W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB NMR Structure Validation Report ⓘ

Feb 19, 2022 – 04:42 PM EST

PDB ID	:	1RL1
Title	:	Solution structure of human Sgt1 CS domain
Authors	:	Lee, Y.-T.; Jacob, J.; Michowski, W.; Nowotny, M.; Kuznicki, J.; Chazin, W.J.
Deposited on	:	2003-11-24

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 ⓘ symbol.

The following versions of software and data (see references ⓘ) 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)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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 was not calculated.

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.
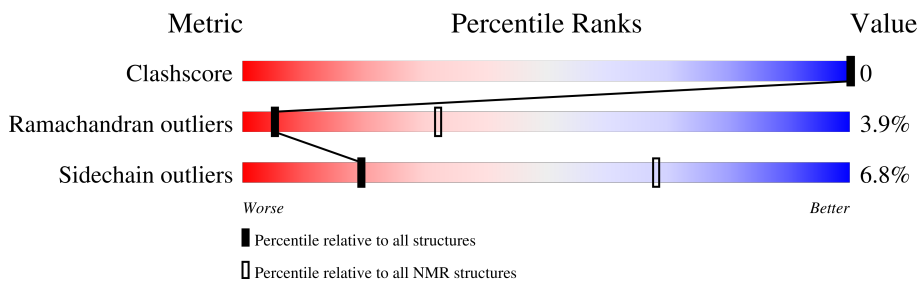
Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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	114	71% 6% · 19%

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 18 is the overall representative, medoid model (most similar to other models). The authors have identified model 19 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:140-A:227 (88)	0.43	18

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 6 clusters and 5 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1396 atoms, of which 644 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Suppressor of G2 allele of SKP1 homolog.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	92	1396	487	644	121	143	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	131	GLY	-	cloning artifact	UNP Q9Y2Z0
A	132	SER	-	cloning artifact	UNP Q9Y2Z0
A	133	HIS	-	cloning artifact	UNP Q9Y2Z0
A	134	MET	-	cloning artifact	UNP Q9Y2Z0

4 Residue-property plots ⓘ

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.

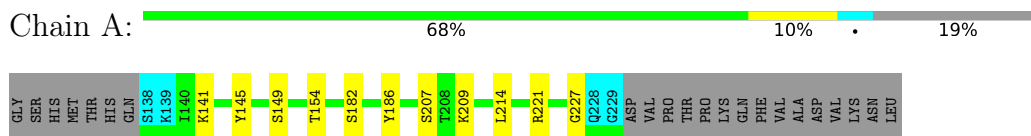• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:
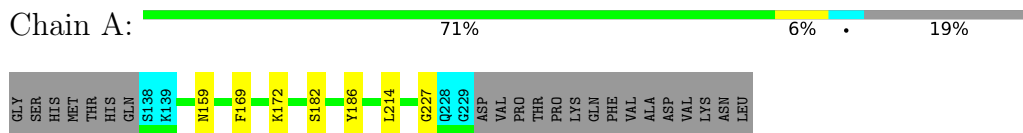
4.2 Scores per residue for each member of the ensemble
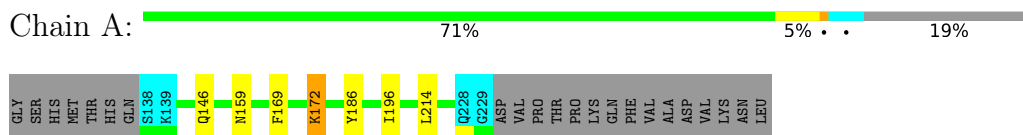
Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1
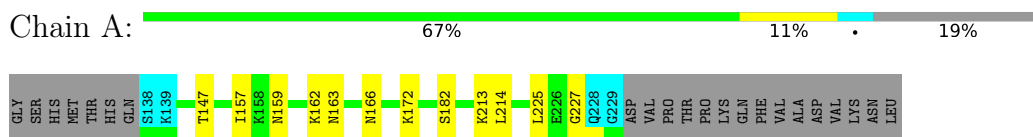
• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

4.2.2 Score per residue for model 2

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

4.2.3 Score per residue for model 3

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A: 67% 11% · 19%

4.2.4 Score per residue for model 4

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A: 70% 7% · 19%

4.2.5 Score per residue for model 5

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A: 70% 6% · · 19%

4.2.6 Score per residue for model 6

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A: 72% · · · 19%

4.2.7 Score per residue for model 7

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A: 66% 10% · · 19%

4.2.8　Score per residue for model 8

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

4.2.9　Score per residue for model 9
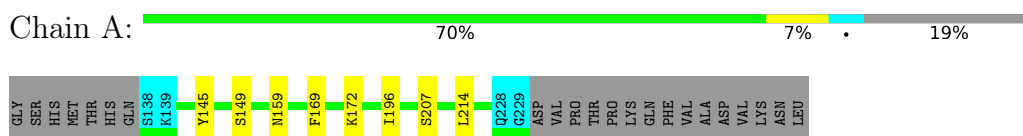
• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

4.2.10　Score per residue for model 10

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

4.2.11　Score per residue for model 11

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

4.2.12　Score per residue for model 12

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

4.2.13 Score per residue for model 13

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

4.2.14 Score per residue for model 14

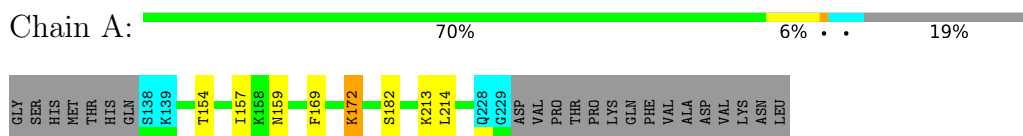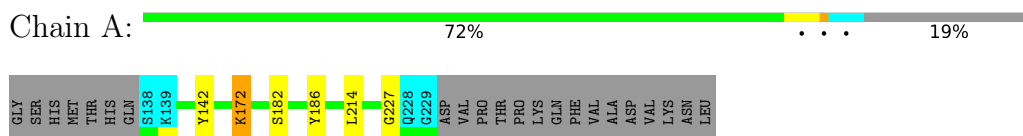• Molecule 1: Suppressor of G2 allele of SKP1 homolog

4.2.15 Score per residue for model 15

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

4.2.16 Score per residue for model 16

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

4.2.17 Score per residue for model 17

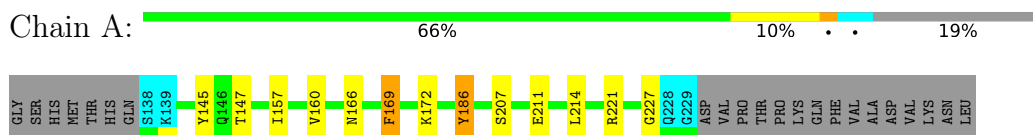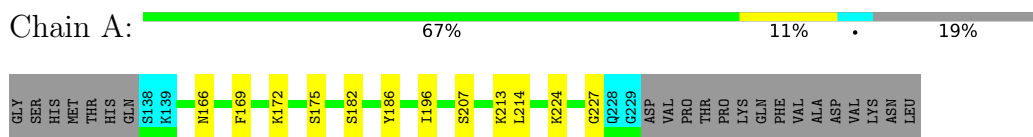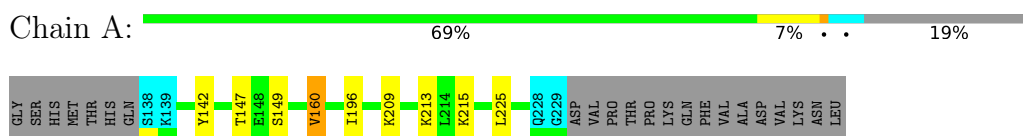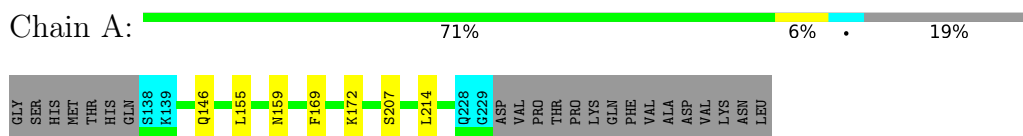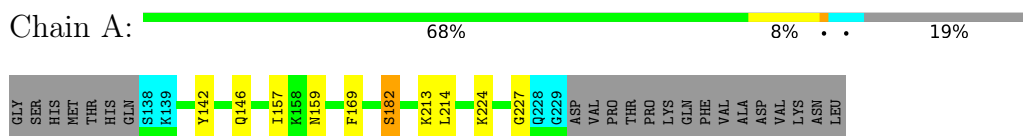• Molecule 1: Suppressor of G2 allele of SKP1 homolog

4.2.18 Score per residue for model 18 (medoid)

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

68% 8% · · 19%

GLY SER HIS MET THR HIS GLN **S138 K139** L155 N163 F169 K172 S182 Y186 K213 L214 E226 G227 **Q228 G229** ASP VAL PRO THR PRO LYS GLN PHE VAL ALA ASP VAL LYS ASN LEU
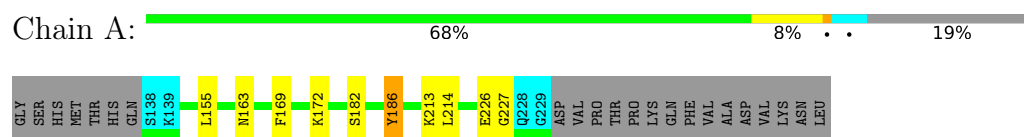
4.2.19 Score per residue for model 19

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

66% 11% · 19%

GLY SER HIS MET THR HIS GLN **S138 K139** S149 I157 **K158 N159** F169 S182 I196 L206 L214 K215 K216 P217 R221 **W222** E223 G227 **Q228 G229** ASP VAL PRO THR PRO LYS GLN PHE VAL ALA ASP VAL LYS ASN LEU
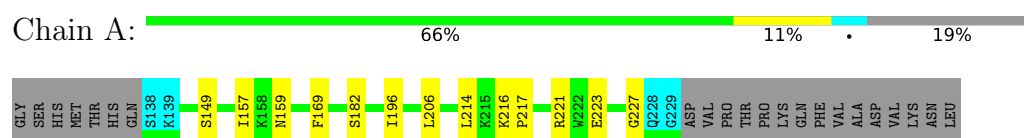
4.2.20 Score per residue for model 20

• Molecule 1: Suppressor of G2 allele of SKP1 homolog

Chain A:

68% 8% · · 19%

GLY SER HIS MET THR HIS GLN **S138 K139** I157 **K158 N159** N163 N166 K172 K179 D185 **Y186** L214 G227 **Q228 G229** ASP VAL PRO THR PRO LYS GLN PHE VAL ALA ASP VAL LYS ASN LEU
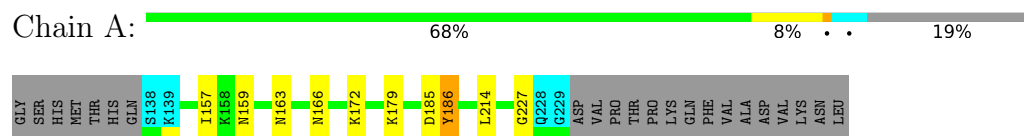
5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *Distance Geometry and restrained Molecular Dynamics.*

Of the 50 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
DYANA	structure solution	1.5
Amber	refinement	7

No chemical shift data was provided.

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	0.53±0.00	0±0/738 (0.0± 0.0%)	0.90±0.02	0±1/998 (0.0± 0.1%)
All	All	0.53	0/14760 (0.0%)	0.90	5/19960 (0.0%)

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	1.4±1.0
All	All	0	28

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	186	TYR	CB-CG-CD2	-7.09	116.74	121.00	18	3
1	A	221	ARG	NE-CZ-NH2	-5.41	117.60	120.30	1	1
1	A	142	TYR	CB-CG-CD2	-5.28	117.83	121.00	6	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	169	PHE	Sidechain	14
1	A	186	TYR	Sidechain	7
1	A	145	TYR	Sidechain	5
1	A	142	TYR	Sidechain	2

6.2 Too-close contacts ⓘ

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	724	622	756	0±0
All	All	14480	12440	15120	1

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:142:TYR:CE1	1:A:153:ILE:CG2	0.40	3.05	13	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	88/114 (77%)	75±2 (86±2%)	9±2 (10±2%)	3±2 (4±2%)	5	32
All	All	1760/2280 (77%)	1508 (86%)	184 (10%)	68 (4%)	5	32

All 12 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	227	GLY	13
1	A	172	LYS	13
1	A	182	SER	9
1	A	157	ILE	8
1	A	159	ASN	7
1	A	207	SER	6
1	A	149	SER	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	225	LEU	3
1	A	196	ILE	2
1	A	160	VAL	1
1	A	223	GLU	1
1	A	217	PRO	1

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	84/107 (79%)	78±2 (93±2%)	6±2 (7±2%)	19	68
All	All	1680/2140 (79%)	1565 (93%)	115 (7%)	19	68

All 38 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	214	LEU	19
1	A	213	LYS	10
1	A	159	ASN	6
1	A	172	LYS	5
1	A	163	ASN	5
1	A	186	TYR	4
1	A	146	GLN	4
1	A	196	ILE	4
1	A	166	ASN	4
1	A	221	ARG	4
1	A	215	LYS	4
1	A	155	LEU	4
1	A	209	LYS	3
1	A	147	THR	3
1	A	224	LYS	3
1	A	182	SER	3
1	A	216	LYS	3
1	A	141	LYS	2
1	A	154	THR	2
1	A	162	LYS	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	160	VAL	2
1	A	142	TYR	2
1	A	206	LEU	2
1	A	145	TYR	1
1	A	169	PHE	1
1	A	211	GLU	1
1	A	175	SER	1
1	A	143	ASP	1
1	A	148	GLU	1
1	A	204	LYS	1
1	A	149	SER	1
1	A	201	SER	1
1	A	189	LYS	1
1	A	190	LEU	1
1	A	226	GLU	1
1	A	223	GLU	1
1	A	179	LYS	1
1	A	185	ASP	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 monosaccharides in this entry.

6.6   Ligand geometry ⓘ

There are no ligands in this entry.

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 (i)

No chemical shift data were provided