W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report (i)

May 29, 2020 – 01:51 am BST

PDB ID	:	2YKA
Title	:	RRM domain of mRNA export adaptor REF2-I bound to HVS ORF57 peptide
Authors	:	Tunnicliffe, R.B.; Hautbergue, G.M.; Kalra, P.; Wilson, S.A.; Golovanov, A.P.
Deposited on	:	2011-05-26

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
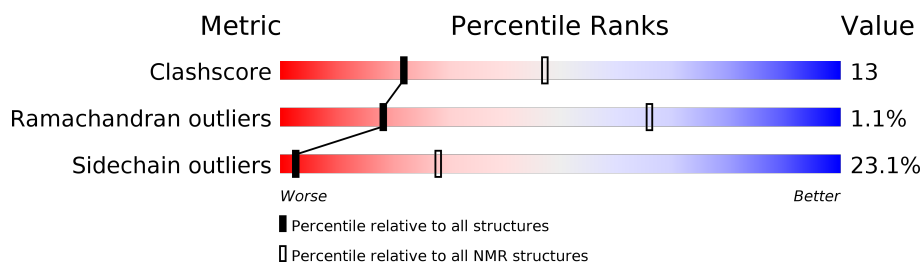with specific help available everywhere you see the (i) symbol.

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 90%.

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		13
Ramachandran outliers		1.1%
Sidechain outliers		23.1%

Worse   *Better*

■ Percentile relative to all structures

▯ Percentile relative to all NMR structures

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	124	37%　　19%　・　43%
2	B	23	30%　9%　61%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 17 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

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:74-A:107, A:114-A:150, B:108-B:116 (80)	0.12	17

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 3 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 3, 4, 5, 6, 9, 11, 12, 14, 15, 17, 18
2	7, 8, 13, 16, 20
3	2, 10, 19

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2202 atoms, of which 1076 are hydrogens and 0 are deuteriums.

• Molecule 1 is a protein called RNA AND EXPORT FACTOR-BINDING PROTEIN 2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	124	1855	587	903	174	185	6	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	40	MET	-	expression tag	UNP Q9JJW6
A	41	ALA	-	expression tag	UNP Q9JJW6
A	42	SER	-	expression tag	UNP Q9JJW6
A	43	MET	-	expression tag	UNP Q9JJW6
A	44	THR	-	expression tag	UNP Q9JJW6
A	45	GLY	-	expression tag	UNP Q9JJW6
A	46	GLY	-	expression tag	UNP Q9JJW6
A	47	GLN	-	expression tag	UNP Q9JJW6
A	48	GLN	-	expression tag	UNP Q9JJW6
A	49	MET	-	expression tag	UNP Q9JJW6
A	50	GLY	-	expression tag	UNP Q9JJW6
A	51	ARG	-	expression tag	UNP Q9JJW6
A	52	ASP	-	expression tag	UNP Q9JJW6
A	156	LEU	-	expression tag	UNP Q9JJW6
A	157	GLU	-	expression tag	UNP Q9JJW6
A	158	HIS	-	expression tag	UNP Q9JJW6
A	159	HIS	-	expression tag	UNP Q9JJW6
A	160	HIS	-	expression tag	UNP Q9JJW6
A	161	HIS	-	expression tag	UNP Q9JJW6
A	162	HIS	-	expression tag	UNP Q9JJW6
A	163	HIS	-	expression tag	UNP Q9JJW6

• Molecule 2 is a protein called 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	23	347	103	173	38	32	1	0

There are 5 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
B	98	GLY	-	expression tag	UNP P13199
B	99	PRO	-	expression tag	UNP P13199
B	100	LEU	-	expression tag	UNP P13199
B	101	GLY	-	expression tag	UNP P13199
B	102	SER	-	expression tag	UNP P13199

4 Residue-property plots (i)

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: RNA AND EXPORT FACTOR-BINDING PROTEIN 2
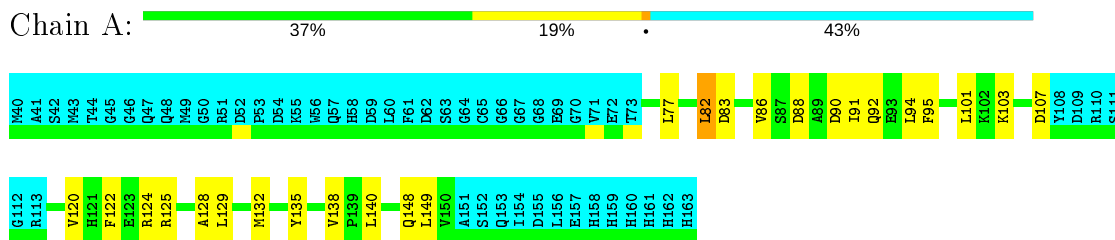
• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN
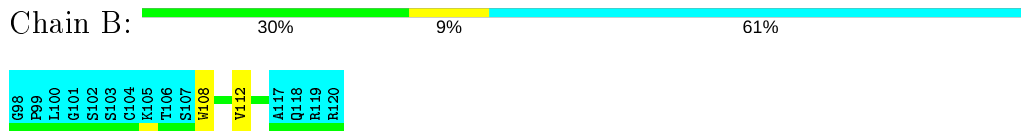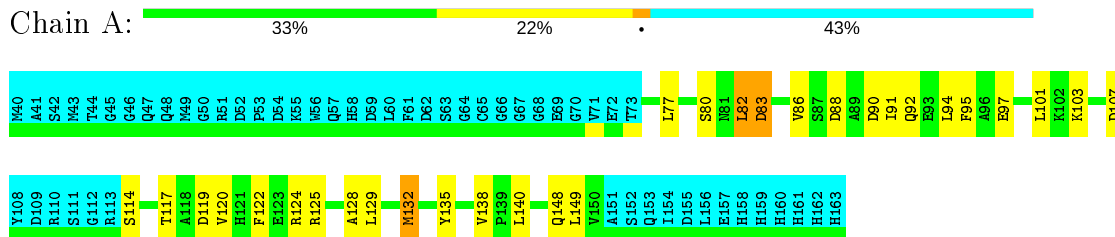
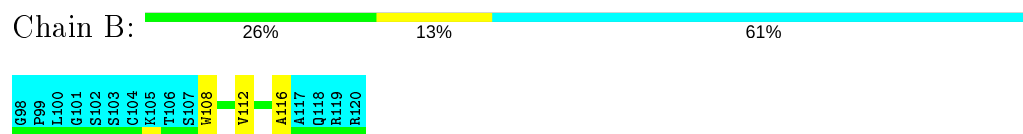

4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1
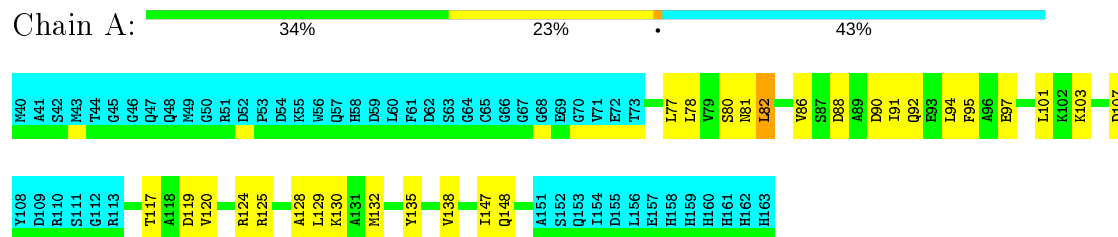
• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2
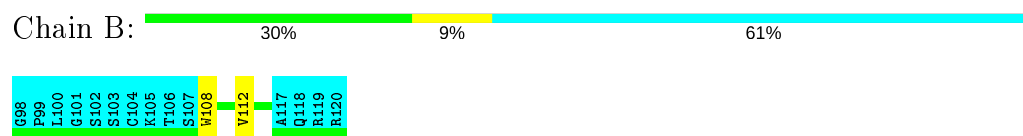

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 26%   13%   61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 V112 A116 A117 Q118 R119 R120

4.2.2   Score per residue for model 2

- Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 34%   23%   •   43%

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 L78 V79 S80 N81 L82 V86 S87 D88 A89 D90 I91 Q92 E93 L94 F95 A96 E97 L101 K102 K103 D107

Y108 D109 R110 S111 G112 R113 T117 A118 D119 V120 R124 R125 A128 L129 K130 A131 M132 Y135 V138 I147 Q148 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

- Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 30%   9%   61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 V112 A117 A118 R119 R120

4.2.3   Score per residue for model 3

- Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 35%   21%   •   43%

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 L82 D83 V86 S87 D88 A89 D90 I91 Q92 E93 L94 F95 L101 K102 K103 D107 Y108 D109 R110 S111

G112 R113 D119 V120 H121 F122 E123 R124 R125 A126 D127 A128 L129 M132 Y135 V138 P139 L140 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

- Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

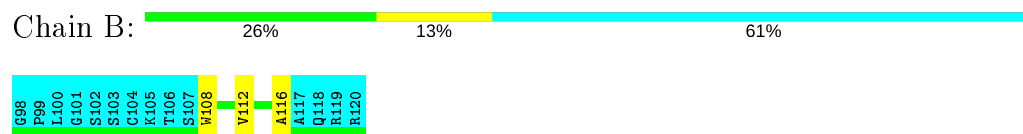Chain B: 26%   13%   61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 V112 A116 A117 Q118 R119 R120

4.2.4   Score per residue for model 4

- Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 33% 23% · 43%

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 L78 L82 D83 V86 S87 D88 A89 D90 I91 Q92 E93 L94 F95 L101 K102 K103 D107 Y108 D109 R110

S111 G112 R113 S114 A118 D119 V120 H121 F122 E123 R124 R125 A128 L129 M132 K133 Q134 Y135 V138 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 35% · 61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 A117 Q118 R119 R120

4.2.5 Score per residue for model 5

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 34% 23% · 43%

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 S80 N81 L82 D83 V86 S87 D88 A89 D90 I91 Q92 E93 L94 F95 A96 E97 L101 K102 K103 D107

Y108 D109 R110 S111 G112 R113 T117 V120 H121 F122 E123 R124 R125 A128 L129 M132 K133 Q134 Y135 V138 P139 L140 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 26% 13% 61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 A109 D110 R111 V112 A117 Q118 R119 R120

4.2.6 Score per residue for model 6

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 36% 19% · 43%

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 L82 V86 S87 D88 A89 D90 I91 Q92 E93 L94 F95 L101 K102 K103 A104 D107 Y108 D109 R110 S111

G112 R113 S114 V120 H121 F122 E123 R124 D127 A128 L129 M132 Y135 V138 P139 L140 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 26% 13% 61%

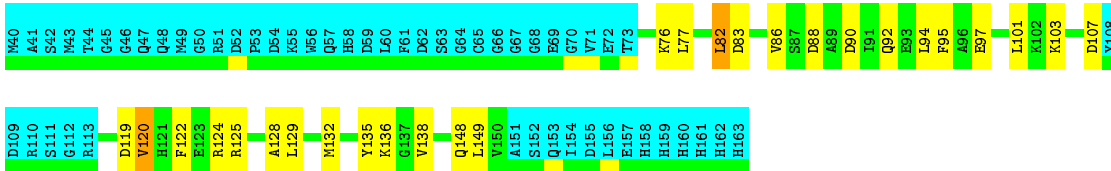G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 V112 R113 A117 Q118 R119 R120

4.2.7 Score per residue for model 7

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A:

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 K76 L77 L82 D83 V86 S87 D88 A89 D90 I91 Q92 E93 L94 F95 A96 E97 L101 K102 K103 D107 Y108

D109 R110 S111 G112 R113 D119 V120 H121 F122 E123 R124 R125 A128 L129 M132 Y135 K136 G137 V138 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B:

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 A109 D110 R111 V112 A117 Q118 R119 R120

4.2.8 Score per residue for model 8

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A:

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 L78 L82 D83 V86 S87 D88 A89 D90 I91 Q92 E93 L94 F95 L101 K102 Y108 D109 R110 S111 G112

R113 S114 L115 V120 H121 F122 E123 R124 R125 A128 L129 K130 A131 M132 K133 Q134 Y135 V138 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B:

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 A109 D110 R111 V112 R113 A117 Q118 R119 R120

4.2.9 Score per residue for model 9

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A:

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 S80 N81 L82 D83 V86 S87 D88 I91 Q92 E93 L94 F95 A96 E97 L101 K102 K103 V106 D107

Y108 D109 R110 S111 G112 R113 T117 A118 D119 V120 R124 R125 L129 M132 Y135 V138 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 35% · 61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 A117 Q118 R119 R120
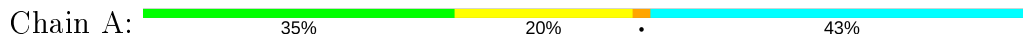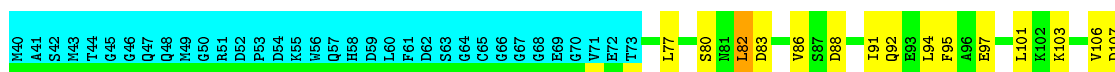
4.2.10 Score per residue for model 10

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 34% 20% · 43%

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 L78 V79 S80 N81 L82 D83 V86 S87 D90 I91 Q92 E93 L94 F95 A96 E97 L101 K102 K103 D107

Y108 D109 R110 S111 G112 R113 V120 H121 F122 E123 R124 R125 A128 L129 M132 Y135 V138 P139 L140 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 30% 9% 61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 V112 A117 Q118 R119 R120

4.2.11 Score per residue for model 11

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 35% 19% · 43%

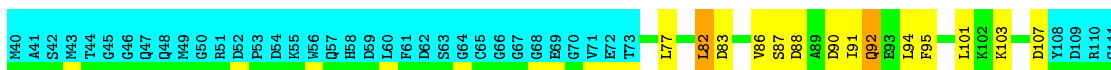M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 L78 L82 D83 V86 S87 D88 A89 D90 I91 Q92 E93 L94 F95 A96 E97 L101 K102 D107 Y108 D109

R110 S111 G112 R113 A118 D119 V120 R124 R125 A128 L129 Q134 Y135 V138 P139 L140 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 30% 9% 61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 A109 D110 A117 Q118 R119 R120

4.2.12 Score per residue for model 12

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 36% 19% • 43%

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 L82 D83 V86 S87 D88 A89 D90 I91 Q92 E93 L94 F95 L101 K102 K103 D107 Y108 D109 R110 S111

G112 R113 S114 V120 H121 F122 E123 R124 A128 L129 M132 K133 D134 Y135 V138 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 22% 17% 61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 W108 A109 D110 R111 V112 A116 A117 Q118 R119 R120

4.2.13 Score per residue for model 13

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 39% 16% • 43%

M40 A41 S42 M43 T44 G45 G46 Q47 Q48 M49 G50 R51 D52 P53 D54 K55 W56 Q57 H58 D59 L60 F61 D62 S63 G64 C65 G66 G67 G68 E69 G70 V71 E72 T73 L77 N81 L82 D83 V86 S87 D88 Q92 E93 L94 F95 L101 K102 K103 D107 Y108 D109 R110 S111 G112

R113 T117 V120 R124 A128 M132 L140 I147 Q148 L149 V150 A151 S152 Q153 I154 D155 L156 E157 H158 H159 H160 H161 H162 H163

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B: 39% 61%

G98 P99 L100 G101 S102 S103 C104 K105 T106 S107 A117 Q118 R119 R120
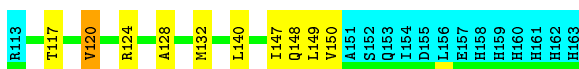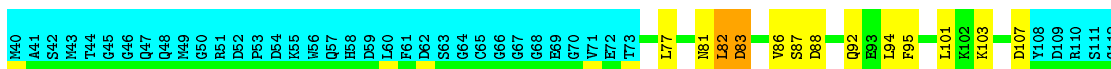
4.2.14 Score per residue for model 14

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A: 34% 22% • 43%

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

4.2.15 Score per residue for model 15

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

4.2.16 Score per residue for model 16
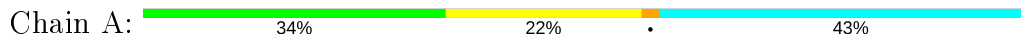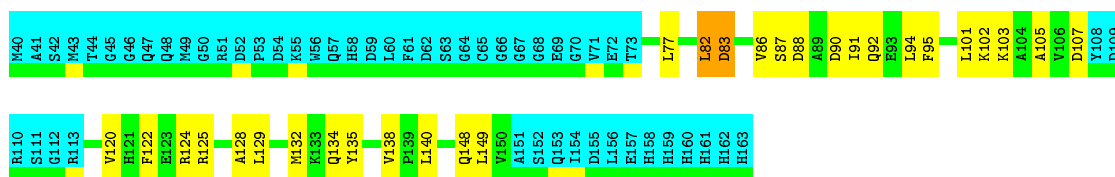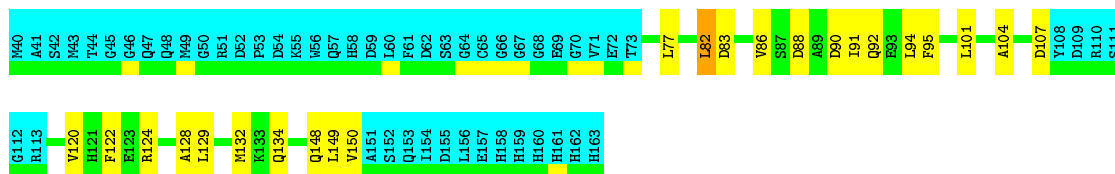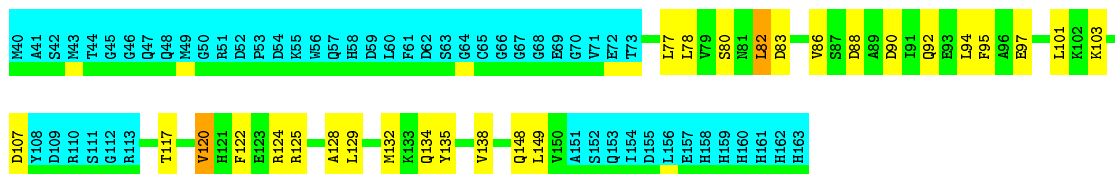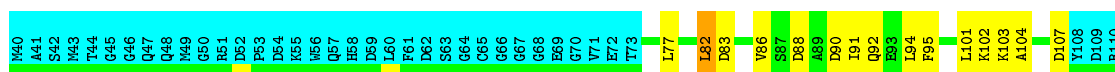
• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

4.2.17　Score per residue for model 17 (medoid)

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

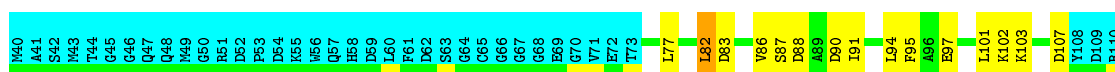Chain A:

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B:

4.2.18　Score per residue for model 18

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A:

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B:

4.2.19　Score per residue for model 19

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A:

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B:

4.2.20 Score per residue for model 20

• Molecule 1: RNA AND EXPORT FACTOR-BINDING PROTEIN 2

Chain A:

• Molecule 2: 52 KDA IMMEDIATE-EARLY PHOSPHOPROTEIN

Chain B:
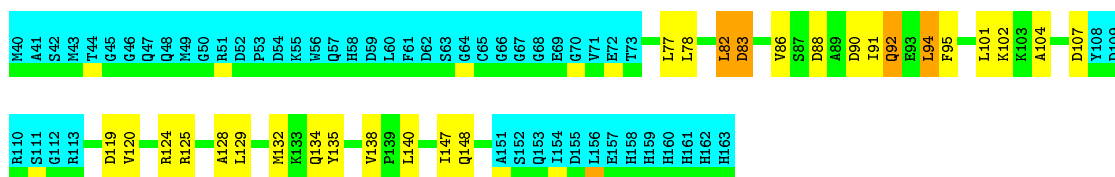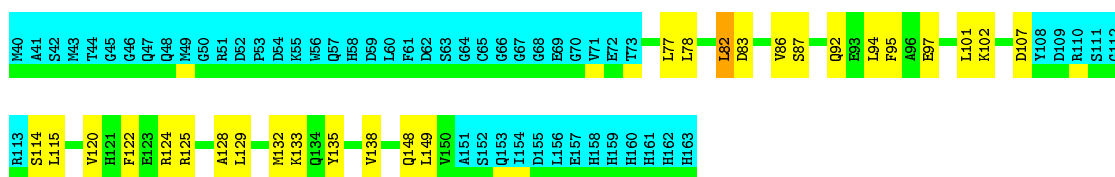
5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 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
CYANA	refinement	2.1
CYANA	structure solution	2.1

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)	input_cs.cif
Number of chemical shift lists	2
Total number of shifts	1612
Number of shifts mapped to atoms	1612
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	90%

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	540	548	548	15±2
2	B	75	70	70	2±1
All	All	12300	12360	12360	316

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 13.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
1:A:91:ILE:HD12	1:A:120:VAL:HG22	0.91	1.39	3	14
1:A:77:LEU:HD11	1:A:128:ALA:HB1	0.80	1.54	4	2
1:A:86:VAL:HG22	2:B:108:TRP:CZ2	0.73	2.18	18	4
1:A:140:LEU:HD11	2:B:112:VAL:HG22	0.72	1.61	10	1
1:A:77:LEU:HD21	1:A:128:ALA:HB1	0.69	1.63	6	17
1:A:82:LEU:HD13	1:A:86:VAL:HG11	0.65	1.68	13	16
1:A:140:LEU:CD1	2:B:112:VAL:HG22	0.64	2.21	10	4
1:A:129:LEU:HD23	1:A:149:LEU:HD22	0.64	1.68	11	2
1:A:76:LYS:C	1:A:77:LEU:HD22	0.64	2.12	7	1
1:A:78:LEU:HD12	1:A:118:ALA:O	0.62	1.94	4	2
1:A:95:PHE:CE2	1:A:120:VAL:HG21	0.61	2.31	4	9
1:A:95:PHE:CB	1:A:101:LEU:HD11	0.61	2.26	15	20
1:A:77:LEU:HD11	1:A:128:ALA:CB	0.61	2.24	4	1

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:95:PHE:HB2	1:A:101:LEU:HD11	0.60	1.73	10	20
1:A:138:VAL:HG21	2:B:116:ALA:HB2	0.60	1.73	20	4
1:A:91:ILE:HD12	1:A:120:VAL:CG1	0.60	2.27	10	1
1:A:95:PHE:CD2	1:A:120:VAL:HG11	0.59	2.32	20	5
1:A:77:LEU:HD21	1:A:128:ALA:CB	0.59	2.28	15	12
2:B:108:TRP:O	2:B:112:VAL:HG23	0.58	1.98	10	12
1:A:94:LEU:HD12	2:B:113:ARG:HG3	0.58	1.75	6	3
1:A:77:LEU:HD13	1:A:149:LEU:HA	0.58	1.76	5	13
1:A:95:PHE:CE2	1:A:120:VAL:HG11	0.56	2.35	20	5
1:A:86:VAL:HG23	2:B:108:TRP:CZ3	0.55	2.36	16	13
1:A:77:LEU:HD23	1:A:122:PHE:CD1	0.54	2.38	17	12
1:A:77:LEU:HD13	1:A:122:PHE:CD1	0.54	2.37	4	1
1:A:149:LEU:HD22	1:A:150:VAL:N	0.54	2.18	10	3
1:A:77:LEU:CD1	1:A:128:ALA:HB1	0.54	2.33	18	1
1:A:83:ASP:CG	1:A:140:LEU:HD23	0.53	2.25	11	6
1:A:80:SER:HA	1:A:117:THR:HG22	0.52	1.81	16	5
1:A:95:PHE:CD2	1:A:120:VAL:HG21	0.52	2.40	12	12
1:A:106:VAL:HG23	1:A:118:ALA:HB2	0.52	1.82	9	1
1:A:149:LEU:HD13	1:A:149:LEU:O	0.52	2.05	10	3
1:A:77:LEU:HD12	1:A:147:ILE:CG2	0.52	2.35	2	3
1:A:77:LEU:HD21	1:A:149:LEU:HD12	0.49	1.84	11	1
1:A:77:LEU:CD1	1:A:149:LEU:HD12	0.49	2.37	18	1
1:A:82:LEU:HD13	1:A:86:VAL:CG1	0.48	2.36	13	3
1:A:91:ILE:HD12	1:A:120:VAL:CG2	0.48	2.33	15	4
1:A:129:LEU:HD13	1:A:129:LEU:O	0.48	2.09	1	4
1:A:125:ARG:HD2	1:A:149:LEU:HD11	0.48	1.86	9	2
1:A:79:VAL:CG1	1:A:82:LEU:HD23	0.47	2.39	10	1
1:A:104:ALA:CB	1:A:120:VAL:HG13	0.47	2.40	17	3
1:A:91:ILE:HD12	1:A:120:VAL:HG12	0.47	1.85	10	1
1:A:92:GLN:HG3	1:A:104:ALA:HB3	0.47	1.87	19	1
1:A:129:LEU:O	1:A:129:LEU:HD13	0.47	2.09	5	5
1:A:129:LEU:HD22	1:A:132:MET:HE3	0.47	1.86	1	1
1:A:86:VAL:HG23	2:B:108:TRP:CH2	0.47	2.45	9	4
1:A:140:LEU:HD13	2:B:112:VAL:HG22	0.46	1.87	5	1
1:A:80:SER:CB	1:A:117:THR:HG22	0.45	2.42	1	2
1:A:82:LEU:O	1:A:82:LEU:HD12	0.45	2.12	6	2
1:A:125:ARG:CD	1:A:149:LEU:HD21	0.45	2.42	17	1
1:A:77:LEU:HD12	1:A:149:LEU:HA	0.45	1.88	18	2
1:A:135:TYR:O	1:A:138:VAL:HG23	0.44	2.12	6	18
1:A:149:LEU:HD13	1:A:149:LEU:C	0.44	2.33	10	3
1:A:82:LEU:HD12	1:A:82:LEU:O	0.43	2.14	15	6

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
1:A:83:ASP:O	1:A:86:VAL:HG12	0.43	2.14	15	1
1:A:103:LYS:NZ	1:A:105:ALA:HB2	0.43	2.29	14	1
1:A:92:GLN:O	1:A:101:LEU:HD12	0.43	2.14	12	2
1:A:129:LEU:HD13	1:A:129:LEU:C	0.42	2.34	7	4
1:A:82:LEU:HB2	1:A:86:VAL:HG21	0.42	1.90	20	4
1:A:138:VAL:HG21	2:B:116:ALA:CB	0.42	2.44	20	2
1:A:104:ALA:HB2	1:A:120:VAL:HG13	0.42	1.90	17	2
1:A:129:LEU:C	1:A:129:LEU:HD13	0.42	2.35	17	1
1:A:129:LEU:HD22	1:A:132:MET:CE	0.41	2.45	1	1
1:A:76:LYS:O	1:A:77:LEU:HD22	0.41	2.15	7	1
1:A:101:LEU:HD21	1:A:122:PHE:CE2	0.41	2.51	10	1
1:A:92:GLN:CG	1:A:104:ALA:HB3	0.41	2.46	19	1
1:A:82:LEU:CD1	1:A:86:VAL:HG11	0.40	2.44	13	1

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	71/124 (57%)	65±1 (91±1%)	6±1 (8±1%)	1±0 (1±1%)	17	64
2	B	9/23 (39%)	9±0 (100±0%)	0±0 (0±0%)	0±0 (0±0%)	100	100
All	All	1600/2940 (54%)	1472 (92%)	111 (7%)	17 (1%)	18	66

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	83	ASP	17

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	56/98 (57%)	42±2 (75±3%)	14±2 (25±3%)	2	25
2	B	6/17 (35%)	6±0 (94±8%)	0±0 (6±8%)	24	73
All	All	1240/2300 (54%)	954 (77%)	286 (23%)	3	28

All 30 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	94	LEU	20
1	A	82	LEU	20
1	A	148	GLN	20
1	A	124	ARG	20
1	A	92	GLN	19
1	A	132	MET	19
1	A	107	ASP	19
1	A	88	ASP	18
1	A	90	ASP	17
1	A	125	ARG	13
1	A	103	LYS	13
1	A	102	LYS	10
1	A	134	GLN	10
1	A	97	GLU	10
1	A	87	SER	7
2	B	110	ASP	7
1	A	129	LEU	7
1	A	119	ASP	7
1	A	114	SER	6
1	A	120	VAL	4
1	A	78	LEU	4
1	A	127	ASP	3
1	A	149	LEU	3
1	A	81	ASN	3
1	A	130	LYS	2
1	A	115	LEU	1
1	A	133	LYS	1
1	A	136	LYS	1
1	A	140	LEU	1
1	A	146	ASP	1

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 90% for the well-defined parts and 82% for the entire structure.

7.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *orf57_3.bmrb.csh*

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	268
Number of shifts mapped to atoms	268
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 (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 9%, i.e. 89 atoms were assigned a chemical shift out of a possible 965. 1 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	39/396 (10%)	16/158 (10%)	15/160 (9%)	8/78 (10%)
Sidechain	39/506 (8%)	25/293 (9%)	13/188 (7%)	1/25 (4%)
Aromatic	11/63 (17%)	6/34 (18%)	4/27 (15%)	1/2 (50%)
Overall	89/965 (9%)	47/485 (10%)	32/375 (9%)	10/105 (10%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 14%, i.e. 248 atoms were assigned a chemical shift out of a possible 1744. 2 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Backbone	107/727 (15%)	44/290 (15%)	42/294 (14%)	21/143 (15%)
Sidechain	130/876 (15%)	85/515 (17%)	40/314 (13%)	5/47 (11%)
Aromatic	11/141 (8%)	6/77 (8%)	4/54 (7%)	1/10 (10%)
Overall	248/1744 (14%)	135/882 (15%)	86/662 (13%)	27/200 (14%)

7.1.4   Statistically unusual chemical shifts ⓘ
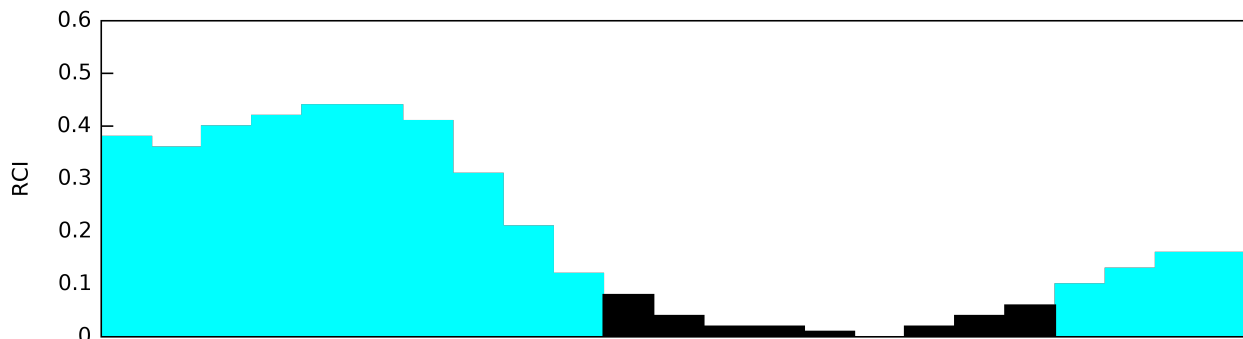
There are no statistically unusual chemical shifts.

7.1.5   Random Coil Index (RCI) plots ⓘ

The images below report *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:
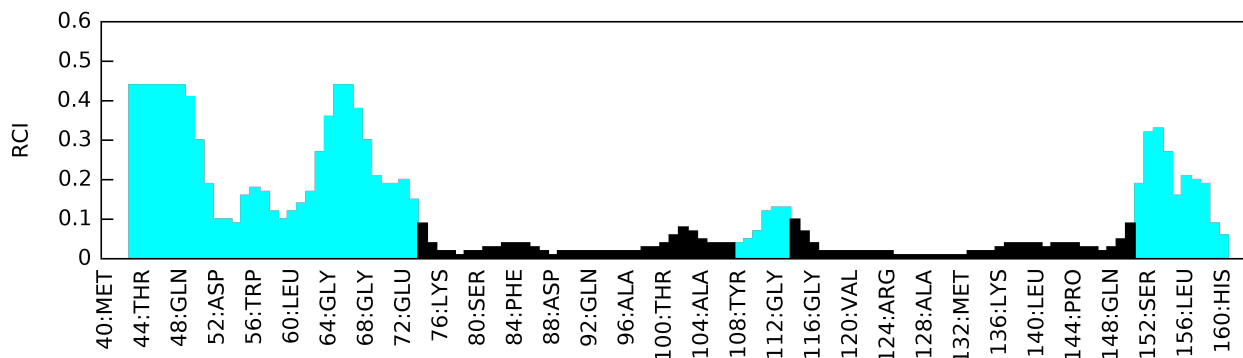
Random coil index (RCI) for chain B:

7.2   Chemical shift list 2

File name: input_cs.cif

Chemical shift list name: *ref_3.bmrb.csh*

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	1344
Number of shifts mapped to atoms	1344
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 ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	105	-0.26 ± 0.10	None needed (< 0.5 ppm)
$^{13}C_\beta$	96	-0.03 ± 0.09	None needed (< 0.5 ppm)
$^{13}C'$	112	-0.06 ± 0.07	None needed (< 0.5 ppm)
^{15}N	113	-0.13 ± 0.24	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 81%, i.e. 781 atoms were assigned a chemical shift out of a possible 965. 14 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	347/396 (88%)	140/158 (89%)	138/160 (86%)	69/78 (88%)
Sidechain	394/506 (78%)	247/293 (84%)	140/188 (74%)	7/25 (28%)
Aromatic	40/63 (63%)	25/34 (74%)	15/27 (56%)	0/2 (0%)
Overall	781/965 (81%)	412/485 (85%)	293/375 (78%)	76/105 (72%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 69%, i.e. 1204 atoms were assigned a chemical shift out of a possible 1744. 16 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	559/727 (77%)	229/290 (79%)	217/294 (74%)	113/143 (79%)
Sidechain	579/876 (66%)	370/515 (72%)	197/314 (63%)	12/47 (26%)
Aromatic	66/141 (47%)	39/77 (51%)	26/54 (48%)	1/10 (10%)
Overall	1204/1744 (69%)	638/882 (72%)	440/662 (66%)	126/200 (63%)

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	99	GLY	C	184.94	183.33 – 164.53	5.9

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: