wwPDB NMR Structure Validation Summary Report (i)

Apr 27, 2016 – 05:32 AM BST

PDB ID	:	3EZA
Title	:	COMPLEX OF THE AMINO TERMINAL DOMAIN OF ENZYME I AND THE HISTIDINE-CONTAINING PHOSPHOCARRIER PROTEIN HPR FROM ESCHERICHIA COLI NMR, RESTRAINED REGULARIZED MEAN STRUCTURE
Authors	:	Clore, G.M.; Garrett, D.S.; Gronenborn, A.M.
Deposited on	:	1998-11-03

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 (i) symbol.

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

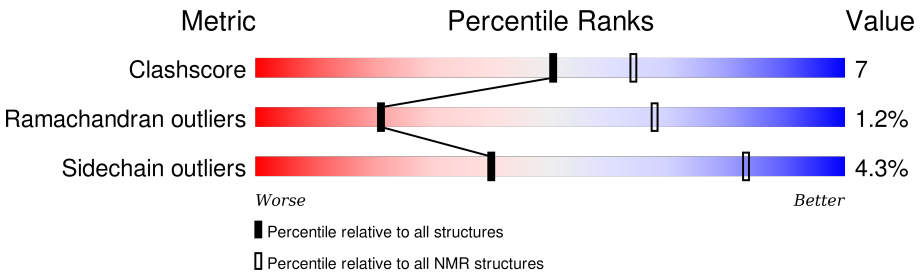
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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 80%.

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	249	87% 11% •
2	B	85	81% 18% •

2 Ensemble composition and analysis (i)

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 5147 atoms, of which 2600 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called PHOSPHOTRANSFERASE SYSTEM, ENZYME I.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	249	3854	1192	1947	322	388	5	0

- Molecule 2 is a protein called HISTIDINE-CONTAINING PHOSPHOCARRIER PROTEIN HPR.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	85	1293	401	653	107	130	2	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	63	ILE	LEU	CONFLICT	UNP P0AA04

4 Residue-property plots (i)

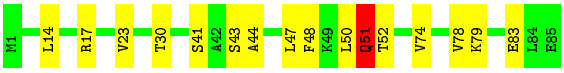
These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain 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: PHOSPHOTRANSFERASE SYSTEM, ENZYME I

Chain A:

- Molecule 2: HISTIDINE-CONTAINING PHOSPHOCARRIER PROTEIN HPR

Chain B:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *SIMULATED ANNEALING*.

Of the 40 calculated structures, 1 were deposited, based on the following criterion: *REGULARIZED MEAN STRUCTURE*.

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

Software name	Classification	Version
CNS	refinement	
CNS (SEE ABOVE)	structure solution	ABOVE)

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 4264
Number of chemical shift lists	4
Total number of shifts	5105
Number of shifts mapped to atoms	4925
Number of unparsed shifts	0
Number of shifts with mapping errors	180
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	80%

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)

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	1.15	3/1922 (0.2%)	2.03	6/2592 (0.2%)
2	B	1.13	0/647 (0.0%)	0.84	0/871 (0.0%)
All	All	1.15	3/2569 (0.1%)	1.81	6/3463 (0.2%)

All bond outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	97	HIS	CG-CD2	8.12	1.49	1.35
1	A	97	HIS	CB-CG	6.95	1.62	1.50
1	A	97	HIS	CE1-NE2	5.05	1.44	1.32

5 of 6 angle outliers are listed below. They are sorted according to the Z-score.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	97	HIS	ND1-CG-CD2	-54.98	29.02	106.00
1	A	97	HIS	CG-ND1-CE1	-51.11	36.65	108.20
1	A	97	HIS	CG-CD2-NE2	-36.71	39.44	109.20
1	A	97	HIS	ND1-CE1-NE2	-34.50	34.00	109.90
1	A	97	HIS	CE1-NE2-CD2	-26.25	40.98	106.60

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	1907	1947	1947	28

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes
2	B	640	653	653	16
All	All	2547	2600	2600	37

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 7.

5 of 37 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:79:LEU:HD11	2:B:48:PHE:CE2	0.74	2.17
1:A:10:GLY:O	1:A:222:ALA:HB3	0.64	1.92
1:A:182:ASP:N	1:A:182:ASP:OD1	0.58	2.35
1:A:79:LEU:HD21	2:B:48:PHE:CE1	0.56	2.35
1:A:122:TYR:OH	1:A:126:ARG:NH1	0.56	2.39

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	247/249 (99%)	233 (94%)	11 (4%)	3 (1%)	21 68
2	B	83/85 (98%)	76 (92%)	6 (7%)	1 (1%)	21 68
All	All	330/334 (99%)	309 (94%)	17 (5%)	4 (1%)	21 68

All 4 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	183	ALA
2	B	51	GLN
1	A	148	ASP
1	A	6	LEU

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	206/206 (100%)	198 (96%)	8 (4%)	43	85
2	B	70/70 (100%)	66 (94%)	4 (6%)	30	75
All	All	276/276 (100%)	264 (96%)	12 (4%)	40	83

5 of 12 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	215	ASP
2	B	43	SER
1	A	216	ASP
1	A	186	ARG
1	A	232	THR

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 80% for the well-defined parts and 80% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 4264

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	2540
Number of shifts mapped to atoms	2420
Number of unparsed shifts	0
Number of shifts with mapping errors	120
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

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

- Residue not found in structure. First 5 (of 120) occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	257	LYS	HB2	1.86	0.03	2
A	255	LYS	HB3	1.87	0.03	2
A	251	ALA	CA	54.4	0.5	1
A	255	LYS	CG	25.3	0.5	1
A	256	LEU	H	7.76	0.03	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	253	-0.67 ± 0.16	Should be applied
$^{13}C_{\beta}$	238	-0.10 ± 0.10	None needed (< 0.5 ppm)
$^{13}C'$	184	-0.37 ± 0.16	None needed (< 0.5 ppm)

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
^{15}N	248	0.70 ± 0.24	Should be applied

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 57%, i.e. 2258 atoms were assigned a chemical shift out of a possible 3968. 0 out of 59 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	1126/1658 (68%)	466/662 (70%)	422/668 (63%)	238/328 (73%)
Sidechain	1085/2163 (50%)	625/1245 (50%)	447/835 (54%)	13/83 (16%)
Aromatic	47/147 (32%)	24/81 (30%)	23/60 (38%)	0/6 (0%)
Overall	2258/3968 (57%)	1115/1988 (56%)	892/1563 (57%)	251/417 (60%)

7.1.4 Statistically unusual chemical shifts (i)

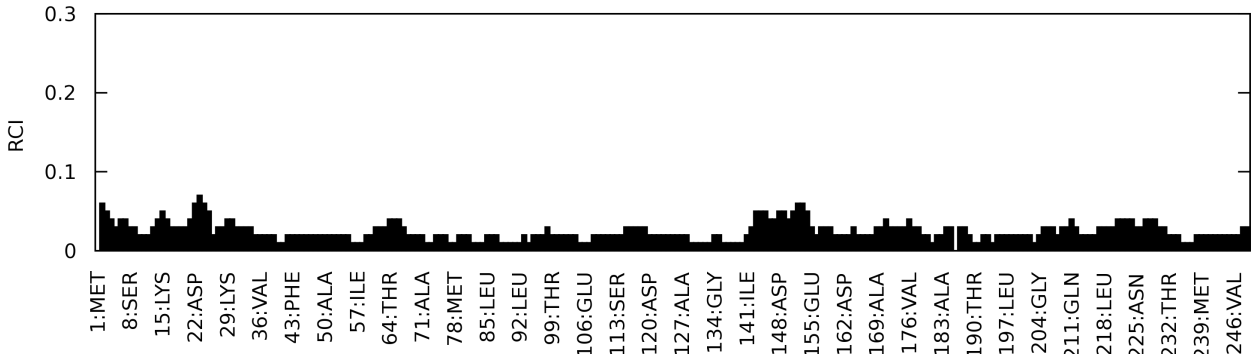
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	165	PRO	CD	43.71	55.31 – 45.41	-6.7

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 A:



7.2 Chemical shift list 2

File name: BMRB entry 4264

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	899
Number of shifts mapped to atoms	899
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

7.2.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	85	-0.59 ± 0.27	Should be applied
$^{13}C_{\beta}$	79	-0.17 ± 0.19	None needed (< 0.5 ppm)
$^{13}C'$	74	0.03 ± 0.04	None needed (< 0.5 ppm)
^{15}N	81	-0.03 ± 0.29	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 21%, i.e. 828 atoms were assigned a chemical shift out of a possible 3968. 0 out of 59 assigned methyl groups (LEU and VAL) were assigned

stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	405/1658 (24%)	165/662 (25%)	159/668 (24%)	81/328 (25%)
Sidechain	387/2163 (18%)	232/1245 (19%)	155/835 (19%)	0/83 (0%)
Aromatic	36/147 (24%)	20/81 (25%)	16/60 (27%)	0/6 (0%)
Overall	828/3968 (21%)	417/1988 (21%)	330/1563 (21%)	81/417 (19%)

7.2.4 Statistically unusual chemical shifts (i)

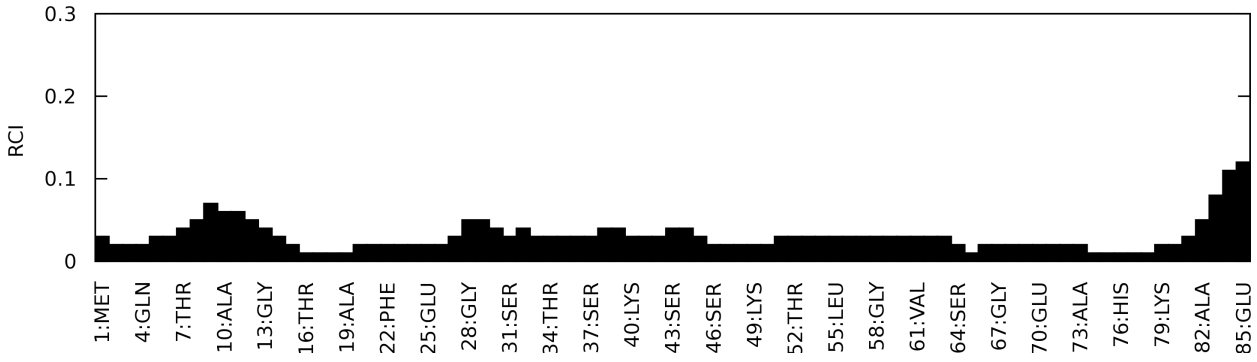
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
2	B	73	ALA	HA	1.92	6.46 – 2.06	-5.3

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 B:



7.3 Chemical shift list 3

File name: BMRB entry 4264

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	1252
Number of shifts mapped to atoms	1192
Number of unparsed shifts	0
Number of shifts with mapping errors	60
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.

- Residue not found in structure. First 5 (of 60) occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	253	LEU	CD1	24.36	0.5	2
A	254	ALA	N	121.06	0.25	1
A	257	LYS	CE	41.65	0.5	1
A	252	GLU	H	8.03	0.03	1
A	256	LEU	CD1	24.42	0.5	2

7.3.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	256	-0.23 ± 0.07	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	238	0.77 ± 0.08	Should be applied
$^{13}C'$	0	—	—
^{15}N	248	1.03 ± 0.15	Should be applied

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 30%, i.e. 1205 atoms were assigned a chemical shift out of a possible 3968. 0 out of 59 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	722/1658 (44%)	238/662 (36%)	246/668 (37%)	238/328 (73%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Sidechain	483/2163 (22%)	28/1245 (2%)	439/835 (53%)	16/83 (19%)
Aromatic	0/147 (0%)	0/81 (0%)	0/60 (0%)	0/6 (0%)
Overall	1205/3968 (30%)	266/1988 (13%)	685/1563 (44%)	254/417 (61%)

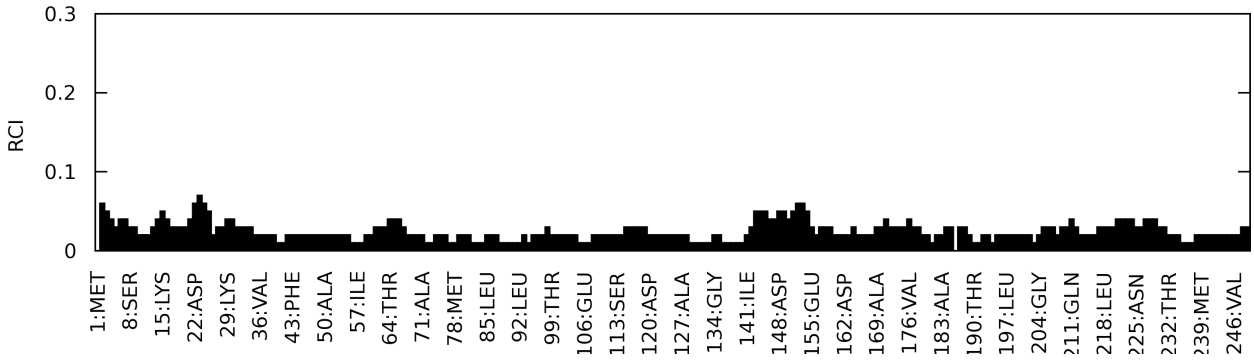
7.3.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

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 A:



7.4 Chemical shift list 4

File name: BMRB entry 4264

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	414
Number of shifts mapped to atoms	414

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

7.4.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	85	-0.16 ± 0.13	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	77	0.63 ± 0.14	Should be applied
$^{13}C'$	0	—	—
^{15}N	82	0.41 ± 0.58	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 10%, i.e. 412 atoms were assigned a chemical shift out of a possible 3968. 0 out of 59 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	249/1658 (15%)	82/662 (12%)	85/668 (13%)	82/328 (25%)
Sidechain	163/2163 (8%)	14/1245 (1%)	142/835 (17%)	7/83 (8%)
Aromatic	0/147 (0%)	0/81 (0%)	0/60 (0%)	0/6 (0%)
Overall	412/3968 (10%)	96/1988 (5%)	227/1563 (15%)	89/417 (21%)

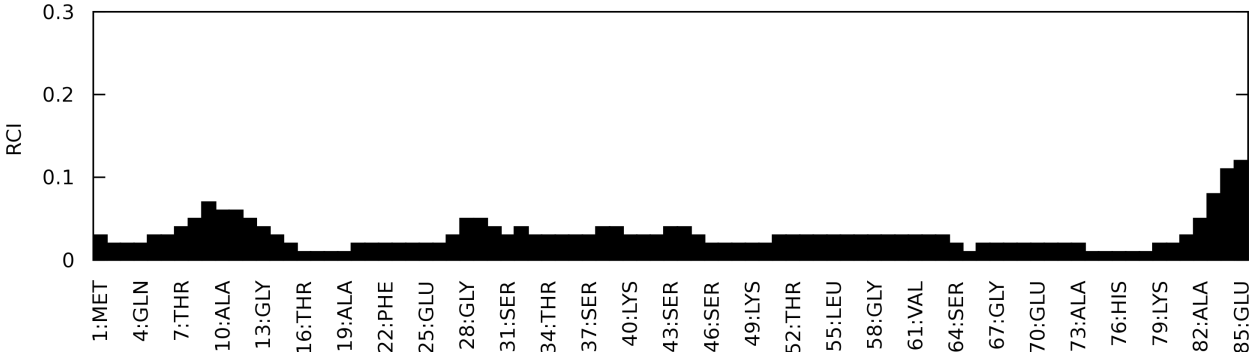
7.4.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

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 B:


RCI
0.3
0.2
0.1
0
1:MET
4:GLN
7:THR
10:ALA
13:GLY
16:THR
19:ALA
22:PHE
25:GLU
28:GLY
31:SER
34:THR
37:SER
40:LYS
43:SER
46:SER
49:LYS
52:THR
55:LEU
58:GLY
61:VAL
64:SER
67:GLY
70:GLU
73:ALA
76:HIS
79:LYS
82:ALA
85:GLU