6, 7, 8, 9
2	1, 5, 10

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 875 atoms, of which 417 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Tyrosine-protein kinase Fyn.

Mol	Chain	Residues	Atoms					Trace
1	A	56	Total	C	H	N	O	0
			875	293	417	70	95	

There are 11 discrepancies between the modelled and reference sequences:

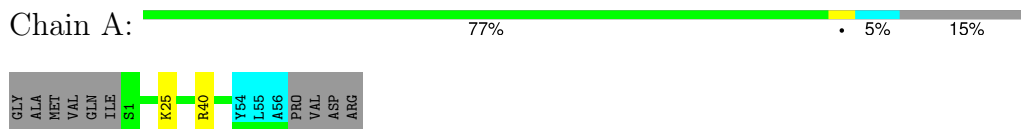
Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP Q05876
A	-4	ALA	-	expression tag	UNP Q05876
A	-3	MET	-	expression tag	UNP Q05876
A	-2	VAL	-	expression tag	UNP Q05876
A	-1	GLN	-	expression tag	UNP Q05876
A	0	ILE	-	expression tag	UNP Q05876
A	1	SER	-	expression tag	UNP Q05876
A	39	VAL	ALA	engineered mutation	UNP Q05876
A	53	PRO	ASN	engineered mutation	UNP Q05876
A	55	LEU	VAL	engineered mutation	UNP Q05876
A	60	ARG	-	expression tag	UNP Q05876

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.

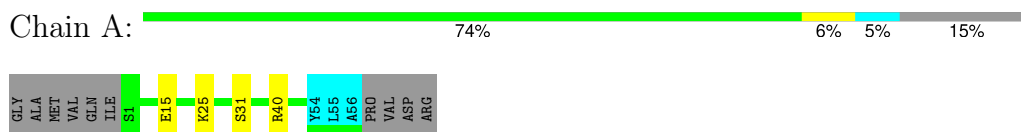
- Molecule 1: Tyrosine-protein kinase Fyn



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

The representative model is number 2. Colouring as in section 4.1 above.

- Molecule 1: Tyrosine-protein kinase Fyn



5 Refinement protocol and experimental data overview

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

Of the 6666 calculated structures, 10 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
ALMOST	structure solution	1.0.4
ALMOST	refinement	1.0.4

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	2
Total number of shifts	941
Number of shifts mapped to atoms	827
Number of unparsed shifts	0
Number of shifts with mapping errors	114
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	76%

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	0.75±0.01	0±0/446 (0.0± 0.0%)	1.07±0.01	1±1/607 (0.2± 0.2%)
All	All	0.75	0/4460 (0.0%)	1.07	11/6070 (0.2%)

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	0.4±0.5
All	All	0	4

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(°)	Ideal(°)	Models	
								Worst	Total
1	A	13	ARG	NE-CZ-NH2	-6.12	117.24	120.30	4	2
1	A	40	ARG	NE-CZ-NH1	5.78	123.19	120.30	3	7
1	A	13	ARG	NE-CZ-NH1	5.44	123.02	120.30	4	2

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	31	SER	Peptide	4

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
All	All	4330	3920	3940	-

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 -.

There are no clashes.

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	52/66 (79%)	50±1 (95±2%)	2±1 (4±2%)	0±0 (1±1%)	29	74
All	All	520/660 (79%)	496 (95%)	21 (4%)	3 (1%)	29	74

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	15	GLU	3

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	47/57 (82%)	46±1 (98±2%)	1±1 (2±2%)	59	93

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	470/570 (82%)	461 (98%)	9 (2%)	59 93

All 4 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	25	LYS	6
1	A	3	LEU	1
1	A	4	PHE	1
1	A	21	HIS	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 76% for the well-defined parts and 76% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

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	664
Number of shifts mapped to atoms	564
Number of unparsed shifts	0
Number of shifts with mapping errors	100
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	4

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 100) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	-4	ALA	HA	4.32	0.01	1
1	A	-4	ALA	HB1	1.35	0.01	1
1	A	-4	ALA	HB2	1.35	0.01	1
1	A	-4	ALA	HB3	1.35	0.01	1
1	A	-4	ALA	C	177.62	0.10	1
1	A	-4	ALA	CA	52.26	0.10	1
1	A	-4	ALA	CB	19.47	0.10	1
1	A	-3	MET	H	8.49	0.01	1
1	A	-3	MET	HA	4.45	0.01	1
1	A	-3	MET	HB2	2.0	0.01	2
1	A	-3	MET	HB3	2.0	0.01	2
1	A	-3	MET	HE1	2.05	0.01	1
1	A	-3	MET	HE2	2.05	0.01	1
1	A	-3	MET	HE3	2.05	0.01	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	-3	MET	C	176.08	0.10	1
1	A	-3	MET	CA	55.36	0.10	1
1	A	-3	MET	CE	16.99	0.10	1
1	A	-3	MET	N	120.72	0.10	1
1	A	-2	VAL	H	8.19	0.01	1
1	A	-2	VAL	HA	4.07	0.01	1
1	A	-2	VAL	HB	1.99	0.01	1
1	A	-2	VAL	HG11	0.86	0.01	1
1	A	-2	VAL	HG12	0.86	0.01	1
1	A	-2	VAL	HG13	0.86	0.01	1
1	A	-2	VAL	HG21	0.87	0.01	1
1	A	-2	VAL	HG22	0.87	0.01	1
1	A	-2	VAL	HG23	0.87	0.01	1
1	A	-2	VAL	C	175.75	0.10	1
1	A	-2	VAL	CA	62.12	0.10	1
1	A	-2	VAL	CG1	21.18	0.10	1
1	A	-2	VAL	CG2	20.62	0.10	1
1	A	-2	VAL	N	122.74	0.10	1
1	A	-1	GLN	H	8.51	0.01	1
1	A	-1	GLN	HA	4.34	0.01	1
1	A	-1	GLN	HB2	1.96	0.01	2
1	A	-1	GLN	HB3	1.96	0.01	2
1	A	-1	GLN	HE21	7.51	0.01	2
1	A	-1	GLN	HE22	6.83	0.01	2
1	A	-1	GLN	HG2	2.3	0.01	2
1	A	-1	GLN	HG3	2.3	0.01	2
1	A	-1	GLN	C	175.88	0.10	1
1	A	-1	GLN	CA	55.5	0.10	1
1	A	-1	GLN	CD	180.37	0.10	1
1	A	-1	GLN	CG	33.4	0.10	1
1	A	-1	GLN	N	124.92	0.10	1
1	A	-1	GLN	NE2	112.61	0.10	1
1	A	0	ILE	H	8.33	0.01	1
1	A	0	ILE	HA	4.18	0.01	1
1	A	0	ILE	HB	1.85	0.01	1
1	A	0	ILE	HD11	0.75	0.01	1
1	A	0	ILE	HD12	0.75	0.01	1
1	A	0	ILE	HD13	0.75	0.01	1
1	A	0	ILE	HG12	1.42	0.01	2
1	A	0	ILE	HG13	1.16	0.01	2
1	A	0	ILE	HG21	0.94	0.01	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	0	ILE	HG22	0.94	0.01	1
1	A	0	ILE	HG23	0.94	0.01	1
1	A	0	ILE	C	176.39	0.10	1
1	A	0	ILE	CA	61.26	0.10	1
1	A	0	ILE	CD1	12.95	0.10	1
1	A	0	ILE	CG2	17.64	0.10	1
1	A	0	ILE	N	123.17	0.10	1
1	A	57	PRO	HA	3.58	0.01	1
1	A	57	PRO	C	177.07	0.10	1
1	A	57	PRO	CA	62.74	0.10	1
1	A	58	VAL	H	7.99	0.01	1
1	A	58	VAL	HA	3.87	0.01	1
1	A	58	VAL	HB	1.88	0.01	1
1	A	58	VAL	HG11	0.83	0.01	1
1	A	58	VAL	HG12	0.83	0.01	1
1	A	58	VAL	HG13	0.83	0.01	1
1	A	58	VAL	HG21	0.78	0.01	1
1	A	58	VAL	HG22	0.78	0.01	1
1	A	58	VAL	HG23	0.78	0.01	1
1	A	58	VAL	C	175.67	0.10	1
1	A	58	VAL	CA	62.97	0.10	1
1	A	58	VAL	CG1	21.38	0.10	1
1	A	58	VAL	CG2	21.58	0.10	1
1	A	58	VAL	N	120.75	0.10	1
1	A	59	ASP	H	8.32	0.01	1
1	A	59	ASP	HA	4.58	0.01	1
1	A	59	ASP	HB2	2.64	0.01	1
1	A	59	ASP	HB3	2.64	0.01	1
1	A	59	ASP	C	174.72	0.10	1
1	A	59	ASP	CA	54.14	0.10	1
1	A	59	ASP	N	122.63	0.10	1
1	A	60	ARG	H	7.61	0.01	1
1	A	60	ARG	HA	4.11	0.01	1
1	A	60	ARG	HB2	1.71	0.01	2
1	A	60	ARG	HB3	1.55	0.01	2
1	A	60	ARG	HD2	3.13	0.01	1
1	A	60	ARG	HD3	3.13	0.01	1
1	A	60	ARG	HE	7.15	0.01	1
1	A	60	ARG	HG2	1.43	0.01	1
1	A	60	ARG	HG3	1.43	0.01	1
1	A	60	ARG	CA	57.03	0.10	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	60	ARG	CD	43.19	0.10	1
1	A	60	ARG	CZ	159.53	0.10	1
1	A	60	ARG	N	124.96	0.10	1
1	A	60	ARG	NE	84.66	0.10	1

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	65	0.19 ± 0.19	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	11	—	None (insufficient data)
$^{13}\text{C}'$	61	-0.01 ± 0.15	None needed (< 0.5 ppm)
^{15}N	61	0.54 ± 0.38	None needed (imprecise)

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 76%, i.e. 529 atoms were assigned a chemical shift out of a possible 699. 0 out of 6 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	263/265 (99%)	108/108 (100%)	104/106 (98%)	51/51 (100%)
Sidechain	240/346 (69%)	199/222 (90%)	37/114 (32%)	4/10 (40%)
Aromatic	26/88 (30%)	22/43 (51%)	2/42 (5%)	2/3 (67%)
Overall	529/699 (76%)	329/373 (88%)	143/262 (55%)	57/64 (89%)

7.1.4 Statistically unusual chemical shifts ⓘ

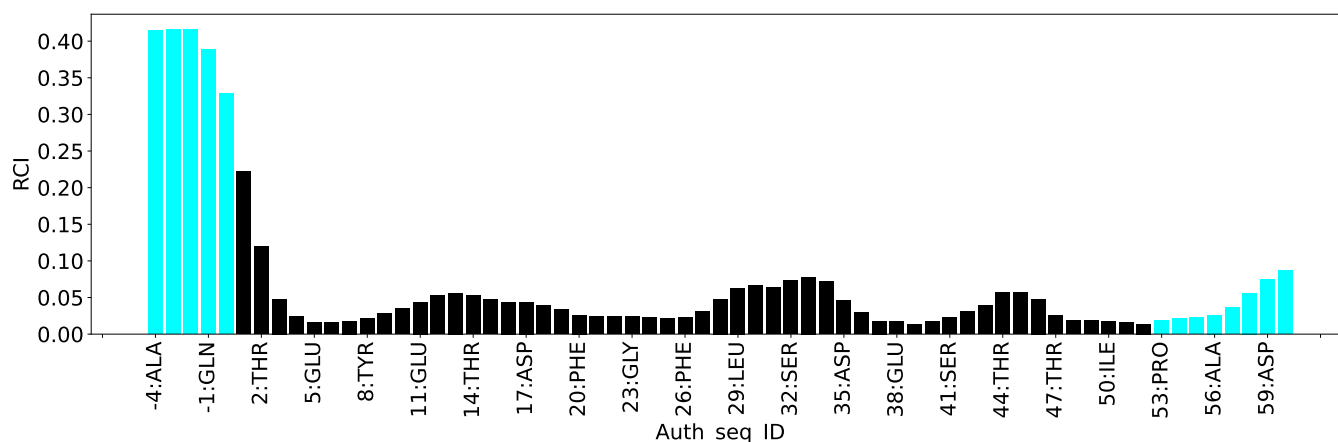
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	52	SER	HB3	1.59	2.49 – 5.20	-8.3
1	A	52	SER	HB2	1.97	2.61 – 5.13	-7.5
1	A	31	SER	HB3	2.19	2.49 – 5.20	-6.1
1	A	10	TYR	HB3	0.56	0.93 – 4.76	-6.0

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:



7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_2*

7.2.1 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	277
Number of shifts mapped to atoms	263
Number of unparsed shifts	0
Number of shifts with mapping errors	14
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 14) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
2	A	57	PRO	HA	4.38	0.10	1
2	A	57	PRO	C	177.03	0.11	1
2	A	57	PRO	CA	63.26	0.10	1
2	A	58	VAL	H	8.11	0.01	1
2	A	58	VAL	HA	4.12	0.10	1
2	A	58	VAL	C	175.89	0.10	1
2	A	58	VAL	CG1	21.38	0.10	1
2	A	58	VAL	CG2	20.24	0.10	1
2	A	58	VAL	N	119.53	0.10	1
2	A	59	ASP	H	8.32	0.05	1
2	A	59	ASP	HA	4.58	0.10	1
2	A	59	ASP	C	175.07	0.10	1
2	A	59	ASP	CA	54.14	0.28	1
2	A	59	ASP	N	122.63	1.56	1

7.2.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, <i>ppm</i>	Suggested action
$^{13}\text{C}_\alpha$	54	0.07 ± 0.17	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	3	—	None (insufficient data)
$^{13}\text{C}'$	52	-0.21 ± 0.23	None needed (< 0.5 ppm)
^{15}N	51	0.14 ± 0.83	None needed (< 0.5 ppm)

7.2.3 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 35%, i.e. 247 atoms were assigned a chemical shift out of a possible 699. 0 out of 6 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	222/265 (84%)	80/108 (74%)	96/106 (91%)	46/51 (90%)
Sidechain	21/346 (6%)	0/222 (0%)	21/114 (18%)	0/10 (0%)
Aromatic	4/88 (5%)	2/43 (5%)	0/42 (0%)	2/3 (67%)
Overall	247/699 (35%)	82/373 (22%)	117/262 (45%)	48/64 (75%)

7.2.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.2.5 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:

