



wwPDB NMR Structure Validation Summary Report

Apr 26, 2016 – 10:37 PM BST

PDB ID : 2K79
Title : Solution Structure of the binary complex between the SH3 and SH2 domain of interleukin-2 tyrosine kinase
Authors : Andreotti, A.H.; Severin, A.J.; Fulton, D.B.
Deposited on : 2008-08-08

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
<http://wwpdb.org/validation/2016/NMRValidationReportHelp>
with specific help available everywhere you see the  symbol.

The following versions of software and data (see [references](#) ) were used in the production of this report:

Cyrange : Kirchner and Güntert (2011)
NmrClust : Kelley et al. (1996)
MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
ShiftChecker : rb-20027457
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027457

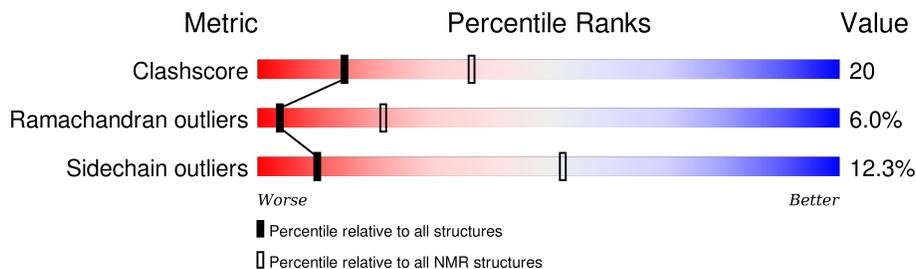
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 81%.

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	63	
2	B	110	

2 Ensemble composition and analysis

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2749 atoms, of which 1345 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called SH3 domain of Tyrosine-protein kinase ITK/TSK.

Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
1	A	63	993	328	473	82	109	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	169	GLY	-	EXPRESSION TAG	UNP Q03526
A	170	SER	-	EXPRESSION TAG	UNP Q03526

- Molecule 2 is a protein called SH2 domain of Tyrosine-protein kinase ITK/TSK.

Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
2	B	108	1756	565	872	150	166	3	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	230	GLY	-	EXPRESSION TAG	UNP Q03526
B	231	SER	-	EXPRESSION TAG	UNP Q03526
B	339	GLY	-	EXPRESSION TAG	UNP Q03526

4 Residue-property plots [i](#)

These plots are provided for all protein, RNA and DNA 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.

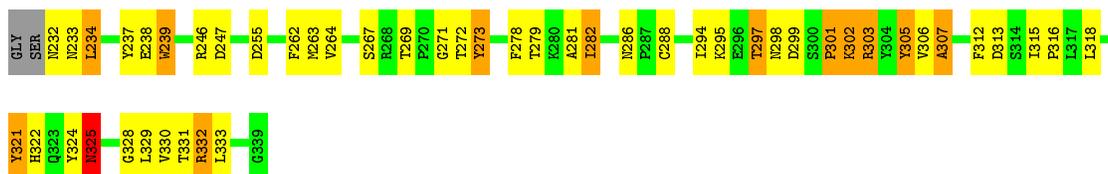
- Molecule 1: SH3 domain of Tyrosine-protein kinase ITK/TSK

Chain A:  62% 37%



- Molecule 2: SH2 domain of Tyrosine-protein kinase ITK/TSK

Chain B:  54% 33% 11%



5 Refinement protocol and experimental data overview

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

Of the ? calculated structures, 1 were deposited, based on the following criterion: ?.

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

Software name	Classification	Version
xplor-nih	refinement	2.19

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)	BMRB entry 15912
Number of chemical shift lists	4
Total number of shifts	3936
Number of shifts mapped to atoms	3782
Number of unparsed shifts	0
Number of shifts with mapping errors	154
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	81%

No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality [i](#)

6.1 Standard geometry [i](#)

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	520	473	470	20
2	B	884	872	870	39
All	All	1404	1345	1340	54

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 54 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:222:PRO:O	1:A:226:LEU:HD12	0.83	1.73
1:A:208:TRP:HE1	2:B:328:GLY:C	0.74	1.85
2:B:306:VAL:HG23	2:B:306:VAL:O	0.70	1.85
1:A:184:THR:OG1	1:A:191:ALA:N	0.56	2.38
2:B:278:PHE:CD1	2:B:279:THR:N	0.56	2.73

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	61/63 (97%)	53 (87%)	8 (13%)	0 (0%)	100	100
2	B	106/110 (96%)	81 (76%)	15 (14%)	10 (9%)	2	11
All	All	167/173 (97%)	134 (80%)	23 (14%)	10 (6%)	4	21

5 of 10 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
2	B	303	ARG
2	B	298	ASN
2	B	332	ARG
2	B	282	ILE
2	B	325	ASN

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	57/57 (100%)	51 (89%)	6 (11%)	13	57
2	B	97/98 (99%)	84 (87%)	13 (13%)	9	50
All	All	154/155 (99%)	135 (88%)	19 (12%)	10	52

5 of 19 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	214	LYS
2	B	267	SER
2	B	321	TYR
2	B	329	LEU
1	A	170	SER

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 carbohydrates 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](#)

The completeness of assignment taking into account all chemical shift lists is 81% for the well-defined parts and 81% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 15912

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 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	1318
Number of shifts mapped to atoms	1249
Number of unparsed shifts	0
Number of shifts with mapping errors	69
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- Residue not found in structure. First 5 (of 69) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
B	112	PRO	C	174.7	0.3	1
B	111	SER	CB	60.91	0.3	1
B	112	PRO	CA	61.13	0.3	1
B	114	ILE	H	7.72	0.02	1
B	111	SER	N	116.28	0.3	1

7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	116	2.40 \pm 0.22	Should be applied
$^{13}\text{C}_\beta$	107	2.62 \pm 0.23	Should be applied
$^{13}\text{C}'$	104	2.98 \pm 0.19	Should be applied

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction \pm precision, ppm	Suggested action
¹⁵ N	107	0.57 \pm 0.32	None needed (imprecise)

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 52%, i.e. 1115 atoms were assigned a chemical shift out of a possible 2150. 13 out of 25 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	516/837 (62%)	208/333 (62%)	207/342 (61%)	101/162 (62%)
Sidechain	518/1086 (48%)	290/640 (45%)	218/398 (55%)	10/48 (21%)
Aromatic	81/227 (36%)	58/117 (50%)	22/101 (22%)	1/9 (11%)
Overall	1115/2150 (52%)	556/1090 (51%)	447/841 (53%)	112/219 (51%)

7.1.4 Statistically unusual chemical shifts [i](#)

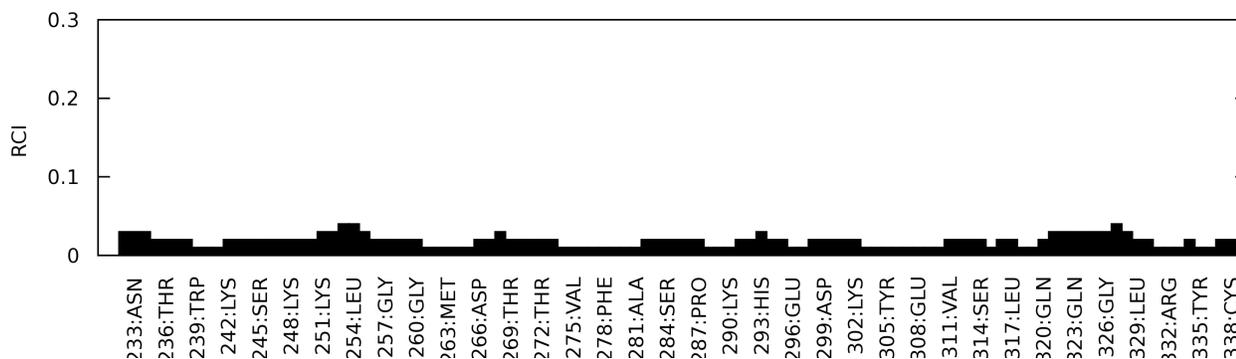
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
2	B	303	ARG	NE	119.27	92.63 – 76.73	21.8
2	B	265	ARG	NE	118.55	92.63 – 76.73	21.3
2	B	333	LEU	HB3	-1.20	3.34 – -0.26	-7.6
2	B	237	TYR	HB3	0.47	4.75 – 0.95	-6.3
2	B	287	PRO	CG	21.76	32.66 – 21.76	-5.0

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.

Random coil index (RCI) for chain B:



7.2 Chemical shift list 2

File name: BMRB entry 15912

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	703
Number of shifts mapped to atoms	693
Number of unparsed shifts	0
Number of shifts with mapping errors	10
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.

- Residue not found in structure. First 5 (of 10) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	64	ASN	CA	52.43	0.3	1
A	64	ASN	HA	4.32	0.02	1
A	64	ASN	N	123.77	0.3	1
A	64	ASN	HD22	6.77	0.02	1
A	64	ASN	HB2	2.63	0.02	2

7.2.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	64	2.90 ± 0.14	Should be applied
$^{13}\text{C}_\beta$	61	2.27 ± 0.30	Should be applied
$^{13}\text{C}'$	58	2.98 ± 0.24	Should be applied
^{15}N	59	0.34 ± 0.60	None needed (< 0.5 ppm)

7.2.3 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 29%, i.e. 629 atoms were assigned a chemical shift out of a possible 2150. 10 out of 25 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	299/837 (36%)	120/333 (36%)	121/342 (35%)	58/162 (36%)
Sidechain	296/1086 (27%)	167/640 (26%)	122/398 (31%)	7/48 (15%)
Aromatic	34/227 (15%)	18/117 (15%)	14/101 (14%)	2/9 (22%)
Overall	629/2150 (29%)	305/1090 (28%)	257/841 (31%)	67/219 (31%)

7.2.4 Statistically unusual chemical shifts [i](#)

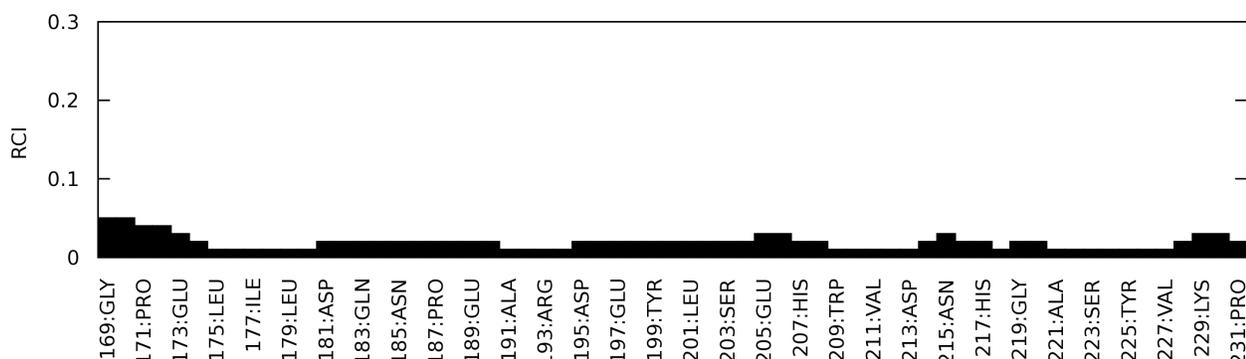
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	223	SER	HB3	1.61	5.25 – 2.45	-8.0
1	A	223	SER	HB2	1.80	5.18 – 2.58	-8.0
1	A	203	SER	HB3	2.16	5.25 – 2.45	-6.0
1	A	222	PRO	HG2	0.32	3.48 – 0.38	-5.2

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.

Random coil index (RCI) for chain A:



7.3 Chemical shift list 3

File name: BMRB entry 15912

Chemical shift list name: *assigned_chem_shift_list_3*

7.3.1 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	1212
Number of shifts mapped to atoms	1147
Number of unparsed shifts	0
Number of shifts with mapping errors	65
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	6

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- Residue not found in structure. First 5 (of 65) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
B	111	SER	CB	60.91	0.3	1
B	112	PRO	CA	61.13	0.3	1
B	114	ILE	H	7.72	0.02	1
B	111	SER	N	116.29	0.3	1
B	114	ILE	HG21	0.66	0.02	1

7.3.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	116	2.62 ± 0.32	Should be applied
$^{13}\text{C}_\beta$	107	2.50 ± 0.25	Should be applied
$^{13}\text{C}'$	0	—	—
^{15}N	107	0.41 ± 0.39	None needed (< 0.5 ppm)

7.3.3 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 47%, i.e. 1016 atoms were assigned a chemical shift out of a possible 2150. 13 out of 25 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	417/837 (50%)	208/333 (62%)	108/342 (32%)	101/162 (62%)
Sidechain	518/1086 (48%)	290/640 (45%)	218/398 (55%)	10/48 (21%)
Aromatic	81/227 (36%)	58/117 (50%)	22/101 (22%)	1/9 (11%)
Overall	1016/2150 (47%)	556/1090 (51%)	348/841 (41%)	112/219 (51%)

7.3.4 Statistically unusual chemical shifts [i](#)

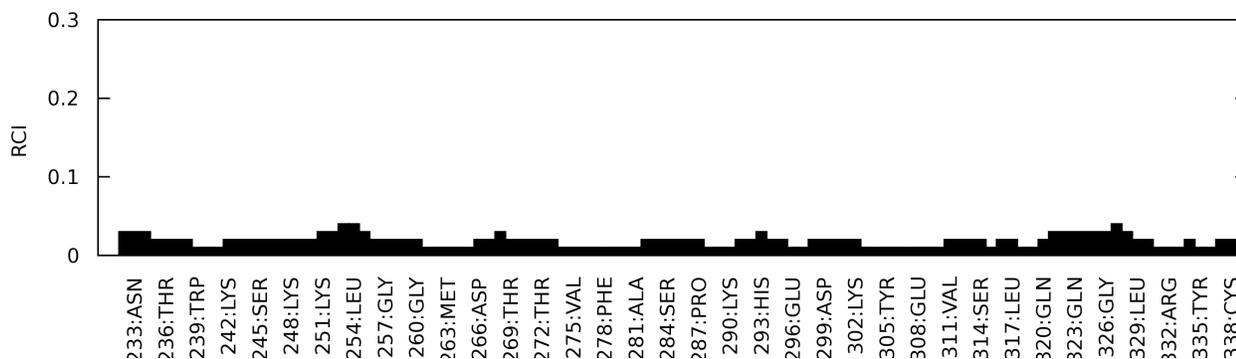
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
2	B	303	ARG	NE	119.29	92.63 – 76.73	21.8
2	B	265	ARG	NE	118.52	92.63 – 76.73	21.3
2	B	333	LEU	HB3	-1.20	3.34 – -0.26	-7.6
2	B	239	TRP	CD1	114.38	136.18 – 116.78	-6.2
2	B	237	TYR	HB3	0.53	4.75 – 0.95	-6.1
2	B	287	PRO	CG	21.76	32.66 – 21.76	-5.0

7.3.5 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.

Random coil index (RCI) for chain B:



7.4 Chemical shift list 4

File name: BMRB entry 15912

Chemical shift list name: *assigned_chem_shift_list_4*

7.4.1 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	703
Number of shifts mapped to atoms	693
Number of unparsed shifts	0
Number of shifts with mapping errors	10
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- Residue not found in structure. First 5 (of 10) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	64	ASN	CA	52.43	0.3	1
A	64	ASN	HA	4.32	0.02	1
A	64	ASN	N	123.77	0.3	1
A	64	ASN	HD22	6.77	0.02	1
A	64	ASN	HB2	2.63	0.02	2

7.4.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	64	2.89 ± 0.19	Should be applied
$^{13}\text{C}_\beta$	61	2.26 ± 0.16	Should be applied
$^{13}\text{C}'$	58	3.01 ± 0.16	Should be applied
^{15}N	59	0.32 ± 0.50	None needed (< 0.5 ppm)

7.4.3 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 29%, i.e. 629 atoms were assigned a chemical shift out of a possible 2150. 10 out of 25 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	299/837 (36%)	120/333 (36%)	121/342 (35%)	58/162 (36%)
Sidechain	296/1086 (27%)	167/640 (26%)	122/398 (31%)	7/48 (15%)
Aromatic	34/227 (15%)	18/117 (15%)	14/101 (14%)	2/9 (22%)
Overall	629/2150 (29%)	305/1090 (28%)	257/841 (31%)	67/219 (31%)

7.4.4 Statistically unusual chemical shifts [i](#)

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.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	223	SER	HB2	1.76	5.18 – 2.58	-8.2
1	A	223	SER	HB3	1.58	5.25 – 2.45	-8.1
1	A	203	SER	HB3	2.15	5.25 – 2.45	-6.1
1	A	222	PRO	HG2	0.21	3.48 – 0.38	-5.5
1	A	222	PRO	HG3	0.18	3.56 – 0.26	-5.2

7.4.5 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.

Random coil index (RCI) for chain A:

