WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report ⓘ

Mar 4, 2017 – 10:10 am GMT

PDB ID	:	5N9U
Title	:	Dehydroascorbate reductase 3A from Populus trichocarpa complexed with GSH.
Authors	:	Roret, T.; Tsan, P.
Deposited on	:	2017-02-27

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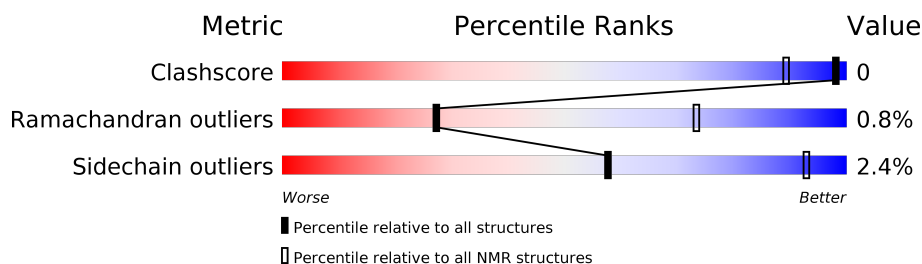with specific help available everywhere you see the ⓘ symbol.

---

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 14%.

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	Percentile Ranks	Value
Clashscore		0
Ramachandran outliers		0.8%
Sidechain outliers		2.4%

Worse   *Better*

■ Percentile relative to all structures
▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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	218	88% 8% • •

2 Ensemble composition and analysis ⓘ

This entry contains 10 models. Model 6 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:2-A:212 (211)	0.23	6

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 and 2 single-model clusters were found.

Cluster number	Models
1	2, 8, 9, 10
2	4, 5, 6, 7
Single-model clusters	1; 3

Full wwPDB NMR Structure Validation Report

3 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 3395 atoms, of which 1714 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Dehydroascorbate reductase family protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	212	3360	1085	1699	272	300	4	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	40	HIS	TYR	conflict	UNP B9HM36
A	171	PRO	THR	conflict	UNP B9HM36
A	213	HIS	-	expression tag	UNP B9HM36
A	214	HIS	-	expression tag	UNP B9HM36
A	215	HIS	-	expression tag	UNP B9HM36
A	216	HIS	-	expression tag	UNP B9HM36
A	217	HIS	-	expression tag	UNP B9HM36
A	218	HIS	-	expression tag	UNP B9HM36

- Molecule 2 is GLUTATHIONE (three-letter code: GSH) (formula: $C_{10}H_{17}N_3O_6S$).
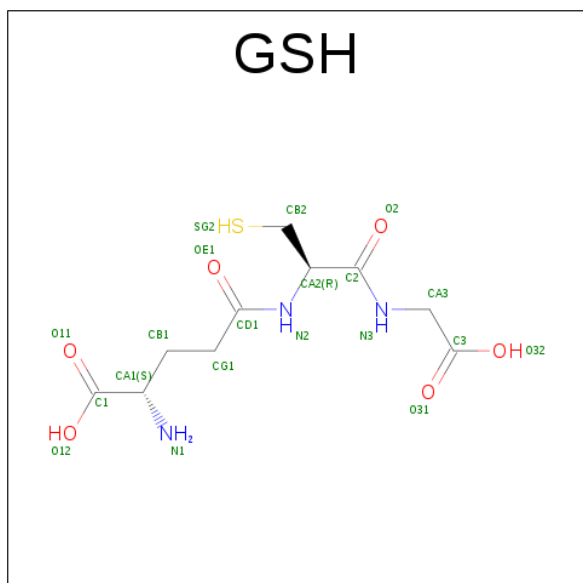
Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	S
2	A	1	35	10	15	3	6	1

4 Residue-property plots ⓘ

4.1 Average score per residue in the NMR ensemble

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.

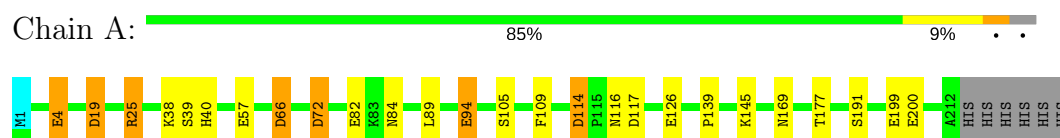• Molecule 1: Dehydroascorbate reductase family protein
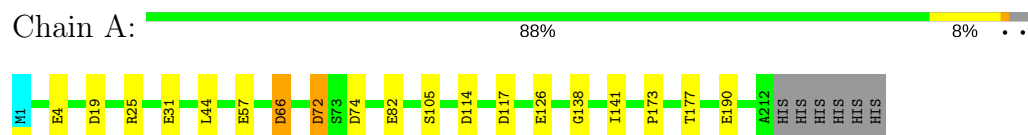
4.2 Scores per residue for each member of the ensemble
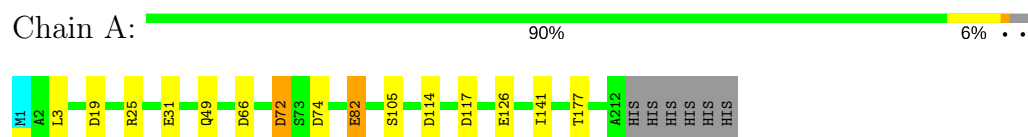
Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Dehydroascorbate reductase family protein

4.2.2 Score per residue for model 2

• Molecule 1: Dehydroascorbate reductase family protein

4.2.3 Score per residue for model 3

• Molecule 1: Dehydroascorbate reductase family protein

Chain A:

4.2.4 Score per residue for model 4

• Molecule 1: Dehydroascorbate reductase family protein

Chain A:

4.2.5 Score per residue for model 5

• Molecule 1: Dehydroascorbate reductase family protein

Chain A:
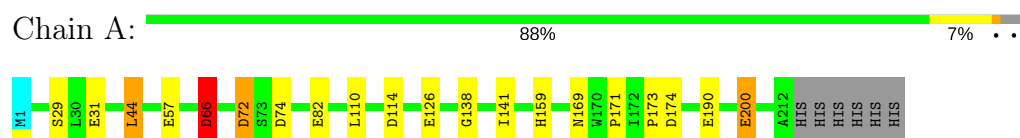
4.2.6 Score per residue for model 6 (medoid)

• Molecule 1: Dehydroascorbate reductase family protein
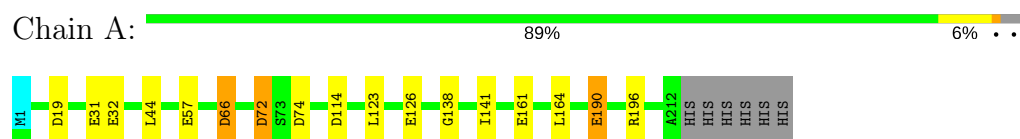
Chain A:
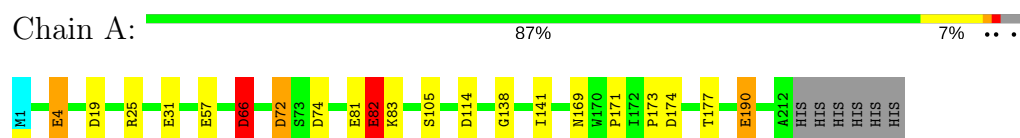
4.2.7 Score per residue for model 7

• Molecule 1: Dehydroascorbate reductase family protein
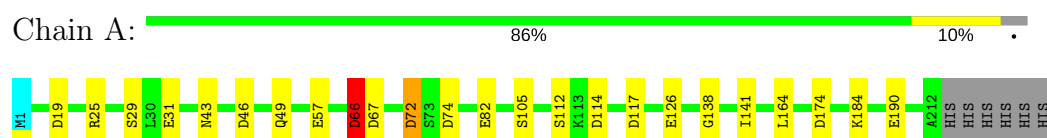
Chain A:
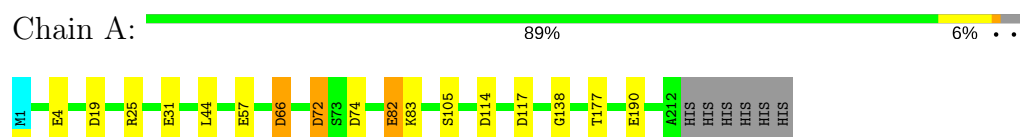
4.2.8 Score per residue for model 8

• Molecule 1: Dehydroascorbate reductase family protein

Chain A:

89%　　　6%

M1 D19 R25 E31 D46 E57 D66 D67 D72 S73 D74 E82 K83 S105 D114 G138 P173 T177 E190 A212 HIS HIS HIS HIS HIS HIS
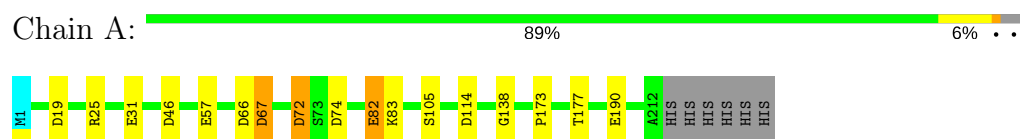
4.2.9 Score per residue for model 9

• Molecule 1: Dehydroascorbate reductase family protein

Chain A:

89%　　　7%

M1 D19 R25 L44 G45 D46 E57 D66 D72 S73 D74 S105 D114 D117 E126 I141 P173 T177 I203 A212 HIS HIS HIS HIS HIS HIS
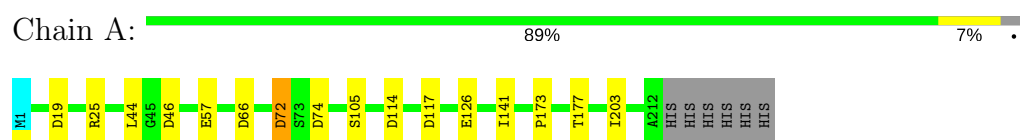
4.2.10 Score per residue for model 10

• Molecule 1: Dehydroascorbate reductase family protein

Chain A:

90%　　　6%

M1 D19 R25 D46 E57 D66 D72 S73 D74 E94 S105 D114 D117 E126 P173 T177 E200 A212 HIS HIS HIS HIS HIS HIS
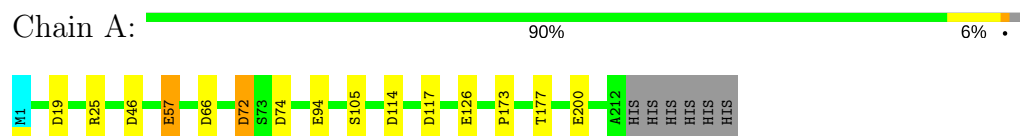
5 Refinement protocol and experimental data overview ⓘ

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

Of the 100 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
YASARA	refinement	

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)	5n9u_cs.str
Number of chemical shift lists	2
Total number of shifts	752
Number of shifts mapped to atoms	752
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	14%

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

6 Model quality ⓘ

6.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GSH

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	0.65±0.01	0±0/1698 (0.0±0.0%)	0.99±0.01	2±1/2309 (0.1±0.0%)
All	All	0.65	0/16980 (0.0%)	0.99	18/23090 (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	14.1±2.7
All	All	0	141

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	25	ARG	NE-CZ-NH2	7.07	123.83	120.30	6	7
1	A	66	ASP	CB-CG-OD1	-5.87	113.02	118.30	6	3
1	A	25	ARG	NE-CZ-NH1	5.81	123.20	120.30	1	1
1	A	82	GLU	OE1-CD-OE2	-5.72	116.44	123.30	6	2
1	A	4	GLU	OE1-CD-OE2	-5.59	116.59	123.30	1	2
1	A	67	ASP	CB-CG-OD2	-5.54	113.31	118.30	8	1
1	A	161	GLU	OE1-CD-OE2	-5.52	116.68	123.30	3	1
1	A	57	GLU	OE1-CD-OE2	-5.01	117.28	123.30	10	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	72	ASP	Sidechain	10
1	A	114	ASP	Sidechain,Mainchain,Peptide	10
1	A	19	ASP	Sidechain,Peptide	9
1	A	74	ASP	Sidechain	9
1	A	105	SER	Mainchain	8
1	A	57	GLU	Sidechain	8
1	A	126	GLU	Sidechain	7
1	A	31	GLU	Sidechain	7
1	A	82	GLU	Sidechain	7
1	A	138	GLY	Mainchain,Peptide	6
1	A	66	ASP	Sidechain	6
1	A	117	ASP	Sidechain	6
1	A	190	GLU	Sidechain	5
1	A	173	PRO	Mainchain	5
1	A	46	ASP	Mainchain	4
1	A	4	GLU	Sidechain	3
1	A	94	GLU	Sidechain	2
1	A	49	GLN	Mainchain	2
1	A	171	PRO	Peptide	2
1	A	67	ASP	Sidechain	1
1	A	116	ASN	Mainchain	1
1	A	184	LYS	Mainchain	1
1	A	25	ARG	Sidechain	1
1	A	43	ASN	Peptide	1
1	A	199	GLU	Sidechain	1
1	A	200	GLU	Sidechain	1
1	A	81	GLU	Sidechain	1

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	1653	1688	1688	1±2
All	All	16730	17030	17020	12

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:38:LYS:HE2	1:A:40:HIS:NE2	0.58	2.12	1	1
1:A:89:LEU:HD22	1:A:145:LYS:NZ	0.48	2.22	1	1
1:A:110:LEU:HD13	1:A:159:HIS:CD2	0.48	2.44	5	1
1:A:94:GLU:CD	1:A:94:GLU:H	0.45	2.15	1	1
1:A:82:GLU:CD	1:A:83:LYS:HE3	0.45	2.32	6	2
1:A:82:GLU:OE1	1:A:83:LYS:HE3	0.43	2.14	8	1
1:A:40:HIS:N	1:A:40:HIS:CD2	0.43	2.87	1	1
1:A:82:GLU:OE2	1:A:83:LYS:HE3	0.42	2.14	7	1
1:A:190:GLU:CD	1:A:190:GLU:H	0.41	2.19	6	1
1:A:39:SER:C	1:A:40:HIS:HD2	0.41	2.18	1	1
1:A:38:LYS:HE2	1:A:40:HIS:CE1	0.40	2.50	1	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	210/218 (96%)	192±1 (92±1%)	16±2 (8±1%)	2±1 (1±0%)	27	73
All	All	2100/2180 (96%)	1924 (92%)	159 (8%)	17 (1%)	27	73

All 6 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	72	ASP	10
1	A	169	ASN	2
1	A	44	LEU	2
1	A	139	PRO	1
1	A	19	ASP	1
1	A	196	ARG	1

6.3.2 Protein sidechains ⓘ

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	182/189 (96%)	178±2 (98±1%)	4±2 (2±1%)	58	93
All	All	1820/1890 (96%)	1776 (98%)	44 (2%)	58	93

All 21 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	177	THR	7
1	A	141	ILE	6
1	A	66	ASP	4
1	A	200	GLU	3
1	A	44	LEU	3
1	A	174	ASP	3
1	A	29	SER	2
1	A	190	GLU	2
1	A	164	LEU	2
1	A	57	GLU	1
1	A	112	SER	1
1	A	84	ASN	1
1	A	114	ASP	1
1	A	109	PHE	1
1	A	3	LEU	1
1	A	32	GLU	1
1	A	191	SER	1
1	A	169	ASN	1
1	A	67	ASP	1
1	A	123	LEU	1
1	A	203	ILE	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 carbohydrates in this entry.

6.6   Ligand geometry ⓘ

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	GSH	A	301	1	11,19,19	0.81±0.09	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	GSH	A	301	1	14,24,24	1.13±0.10	0±0 (0±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GSH	A	301	1	-	0±0,18,24,24	0±0,0,0,0

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 14% for the well-defined parts and 14% for the entire structure.

7.1 Chemical shift list 1

File name: 5n9u_cs.str

Chemical shift list name: *PtDHAR3A-GSH.txt*

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	376
Number of shifts mapped to atoms	376
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
$^{13}C_\alpha$	0	—	None (insufficient data)
$^{13}C_\beta$	0	—	None (insufficient data)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	188	0.95 ± 0.28	Should be applied

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 14%, i.e. 376 atoms were assigned a chemical shift out of a possible 2609. 0 out of 42 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	376/1019 (37%)	188/404 (47%)	0/422 (0%)	188/193 (97%)
Sidechain	0/1379 (0%)	0/810 (0%)	0/527 (0%)	0/42 (0%)

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	0/211 (0%)	0/114 (0%)	0/86 (0%)	0/11 (0%)
Overall	376/2609 (14%)	188/1328 (14%)	0/1035 (0%)	188/246 (76%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 14%, i.e. 376 atoms were assigned a chemical shift out of a possible 2622. 0 out of 42 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

|	Total	^{1}H	^{13}C	^{15}N
Backbone	376/1024 (37%)	188/406 (46%)	0/424 (0%)	188/194 (97%)
Sidechain	0/1387 (0%)	0/815 (0%)	0/530 (0%)	0/42 (0%)
Aromatic	0/211 (0%)	0/114 (0%)	0/86 (0%)	0/11 (0%)
Overall	376/2622 (14%)	188/1335 (14%)	0/1040 (0%)	188/247 (76%)

7.1.4　Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.1.5　Random Coil Index (RCI) plots ⓘ

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:
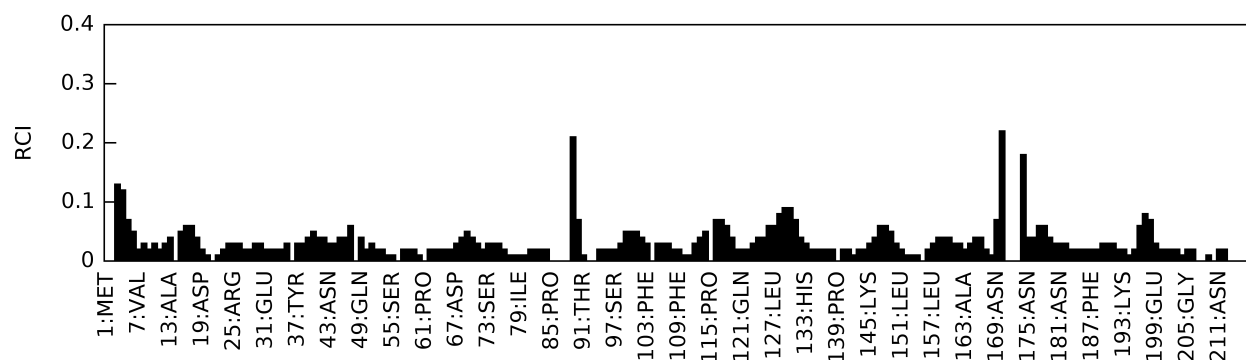
7.2　Chemical shift list 2

File name: 5n9u_cs.str

Chemical shift list name: *PtDHAR3A.txt*

7.2.1　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	376
Number of shifts mapped to atoms	376
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

7.2.2　Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	0	—	None (insufficient data)
$^{13}C_\beta$	0	—	None (insufficient data)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	188	0.93 ± 0.34	Should be applied

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 14%, i.e. 376 atoms were assigned a chemical shift out of a possible 2609. 0 out of 42 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	376/1019 (37%)	188/404 (47%)	0/422 (0%)	188/193 (97%)
Sidechain	0/1379 (0%)	0/810 (0%)	0/527 (0%)	0/42 (0%)
Aromatic	0/211 (0%)	0/114 (0%)	0/86 (0%)	0/11 (0%)
Overall	376/2609 (14%)	188/1328 (14%)	0/1035 (0%)	188/246 (76%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 14%, i.e. 376 atoms were assigned a chemical shift out of a possible 2622. 0 out of 42 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	376/1024 (37%)	188/406 (46%)	0/424 (0%)	188/194 (97%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Sidechain	0/1387 (0%)	0/815 (0%)	0/530 (0%)	0/42 (0%)
Aromatic	0/211 (0%)	0/114 (0%)	0/86 (0%)	0/11 (0%)
Overall	376/2622 (14%)	188/1335 (14%)	0/1040 (0%)	188/247 (76%)

7.2.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.2.5 Random Coil Index (RCI) plots ⓘ

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: