Full wwPDB NMR Structure Validation Report (i)

Jun 24, 2024 – 07:13 PM EDT

PDB ID	:	6SYK
BMRB ID	:	34438
Title	:	Guanine-rich oligonucleotide with 5'- and 3'-GC ends form G-quadruplex with A(GGGG)A hexad, GCGC- and G-quartets and two symmetric GG and AA base pair
Authors	:	Pavc, D.; Wang, B.; Spindler, L.; Drevensek-Olenik, I.; Plavec, J.; Sket, P.
Deposited on	:	2019-09-30

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:

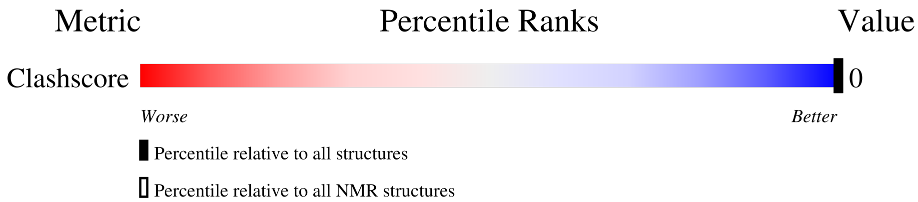
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 23%.

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

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	14	21% 64% 14%
1	B	14	14% 71% 14%

2 Ensemble composition and analysis (i)

This entry contains 10 models. This entry does not contain polypeptide chains, therefore identification of well-defined residues and clustering analysis are not possible. All residues are included in the validation scores.

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 906 atoms, of which 312 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called GCnCG.

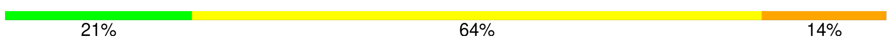
Mol	Chain	Residues	Atoms						Trace
1	A	14	Total 453	C 138	H 156	N 66	O 80	P 13	0
1	B	14	Total 453	C 138	H 156	N 66	O 80	P 13	0

4 Residue-property plots (i)

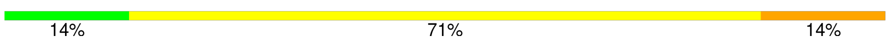
4.1 Average score per residue in the NMR ensemble

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: GCnCG

Chain A:


- Molecule 1: GCnCG

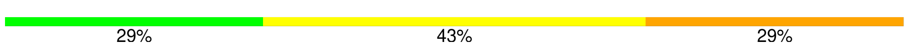
Chain B:


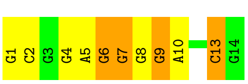
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: GCnCG

Chain A:




- Molecule 1: GCnCG

Chain B:

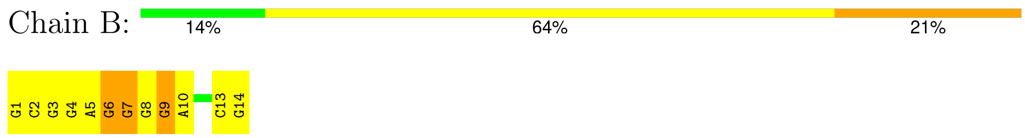

4.2.2 Score per residue for model 2

- Molecule 1: GCnCG

Chain A:

- Molecule 1: GCnCG

Chain B:



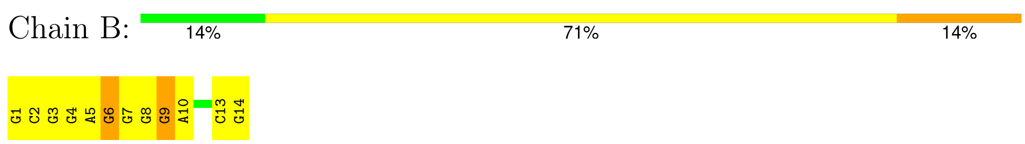
4.2.3 Score per residue for model 3

- Molecule 1: GCnCG

Chain A:

- Molecule 1: GCnCG

Chain B:



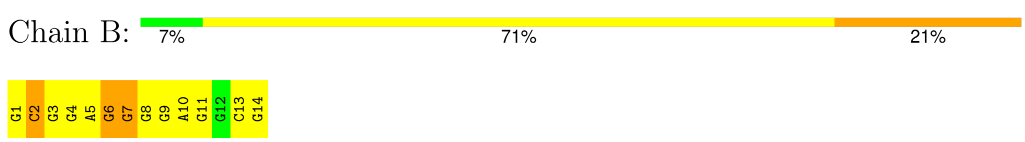
4.2.4 Score per residue for model 4

- Molecule 1: GCnCG

Chain A:

- Molecule 1: GCnCG

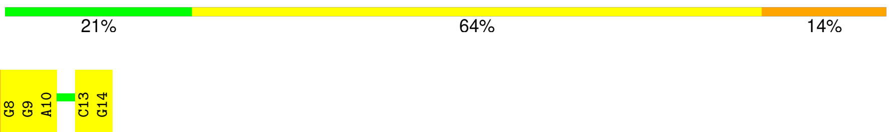
Chain B:



4.2.5 Score per residue for model 5

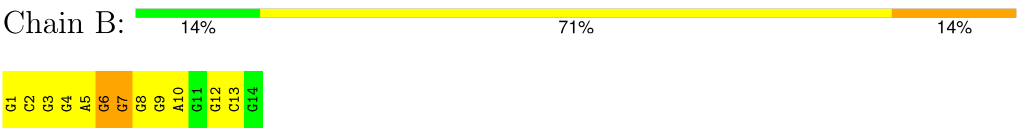
• Molecule 1: GCnCG

Chain A:



• Molecule 1: GCnCG

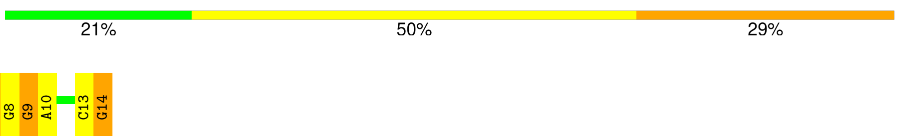
Chain B:



4.2.6 Score per residue for model 6

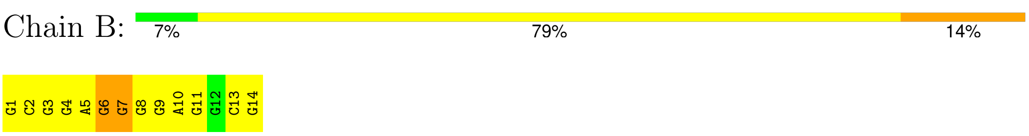
• Molecule 1: GCnCG

Chain A:



• Molecule 1: GCnCG

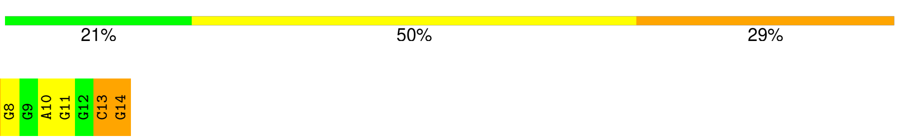
Chain B:



4.2.7 Score per residue for model 7

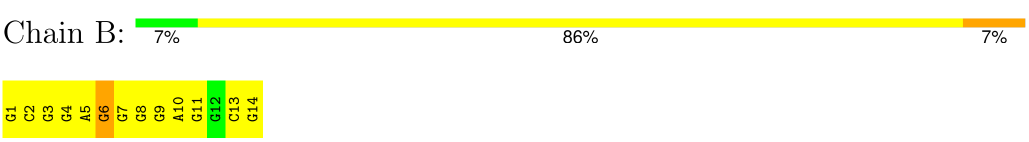
• Molecule 1: GCnCG

Chain A:



• Molecule 1: GCnCG

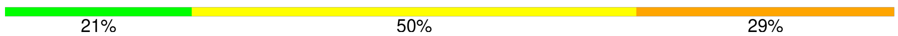
Chain B:



4.2.8 Score per residue for model 8

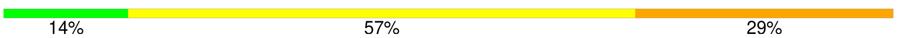
- Molecule 1: GCnCG

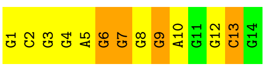
Chain A:



- Molecule 1: GCnCG

Chain B:

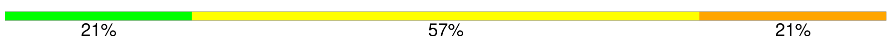




4.2.9 Score per residue for model 9

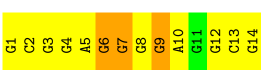
- Molecule 1: GCnCG

Chain A:



- Molecule 1: GCnCG

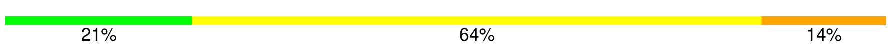
Chain B:

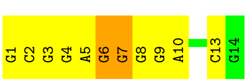


4.2.10 Score per residue for model 10

- Molecule 1: GCnCG

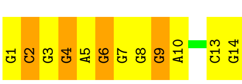
Chain A:





- Molecule 1: GCnCG

Chain B:



5 Refinement protocol and experimental data overview (i)

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

Of the 200 calculated structures, 10 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
Amber	structure calculation	14
Amber	refinement	14

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	145
Number of shifts mapped to atoms	145
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	23%

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	1.65±0.01	0±0/336 (0.0± 0.0%)	2.40±0.02	23±2/520 (4.4± 0.3%)
1	B	1.67±0.01	0±0/336 (0.0± 0.0%)	2.41±0.03	24±2/520 (4.6± 0.4%)
All	All	1.66	0/6720 (0.0%)	2.41	464/10400 (4.5%)

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	3.2±1.0
1	B	0.0±0.0	4.0±1.0
All	All	0	72

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	B	6	DG	O4'-C1'-N9	12.14	116.50	108.00	3	10
1	A	1	DG	O4'-C1'-N9	12.01	116.41	108.00	4	10
1	B	1	DG	O4'-C1'-N9	11.56	116.09	108.00	9	10
1	A	6	DG	O4'-C1'-N9	11.22	115.85	108.00	1	10
1	B	5	DA	N1-C6-N6	-10.96	112.03	118.60	9	10
1	A	5	DA	N1-C6-N6	-10.49	112.30	118.60	10	10
1	B	10	DA	N1-C6-N6	-10.23	112.46	118.60	6	10
1	A	10	DA	N1-C6-N6	-9.84	112.70	118.60	3	10
1	B	13	DC	O4'-C1'-N1	9.10	114.37	108.00	8	2
1	A	8	DG	O4'-C1'-N9	8.53	113.97	108.00	6	5
1	B	3	DG	O4'-C1'-N9	8.42	113.90	108.00	8	6
1	A	3	DG	O4'-C1'-N9	8.37	113.86	108.00	5	2
1	A	5	DA	C5-C6-N1	8.31	121.86	117.70	1	10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	B	8	DG	O4'-C1'-N9	8.31	113.82	108.00	10	5
1	B	10	DA	C5-C6-N1	8.09	121.75	117.70	7	10
1	B	5	DA	C5-C6-N1	8.02	121.71	117.70	8	10
1	A	13	DC	N3-C2-O2	-7.81	116.43	121.90	7	10
1	A	10	DA	C5-C6-N1	7.80	121.60	117.70	7	10
1	A	5	DA	C4-C5-C6	-7.70	113.15	117.00	3	10
1	A	1	DG	C1'-O4'-C4'	-7.60	102.50	110.10	1	10
1	B	5	DA	C4-C5-C6	-7.59	113.20	117.00	1	10
1	B	2	DC	O4'-C1'-N1	7.55	113.28	108.00	2	2
1	B	3	DG	O4'-C4'-C3'	-7.50	101.50	104.50	5	1
1	B	1	DG	C1'-O4'-C4'	-7.46	102.64	110.10	9	10
1	A	2	DC	N3-C2-O2	-7.40	116.72	121.90	8	10
1	B	13	DC	N3-C2-O2	-7.39	116.73	121.90	9	10
1	B	2	DC	N3-C2-O2	-7.22	116.85	121.90	1	10
1	A	7	DG	O4'-C1'-N9	6.91	112.84	108.00	1	10
1	A	5	DA	O4'-C1'-N9	6.88	112.82	108.00	5	4
1	B	7	DG	O4'-C1'-N9	6.85	112.79	108.00	10	8
1	B	9	DG	O4'-C1'-N9	6.81	112.77	108.00	7	5
1	A	9	DG	O4'-C1'-N9	6.78	112.75	108.00	10	6
1	A	2	DC	O4'-C1'-N1	6.61	112.63	108.00	1	4
1	A	13	DC	O4'-C1'-N1	6.55	112.59	108.00	8	2
1	B	10	DA	C4-C5-C6	-6.51	113.75	117.00	7	10
1	B	5	DA	O4'-C1'-N9	6.49	112.54	108.00	6	4
1	A	10	DA	C4-C5-C6	-6.41	113.79	117.00	10	10
1	B	3	DG	C1'-O4'-C4'	-6.31	103.79	110.10	6	5
1	A	6	DG	N1-C6-O6	-6.16	116.21	119.90	1	10
1	A	13	DC	N1-C2-O2	6.12	122.57	118.90	7	7
1	B	6	DG	N1-C6-O6	-6.04	116.27	119.90	1	10
1	B	12	DG	O4'-C1'-N9	6.03	112.22	108.00	5	2
1	A	5	DA	O4'-C4'-C3'	5.96	109.58	106.00	4	1
1	B	13	DC	N1-C2-O2	5.95	122.47	118.90	9	7
1	A	3	DG	O4'-C4'-C3'	-5.90	102.14	104.50	7	3
1	A	2	DC	N1-C2-O2	5.85	122.41	118.90	6	10
1	B	2	DC	N3-C4-C5	5.81	124.22	121.90	5	3
1	A	3	DG	N1-C6-O6	-5.80	116.42	119.90	10	4
1	A	14	DG	N1-C6-O6	-5.78	116.43	119.90	8	8
1	B	8	DG	C5-C6-N1	5.75	114.38	111.50	8	6
1	A	8	DG	C5-C6-N1	5.74	114.37	111.50	7	8
1	B	4	DG	N1-C6-O6	-5.73	116.46	119.90	1	3
1	B	2	DC	N1-C2-O2	5.70	122.32	118.90	8	10
1	B	14	DG	N1-C6-O6	-5.69	116.48	119.90	4	6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	B	9	DG	C5-C6-N1	5.61	114.30	111.50	9	5
1	A	1	DG	N3-C4-C5	-5.59	125.81	128.60	4	10
1	A	2	DC	N3-C4-C5	5.58	124.13	121.90	6	6
1	B	1	DG	N3-C4-C5	-5.55	125.83	128.60	7	10
1	A	1	DG	N3-C4-N9	5.51	129.31	126.00	4	6
1	A	3	DG	C1'-O4'-C4'	-5.51	104.59	110.10	2	2
1	B	1	DG	N3-C4-N9	5.50	129.30	126.00	2	8
1	A	9	DG	C5-C6-N1	5.44	114.22	111.50	1	4
1	A	4	DG	O4'-C1'-N9	5.39	111.77	108.00	9	2
1	B	3	DG	N1-C6-O6	-5.26	116.75	119.90	5	6
1	A	8	DG	O4'-C4'-C3'	5.23	109.14	106.00	2	1
1	B	8	DG	N1-C6-O6	-5.13	116.82	119.90	4	2
1	A	4	DG	N1-C6-O6	-5.12	116.83	119.90	10	1
1	B	5	DA	O4'-C4'-C3'	5.11	109.07	106.00	2	1
1	B	5	DA	C6-C5-N7	5.10	135.87	132.30	1	1
1	B	13	DC	N3-C4-C5	5.08	123.93	121.90	5	1
1	B	8	DG	N3-C4-C5	-5.07	126.06	128.60	10	2
1	B	6	DG	C5-C6-N1	5.05	114.03	111.50	8	3
1	B	8	DG	O4'-C4'-C3'	5.04	109.02	106.00	6	1
1	B	12	DG	N3-C2-N2	-5.02	116.39	119.90	9	1
1	A	5	DA	C6-C5-N7	5.01	135.80	132.30	7	1
1	B	3	DG	C5-C6-N1	5.00	114.00	111.50	2	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	6	DG	Sidechain	10
1	A	7	DG	Sidechain	10
1	B	6	DG	Sidechain	10
1	B	7	DG	Sidechain	8
1	B	4	DG	Sidechain	8
1	B	9	DG	Sidechain	5
1	A	14	DG	Sidechain	4
1	A	13	DC	Sidechain	3
1	B	11	DG	Sidechain	3
1	A	9	DG	Sidechain	2
1	A	4	DG	Sidechain	2
1	B	14	DG	Sidechain	2
1	B	2	DC	Sidechain	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	B	5	DA	Sidechain	1
1	A	11	DG	Sidechain	1
1	B	13	DC	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
All	All	5940	3120	3120	-

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)

There are no protein molecules in this entry.

6.3.2 Protein sidechains (i)

There are no protein molecules in this entry.

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 monosaccharides 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 23% for the well-defined parts and 23% 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 (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	145
Number of shifts mapped to atoms	145
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	6

7.1.2 Chemical shift referencing (i)

No chemical shift referencing corrections were calculated (not enough data).

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 23%, i.e. 132 atoms were assigned a chemical shift out of a possible 580. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Sugar	95/336 (28%)	95/196 (48%)	0/140 (0%)	0/0 (—%)
Base	37/244 (15%)	37/160 (23%)	0/36 (0%)	0/48 (0%)
Overall	132/580 (23%)	132/356 (37%)	0/176 (0%)	0/48 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 23%, i.e. 132 atoms were assigned a chemical shift out of a possible 580. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Sugar	95/336 (28%)	95/196 (48%)	0/140 (0%)	0/0 (—%)
Base	37/244 (15%)	37/160 (23%)	0/36 (0%)	0/48 (0%)
Overall	132/580 (23%)	132/356 (37%)	0/176 (0%)	0/48 (0%)

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.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	4	DG	H5"	1.98	2.89 – 5.33	-8.7
1	A	4	DG	H4'	3.08	3.37 – 5.35	-6.5
1	A	9	DG	H4'	3.10	3.37 – 5.35	-6.4
1	A	4	DG	H5'	2.75	2.99 – 5.28	-6.1
1	A	1	DG	H8	6.26	6.49 – 9.09	-5.9
1	A	1	DG	H1'	4.32	4.50 – 7.35	-5.6

7.1.5 Random Coil Index (RCI) plots (i)

No *random coil index*(RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins