



# Full wwPDB NMR Structure Validation Report (i)

Apr 27, 2016 – 01:37 AM BST

PDB ID : 2LSJ  
Title : Solution structure of the mouse Rev1 CTD in complex with the Rev1-interacting Region (RIR)of Pol Kappa  
Authors : Liu, J.; Wojtaszek, J.; Zhou, P.  
Deposited on : 2012-05-01

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 (i) symbol.

---

The following versions of software and data (see [references \(1\)](#)) were used in the production of this report:

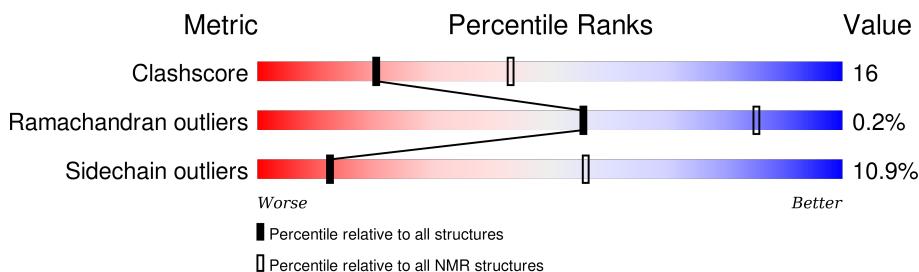
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal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

The following experimental techniques were used to determine the structure:  
*SOLUTION NMR*

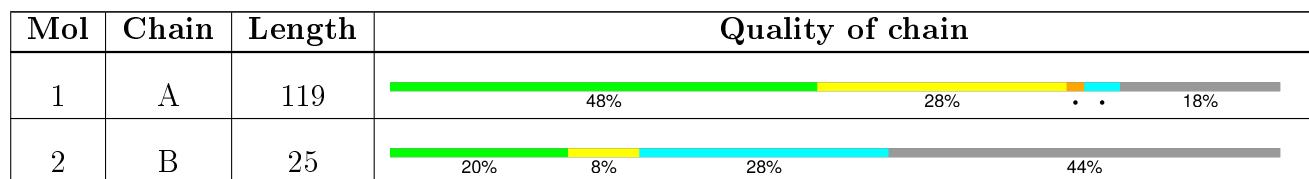
The overall completeness of chemical shifts assignment is 84%.

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$



## 2 Ensemble composition and analysis

This entry contains 20 models. Model 12 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:25-A:116, B:208-B:214 (99)	0.15	12

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 4 clusters. No single-model clusters were found.

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

### 3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1822 atoms, of which 919 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called DNA repair protein REV1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	97	1575	502	794	121	153	5	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP Q920Q2
A	2	SER	-	EXPRESSION TAG	UNP Q920Q2
A	3	GLY	-	EXPRESSION TAG	UNP Q920Q2
A	4	GLY	-	EXPRESSION TAG	UNP Q920Q2

- Molecule 2 is a protein called DNA polymerase kappa.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O		
2	B	14	247	76	125	24	22		0

There are 2 discrepancies between the modelled and reference sequences:

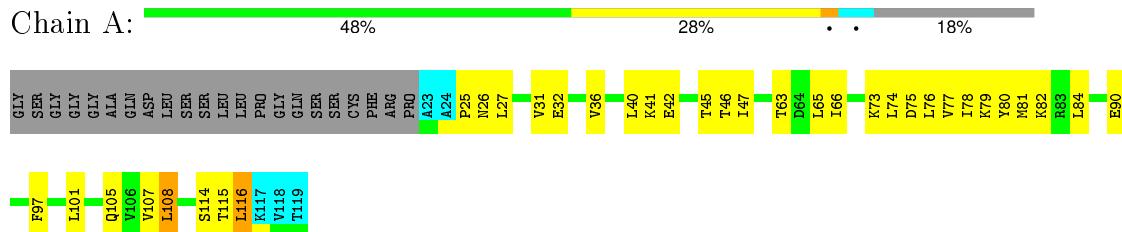
Chain	Residue	Modelled	Actual	Comment	Reference
B	201	SER	-	EXPRESSION TAG	UNP Q9QUG2
B	202	HIS	-	EXPRESSION TAG	UNP Q9QUG2

## 4 Residue-property plots [\(i\)](#)

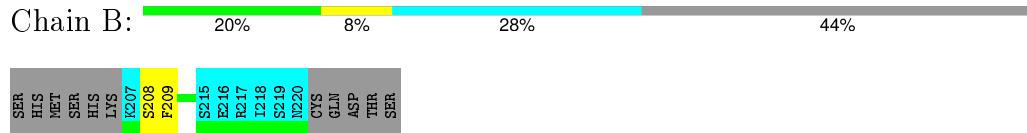
#### 4.1 Average score per residue in the NMR ensemble

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: DNA repair protein REV1



- Molecule 2: DNA polymerase kappa

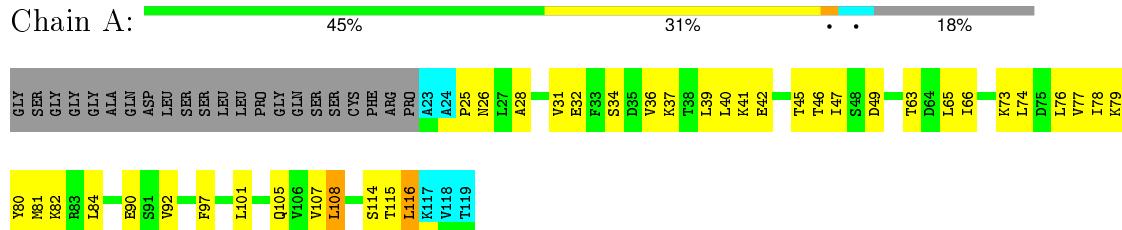


#### 4.2 Scores per residue for each member of the ensemble

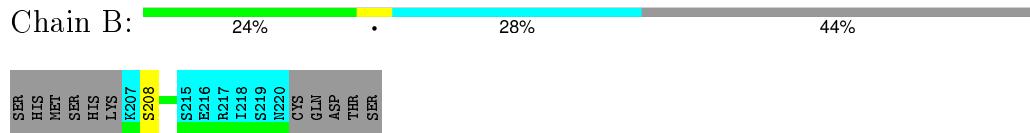
Colouring as in section 4.1 above.

#### 4.2.1 Score per residue for model 1

- Molecule 1: DNA repair protein REV1

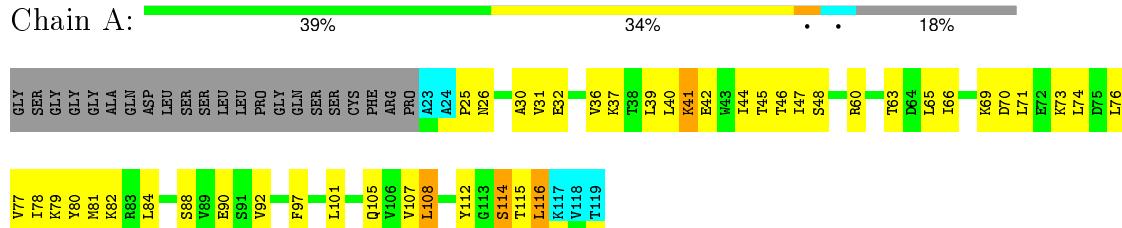


- Molecule 2: DNA polymerase kappa

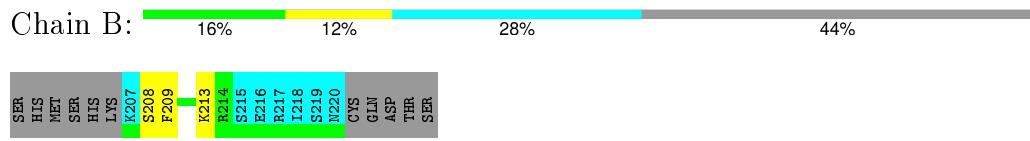


#### 4.2.2 Score per residue for model 2

- Molecule 1: DNA repair protein REV1

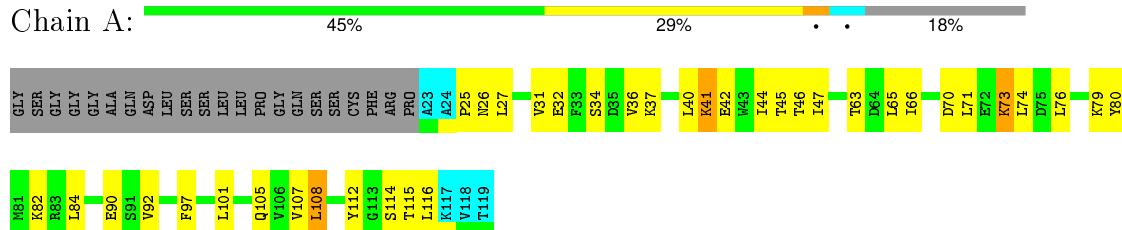


- Molecule 2: DNA polymerase kappa

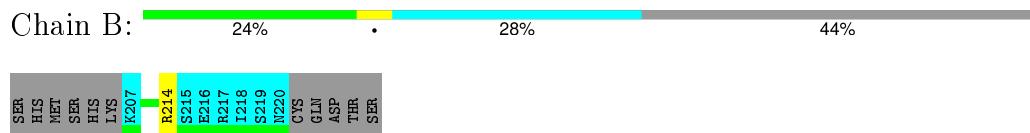


#### 4.2.3 Score per residue for model 3

- Molecule 1: DNA repair protein REV1

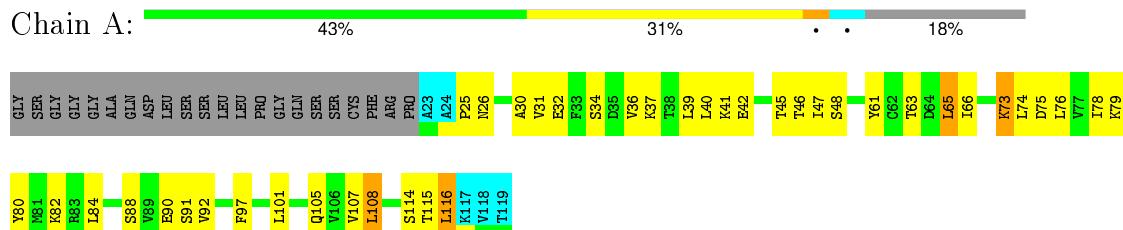


- Molecule 2: DNA polymerase kappa

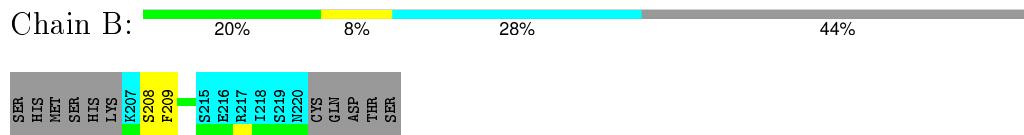


#### 4.2.4 Score per residue for model 4

- Molecule 1: DNA repair protein REV1

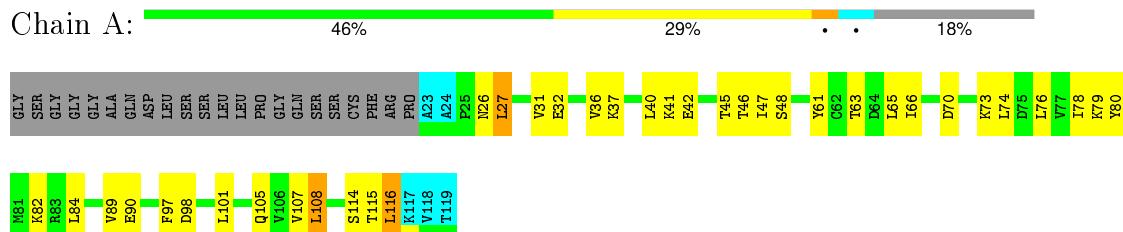


- Molecule 2: DNA polymerase kappa

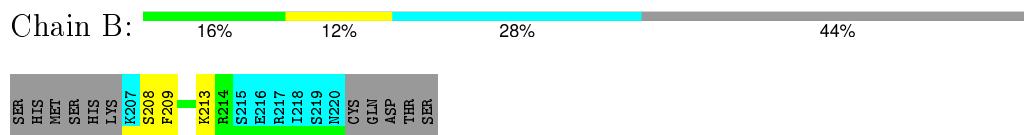


#### 4.2.5 Score per residue for model 5

- Molecule 1: DNA repair protein REV1

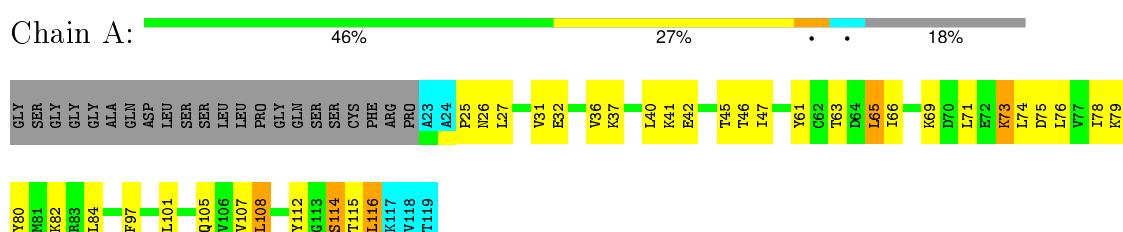


- Molecule 2: DNA polymerase kappa

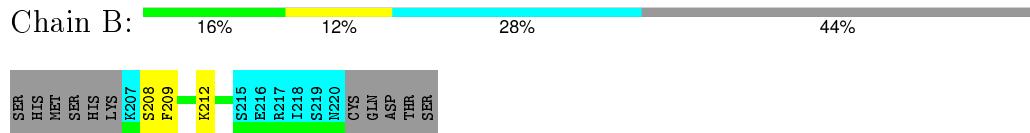


#### 4.2.6 Score per residue for model 6

- Molecule 1: DNA repair protein REV1

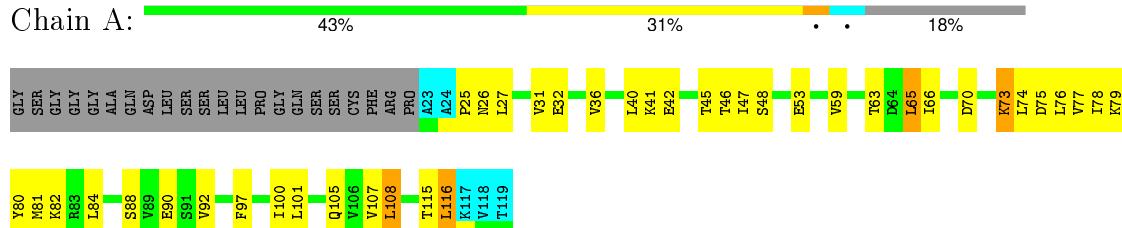


- Molecule 2: DNA polymerase kappa

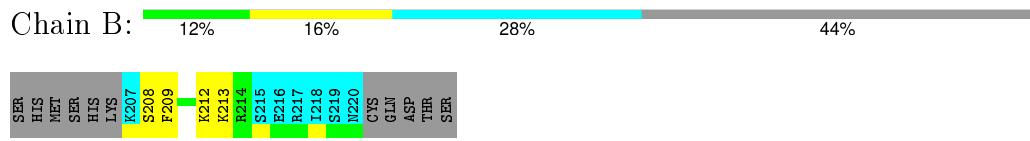


#### 4.2.7 Score per residue for model 7

- Molecule 1: DNA repair protein REV1

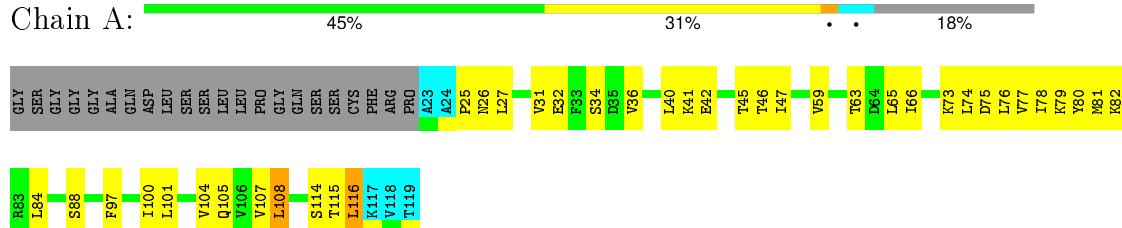


- Molecule 2: DNA polymerase kappa

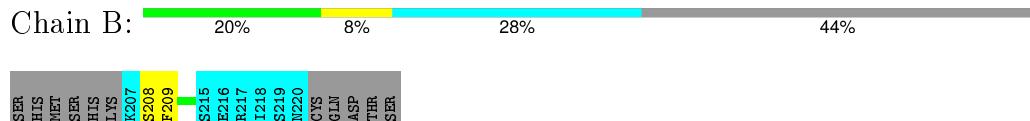


#### 4.2.8 Score per residue for model 8

- Molecule 1: DNA repair protein REV1



- Molecule 2: DNA polymerase kappa



#### 4.2.9 Score per residue for model 9

- Molecule 1: DNA repair protein REV1

Chain A: GLY SER GLY GLY GLY GLN ASP ILEU SER SER ILEU LEU PRO GLY GLN SER SER ARG PRO AC23 AC24 F25 M26 L27 AC28 Y34 K37 T38 L39 L40 K41 E42 T45 Y46 D44 R50 T53 D54 L55 K56 V57 K73 L74 D75 L76 V77 I78 K79 Y80 K82

The diagram shows the primary sequence of Chain A with color-coded segments: blue for the first 47%, yellow for the next 27%, orange for the final 18%. Specific residues are highlighted in green: AC23, AC24, F25, M26, L27, AC28, Y34, K37, T38, L39, L40, K41, E42, T45, Y46, D44, R50, T53, D54, L55, K56, V57, K73, L74, D75, L76, V77, I78, K79, Y80, and K82.

- Molecule 2: DNA polymerase kappa

Sequence logo for Chain B showing the distribution of amino acids at each position. The x-axis represents positions 1 through 20. The y-axis shows the probability of each amino acid (SER, HIS, MET, THR, ASP, GLN, CYS).

Position	SER	HIS	MET	THR	ASP	GLN	CYS
1	16%						
2		8%					
3			•				
4				28%			
5							44%
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							

#### 4.2.10 Score per residue for model 10

- Molecule 1: DNA repair protein REV1

- Molecule 2: DNA polymerase kappa

Chain B: 20% 8% 28% 44%

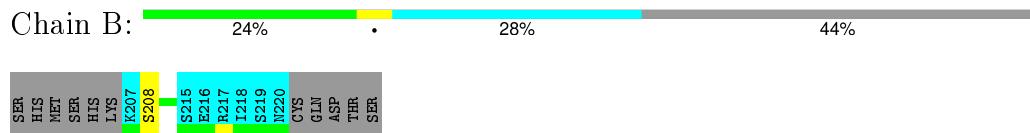
SER	HIS	MET	SER	HIS	K207	S208	F209	K212	S215	E216	R217	I218	S219	N220	CYS	GLN	ASP	THR	SER
-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----

#### 4.2.11 Score per residue for model 11

- Molecule 1: DNA repair protein REV1

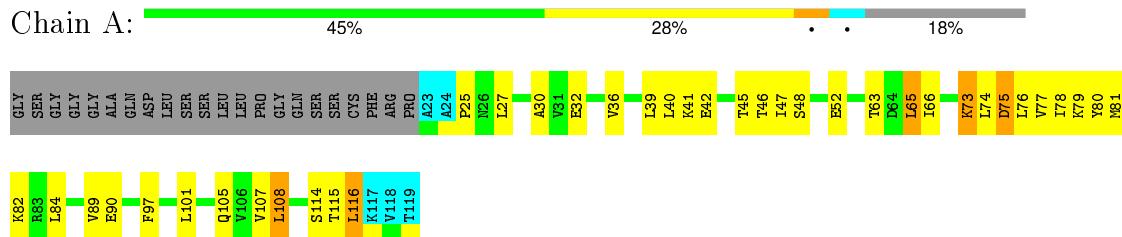
Sequence logo for Chain A showing amino acid probabilities at each position. The x-axis represents the sequence positions, and the y-axis represents the probability of each amino acid (GLY, SER, GLY, ASP, GLY, GLY, GLY, GLN, GLY, SER, SER, LEU, PRO, CYS, PHE, ARG, PRO, ILE, L23, L24, P25, N26, V31, E32, V36, K37, I40, K41, E42, T45, T46, D49, I55, V58, V59, H60, Y61, Q62, T63, I66, L71, E72, R73, K74, D75, L76, V77). The logo uses a color scheme where green represents GLY, blue represents SER, yellow represents GLY, red represents ASP, orange represents GLY, purple represents GLY, pink represents GLN, brown represents SER, light blue represents SER, grey represents LEU, dark blue represents PRO, cyan represents CYS, magenta represents PHE, dark red represents ARG, light red represents PRO, light blue represents ILE, light green represents L23, light blue represents L24, light green represents P25, light blue represents N26, light green represents V31, light green represents E32, light green represents V36, light blue represents K37, light green represents I40, light blue represents K41, light green represents E42, light green represents T45, light blue represents T46, light green represents D49, light green represents I55, light green represents V58, light blue represents V59, light green represents H60, light green represents Y61, light green represents Q62, light blue represents T63, light green represents I66, light green represents L71, light green represents E72, light blue represents R73, light green represents K74, light green represents D75, light green represents L76, light green represents V77.

- Molecule 2: DNA polymerase kappa

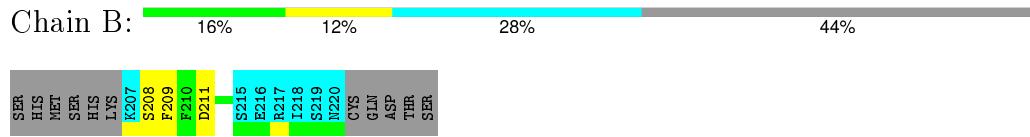


#### 4.2.12 Score per residue for model 12 (medoid)

- Molecule 1: DNA repair protein REV1

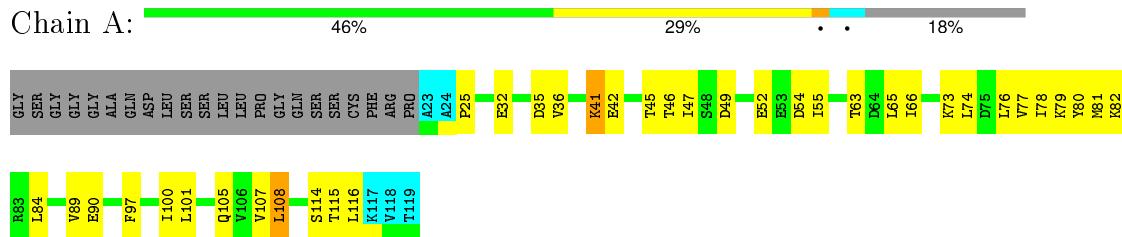


- Molecule 2: DNA polymerase kappa

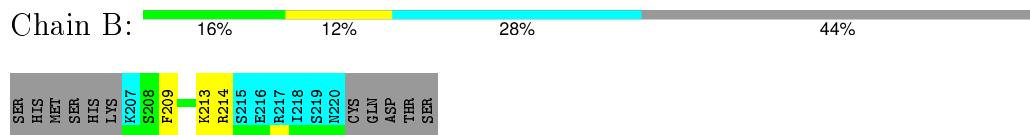


#### 4.2.13 Score per residue for model 13

- Molecule 1: DNA repair protein REV1



- Molecule 2: DNA polymerase kappa

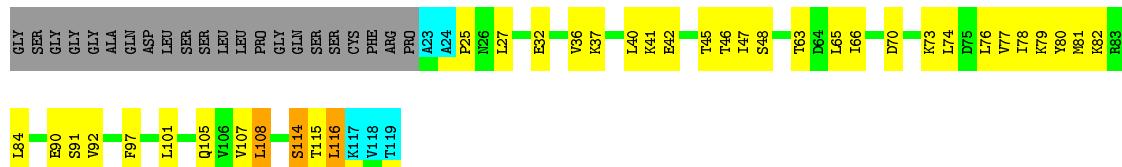


#### 4.2.14 Score per residue for model 14

- Molecule 1: DNA repair protein REV1

A horizontal bar chart titled "Chain A" showing its distribution across four categories. The categories are represented by colored segments of a bar: green (46%), yellow (29%), orange (18%), and cyan (18%).

Category	Percentage
Green	46%
Yellow	29%
Orange	18%
Cyan	18%



- Molecule 2: DNA polymerase kappa

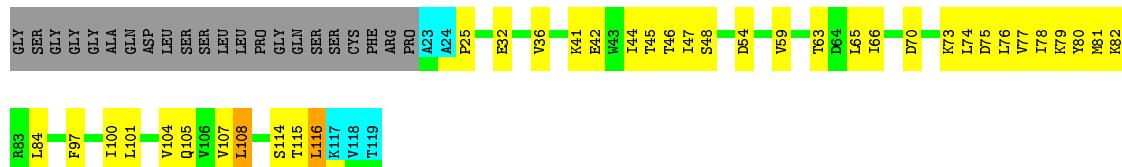
Chain B:  20% 8% 28% 44%



#### 4.2.15 Score per residue for model 15

- Molecule 1: DNA repair protein REV1

A horizontal progress bar for Chain A. The bar is divided into four colored segments: green (46%), yellow (29%), orange (3%), and grey (18%).



- Molecule 2: DNA polymerase kappa

A horizontal bar chart titled "Chain B" showing the percentage distribution across four categories. The categories are represented by green bars with black outlines. The percentages are labeled above each bar: 16% (shortest bar), 12% (second bar from left), 28% (third bar from left), and 44% (longest bar). The total length of the bars is 100%.

Category	Percentage
1	16%
2	12%
3	28%
4	44%

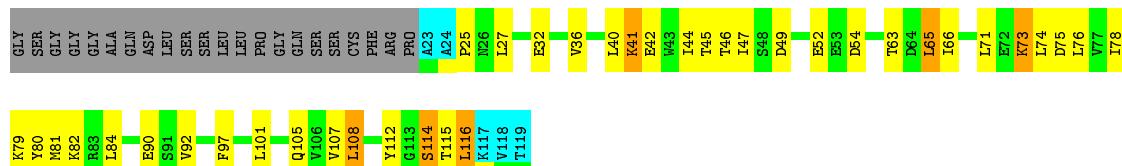


#### 4.2.16 Score per residue for model 16

- Molecule 1: DNA repair protein REV1

Chain A: 100%

A horizontal progress bar consisting of a green segment followed by a yellow segment. The green segment is labeled "45%" and the yellow segment is labeled "28%". To the right of the yellow segment is a small orange segment labeled "5%", followed by a teal segment with a white dot labeled "•", and finally a light blue segment labeled "18%".



- Molecule 2: DNA polymerase kappa

Chain B:  24% : 28% 44%

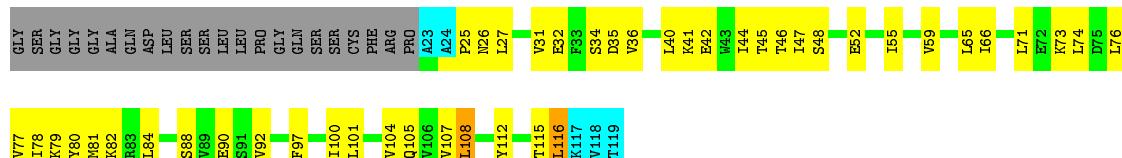


#### 4.2.17 Score per residue for model 17

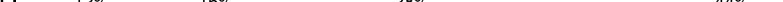
- Molecule 1: DNA repair protein REV1

A horizontal bar chart titled "Chain A" showing its distribution across four categories. The categories are represented by colored segments: red (leftmost), green, yellow, and blue (rightmost). The total length of the bar is 100%.

Category	Percentage
Red	39%
Green	36%
Yellow	• •
Blue	18%



- Molecule 2: DNA polymerase kappa

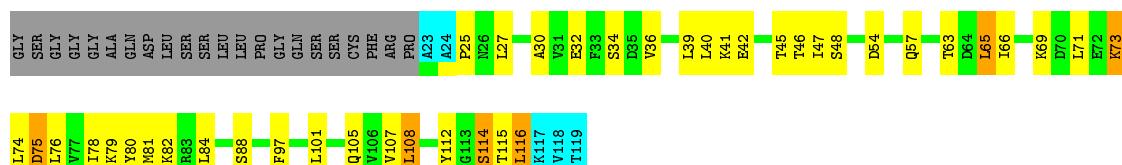
Chain B:  12% 16% 28% 44%



#### 4.2.18 Score per residue for model 18

- Molecule 1: DNA repair protein REV1

Chain A: 43% 29% 5% • 18%



- Molecule 2: DNA polymerase kappa

Chain B:  24% 2% 28% 44%



#### 4.2.19 Score per residue for model 19

- Molecule 1: DNA repair protein REV1

- Molecule 2: DNA polymerase kappa

#### 4.2.20 Score per residue for model 20

- Molecule 1: DNA repair protein REV1

Chain A: 43% 33% • • 18%

M81 K82 L84 S88 V89 E90 S91 V92 F97 I100 L104 V104 S114 T115 L116 K117 T118 T119 Y41 T42 T43 T44 T45 T46 T47 S48 D52 D53 D54 D55 L56 V59 T63 D64 L65 T66 K73 L74 D75 L76 T77 L78 T79 K79 T80

- Molecule 2: DNA polymerase kappa

## 5 Refinement protocol and experimental data overview i

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
CYANA	structure solution	
CYANA	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)	2lsj_cs.str
Number of chemical shift lists	1
Total number of shifts	1552
Number of shifts mapped to atoms	1552
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	84%

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	748	755	755	26±3
2	B	65	66	66	1±1
All	All	16260	16420	16420	523

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 16.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:78:ILE:HG21	1:A:116:LEU:HD22	0.78	1.54	2	16
1:A:97:PHE:CZ	1:A:101:LEU:HD11	0.71	2.21	7	20
1:A:42:GLU:O	1:A:46:THR:HG22	0.68	1.89	9	20
1:A:63:THR:HG22	1:A:107:VAL:HG11	0.67	1.66	9	15
1:A:28:ALA:HB3	1:A:39:LEU:HD13	0.67	1.67	9	2
1:A:108:LEU:HD21	1:A:115:THR:HA	0.66	1.67	13	20
1:A:65:LEU:HD23	1:A:74:LEU:HB2	0.65	1.68	7	19
1:A:71:LEU:HD11	1:A:112:TYR:CE2	0.64	2.28	18	6
1:A:32:GLU:O	1:A:36:VAL:HG23	0.64	1.93	2	18
1:A:65:LEU:HD11	1:A:73:LYS:HE2	0.61	1.71	12	5
1:A:105:GLN:HA	1:A:108:LEU:HD23	0.61	1.71	12	20
1:A:80:TYR:CE1	1:A:84:LEU:HD11	0.59	2.32	16	20
1:A:28:ALA:HB3	1:A:39:LEU:CD1	0.58	2.28	9	1

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:90:GLU:HB3	1:A:92:VAL:HG12	0.58	1.75	4	8
1:A:39:LEU:HD22	2:B:213:LYS:CE	0.58	2.27	9	1
1:A:44:ILE:HG22	1:A:44:ILE:O	0.58	1.98	3	6
1:A:71:LEU:HD11	1:A:112:TYR:CZ	0.58	2.34	18	7
1:A:37:LYS:HA	1:A:40:LEU:HD12	0.57	1.75	3	10
1:A:108:LEU:HD12	1:A:108:LEU:O	0.57	2.00	17	10
1:A:108:LEU:O	1:A:108:LEU:HD12	0.57	2.00	7	10
1:A:36:VAL:HG21	1:A:61:TYR:OH	0.56	2.01	6	3
1:A:73:LYS:HA	1:A:76:LEU:HD12	0.56	1.78	1	20
1:A:47:ILE:HD11	2:B:209:PHE:CD1	0.56	2.35	8	15
1:A:75:ASP:HA	1:A:116:LEU:HD21	0.56	1.77	12	3
1:A:26:ASN:ND2	1:A:31:VAL:HG22	0.55	2.17	2	12
1:A:73:LYS:O	1:A:77:VAL:HG23	0.54	2.03	17	4
1:A:41:LYS:O	1:A:45:THR:HG23	0.53	2.04	9	20
1:A:108:LEU:HD11	1:A:114:SER:O	0.53	2.04	20	18
1:A:27:LEU:HD21	1:A:40:LEU:HD21	0.52	1.80	7	8
1:A:41:LYS:HA	1:A:84:LEU:HD13	0.52	1.82	16	20
1:A:28:ALA:HB1	2:B:214:ARG:HB3	0.51	1.82	20	1
1:A:108:LEU:HD12	1:A:108:LEU:C	0.51	2.26	17	10
1:A:97:PHE:CE2	1:A:101:LEU:HD11	0.51	2.40	18	12
1:A:108:LEU:C	1:A:108:LEU:HD12	0.50	2.27	7	10
1:A:52:GLU:HA	1:A:55:ILE:HD12	0.50	1.83	17	1
1:A:77:VAL:HG12	1:A:81:MET:CE	0.50	2.36	8	12
1:A:63:THR:HG22	1:A:107:VAL:CG1	0.49	2.36	6	13
1:A:75:ASP:HA	1:A:116:LEU:HD11	0.49	1.82	18	1
1:A:55:ILE:O	1:A:58:VAL:HG22	0.49	2.08	11	1
1:A:90:GLU:HB3	1:A:92:VAL:HG22	0.48	1.85	2	2
1:A:55:ILE:HG23	1:A:100:ILE:HD11	0.48	1.85	13	1
1:A:89:VAL:HG13	1:A:90:GLU:CD	0.48	2.28	12	3
1:A:66:ILE:HD13	1:A:107:VAL:HG13	0.47	1.86	8	20
1:A:28:ALA:HB2	2:B:210:PHE:HB3	0.47	1.86	19	1
1:A:46:THR:HG23	1:A:47:ILE:HD12	0.47	1.86	17	6
1:A:59:VAL:HG22	1:A:100:ILE:HA	0.47	1.87	8	7
1:A:100:ILE:O	1:A:104:VAL:HG23	0.46	2.11	10	6
1:A:36:VAL:HG11	1:A:61:TYR:CE2	0.46	2.46	11	1
1:A:39:LEU:HD22	2:B:213:LYS:HE2	0.46	1.87	9	1
1:A:46:THR:HG23	1:A:47:ILE:CD1	0.45	2.42	3	7
1:A:78:ILE:CG2	1:A:116:LEU:HD22	0.45	2.35	2	6
1:A:27:LEU:HD12	1:A:61:TYR:CG	0.45	2.46	5	1
1:A:27:LEU:HD22	1:A:36:VAL:HG13	0.44	1.89	14	3
1:A:30:ALA:HB2	1:A:39:LEU:HD11	0.44	1.90	19	5

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:65:LEU:HD21	1:A:73:LYS:HG3	0.43	1.88	19	2
1:A:78:ILE:HD13	1:A:81:MET:CE	0.43	2.43	18	7
1:A:59:VAL:O	1:A:63:THR:HG23	0.43	2.13	8	1
1:A:65:LEU:HD21	1:A:73:LYS:CG	0.43	2.43	12	2
1:A:39:LEU:HD22	2:B:213:LYS:NZ	0.42	2.29	9	1
1:A:52:GLU:O	1:A:56:LEU:HD23	0.42	2.14	20	1
1:A:39:LEU:HD22	2:B:213:LYS:HZ1	0.42	1.73	9	1
1:A:63:THR:HG23	1:A:104:VAL:HG22	0.42	1.91	11	2
1:A:71:LEU:HD11	1:A:112:TYR:CE1	0.42	2.49	3	1
1:A:44:ILE:O	1:A:44:ILE:HG22	0.42	2.14	16	2
1:A:63:THR:CG2	1:A:107:VAL:HG11	0.42	2.41	9	4
1:A:59:VAL:HG22	1:A:100:ILE:HG23	0.42	1.90	11	1
1:A:26:ASN:CG	1:A:31:VAL:HG22	0.41	2.35	9	1
1:A:65:LEU:HD23	1:A:74:LEU:CB	0.41	2.43	4	3
1:A:65:LEU:HD11	1:A:73:LYS:CE	0.41	2.44	7	1
1:A:27:LEU:HD13	1:A:61:TYR:CG	0.41	2.50	6	1
1:A:78:ILE:HG21	1:A:116:LEU:CD2	0.41	2.43	9	2
1:A:66:ILE:HG23	1:A:112:TYR:CE1	0.41	2.50	3	1
1:A:63:THR:HG22	1:A:104:VAL:HA	0.40	1.93	11	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	92/119 (77%)	89±1 (97±1%)	3±1 (3±1%)	0±0 (0±0%)	56 85
2	B	7/25 (28%)	6±0 (89±6%)	1±0 (11±6%)	0±0 (0±0%)	100 100
All	All	1980/2880 (69%)	1906 (96%)	70 (4%)	4 (0%)	56 85

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	88	SER	4

### 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	87/106 (82%)	78±2 (90±2%)	9±2 (10±2%)	14 59
2	B	7/25 (28%)	5±1 (77±11%)	2±1 (23±11%)	3 30
All	All	1880/2620 (72%)	1675 (89%)	205 (11%)	12 56

All 31 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	79	LYS	20
1	A	116	LEU	20
1	A	108	LEU	20
1	A	82	LYS	20
2	B	208	SER	14
1	A	48	SER	10
1	A	75	ASP	10
1	A	73	LYS	9
1	A	34	SER	8
2	B	213	LYS	8
1	A	65	LEU	7
1	A	54	ASP	6
1	A	70	ASP	6
1	A	114	SER	5
1	A	49	ASP	4
2	B	212	LYS	4
2	B	214	ARG	4
1	A	41	LYS	4
1	A	69	LYS	4
1	A	52	GLU	4
1	A	88	SER	4
1	A	91	SER	2
1	A	60	ARG	2
2	B	211	ASP	2
1	A	35	ASP	2
1	A	57	GLN	1
1	A	83	ARG	1
1	A	27	LEU	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	102	ASP	1
1	A	98	ASP	1
1	A	53	GLU	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 84% for the well-defined parts and 84% for the entire structure.

### 7.1 Chemical shift list 1

File name: 2lsj\_cs.str

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	1552
Number of shifts mapped to atoms	1552
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

#### 7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	133	-0.57 $\pm$ 0.17	Should be applied
$^{13}\text{C}_\beta$	128	0.30 $\pm$ 0.06	None needed (< 0.5 ppm)
$^{13}\text{C}'$	0	—	—
$^{15}\text{N}$	121	0.32 $\pm$ 0.16	None needed (< 0.5 ppm)

#### 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 84%, i.e. 1070 atoms were assigned a chemical shift out of a possible 1268. 22 out of 22 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	382/491 (78%)	191/196 (97%)	99/198 (50%)	92/97 (95%)
Sidechain	611/684 (89%)	375/396 (95%)	227/262 (87%)	9/26 (35%)

*Continued on next page...*

*Continued from previous page...*

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	77/93 (83%)	49/49 (100%)	26/42 (62%)	2/2 (100%)
Overall	1070/1268 (84%)	615/641 (96%)	352/502 (70%)	103/125 (82%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 84%, i.e. 1184 atoms were assigned a chemical shift out of a possible 1415. 23 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	428/551 (78%)	214/220 (97%)	111/222 (50%)	103/109 (94%)
Sidechain	679/771 (88%)	418/447 (94%)	252/292 (86%)	9/32 (28%)
Aromatic	77/93 (83%)	49/49 (100%)	26/42 (62%)	2/2 (100%)
Overall	1184/1415 (84%)	681/716 (95%)	389/556 (70%)	114/143 (80%)

#### 7.1.4 Statistically unusual chemical shifts [\(i\)](#)

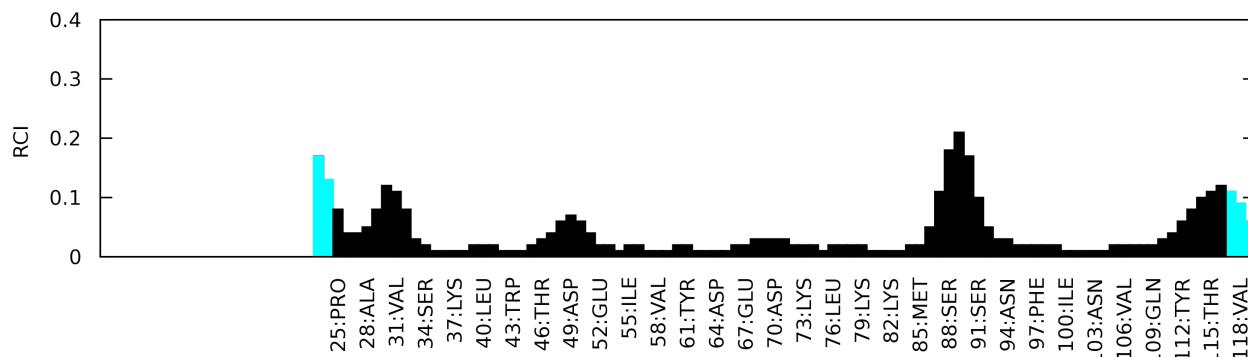
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	50	PRO	HG2	-0.19	3.48 – 0.38	-6.8
1	A	93	TRP	HH2	4.37	8.94 – 5.04	-6.7

#### 7.1.5 Random Coil Index (RCI) plots [\(i\)](#)

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:

