Full wwPDB NMR Structure Validation Report (i)

Mar 1, 2022 – 05:38 PM EST

PDB ID	:	2GQM
Title	:	Solution structure of Human Cu(I)-Sco1
Authors	:	Banci, L.; Bertini, I.; Calderone, V.; Ciofi-Baffoni, S.; Mangani, S.; Palumaa, P.; Martinelli, M.; Wang, S.; Structural Proteomics in Europe (SPINE)
Deposited on	:	2006-04-21

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 following versions of software and data (see references (i)) were used in the production of this report:

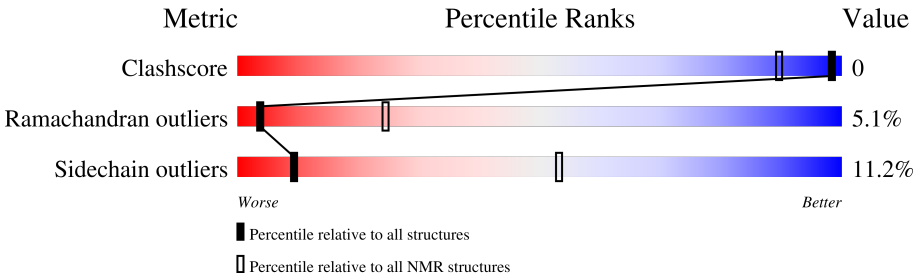
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 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	173	80% 8% • 11%

2 Ensemble composition and analysis (i)

This entry contains 30 models. Model 13 is the overall representative, medoid model (most similar to other models).

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:138-A:248, A:256-A:298 (154)	0.77	13

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	2, 5, 8, 12, 13, 14, 15, 16, 18, 19, 25, 29
2	1, 4, 9, 21, 27
3	3, 22
4	6, 7
5	26, 28
6	11, 23
Single-model clusters	10; 17; 20; 24; 30

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2750 atoms, of which 1360 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called SCO1 protein homolog, mitochondrial.

Mol	Chain	Residues	Atoms						Trace
1	A	173	Total	C	H	N	O	S	0
			2749	890	1360	224	270	5	

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	129	SER	-	cloning artifact	UNP O75880
A	130	PHE	-	cloning artifact	UNP O75880
A	131	THR	-	cloning artifact	UNP O75880

- Molecule 2 is COPPER (I) ION (three-letter code: CU1) (formula: Cu).

Mol	Chain	Residues	Atoms	
2	A	1	Total	Cu
			1	1

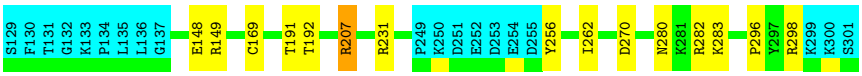
4 Residue-property plots

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: SCO1 protein homolog, mitochondrial

Chain A:



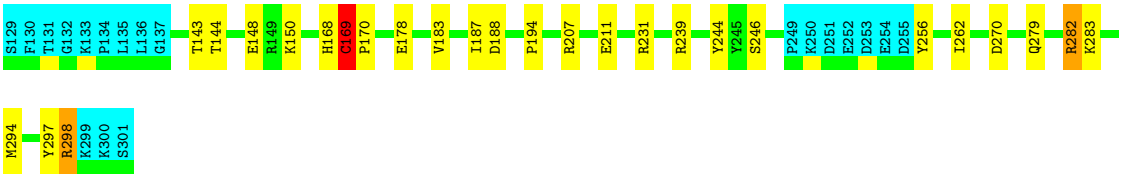
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: SCO1 protein homolog, mitochondrial

Chain A:



4.2.2 Score per residue for model 2

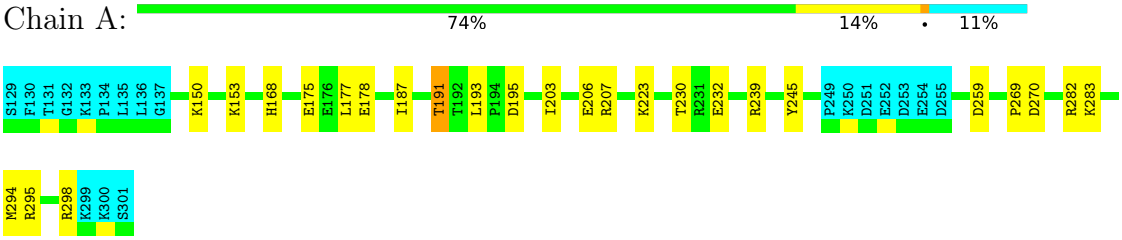
• Molecule 1: SCO1 protein homolog, mitochondrial

Chain A:

4.2.3 Score per residue for model 3

• Molecule 1: SCO1 protein homolog, mitochondrial

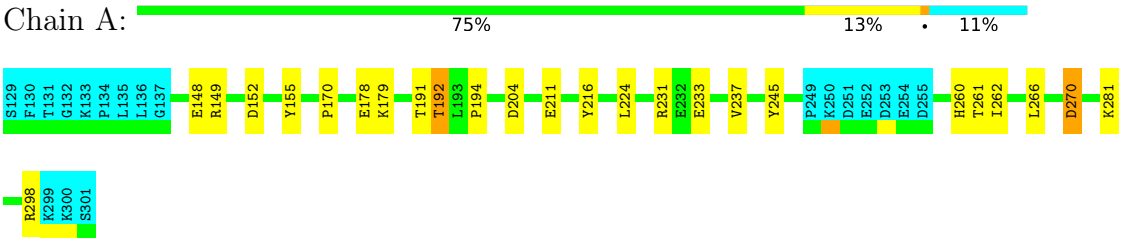
Chain A:



4.2.4 Score per residue for model 4

• Molecule 1: SCO1 protein homolog, mitochondrial

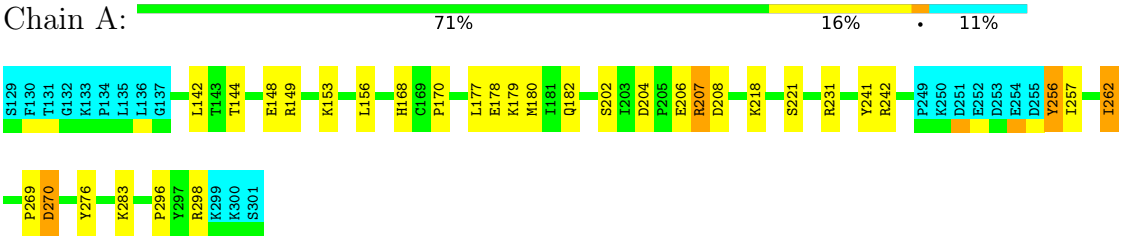
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: SCO1 protein homolog, mitochondrial

Chain A:



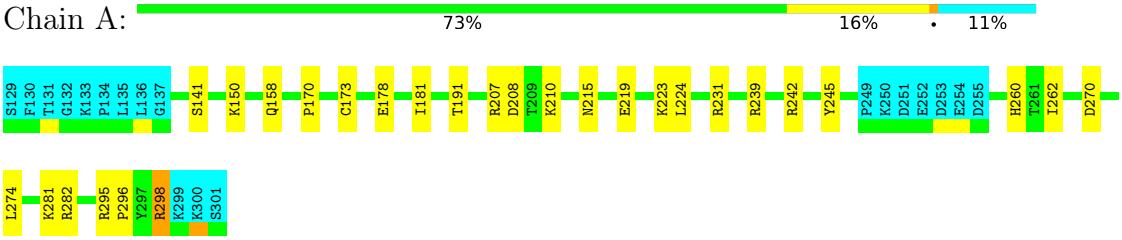
4.2.6 Score per residue for model 6

• Molecule 1: SCO1 protein homolog, mitochondrial

Chain A:

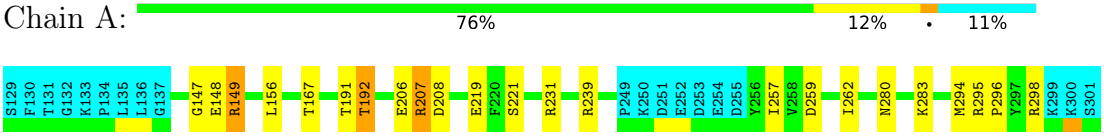
4.2.7 Score per residue for model 7

• Molecule 1: SCO1 protein homolog, mitochondrial



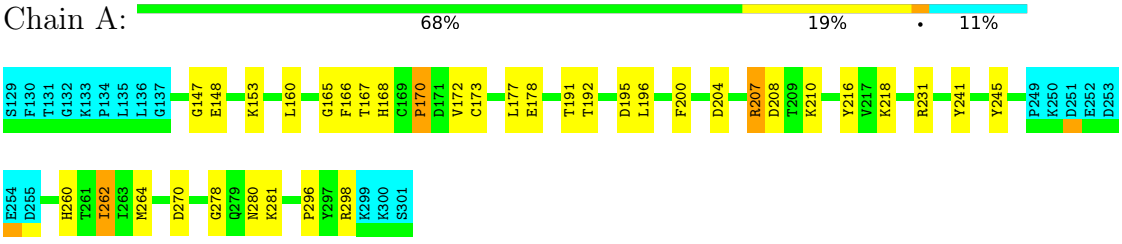
4.2.8 Score per residue for model 8

• Molecule 1: SCO1 protein homolog, mitochondrial



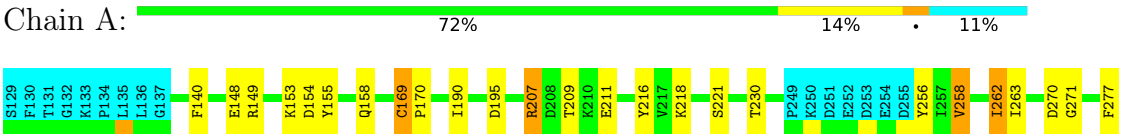
4.2.9 Score per residue for model 9

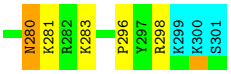
• Molecule 1: SCO1 protein homolog, mitochondrial



4.2.10 Score per residue for model 10

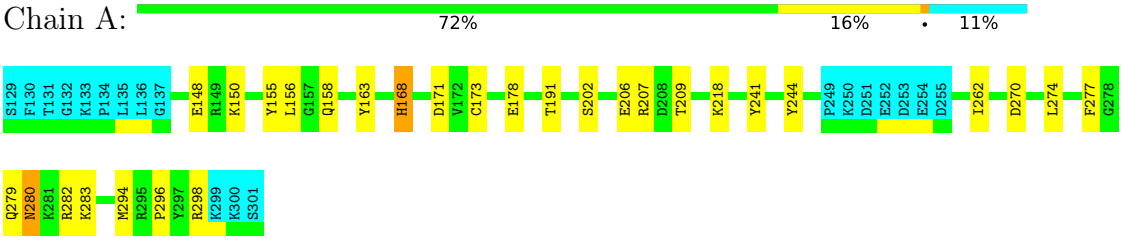
• Molecule 1: SCO1 protein homolog, mitochondrial





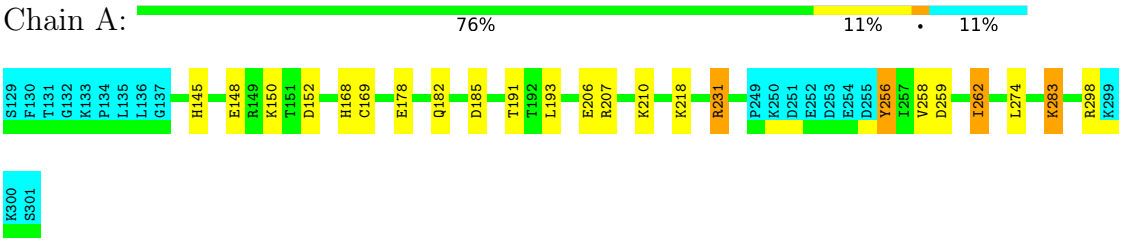
4.2.11 Score per residue for model 11

• Molecule 1: SCO1 protein homolog, mitochondrial



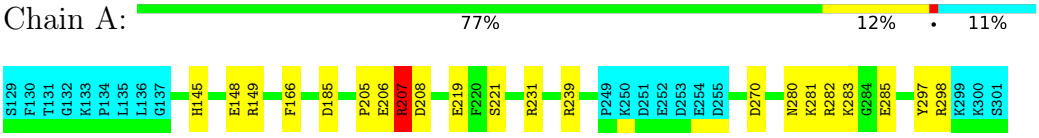
4.2.12 Score per residue for model 12

• Molecule 1: SCO1 protein homolog, mitochondrial



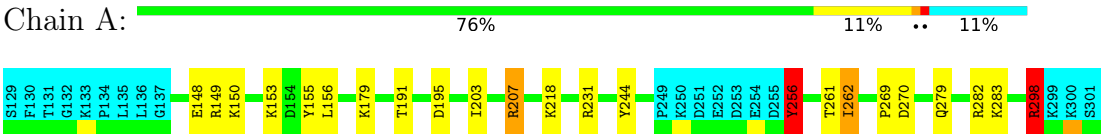
4.2.13 Score per residue for model 13 (medoid)

• Molecule 1: SCO1 protein homolog, mitochondrial



4.2.14 Score per residue for model 14

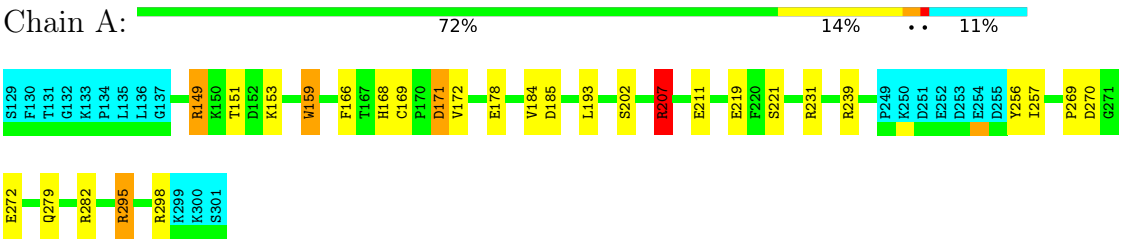
• Molecule 1: SCO1 protein homolog, mitochondrial



4.2.15 Score per residue for model 15

• Molecule 1: SCO1 protein homolog, mitochondrial

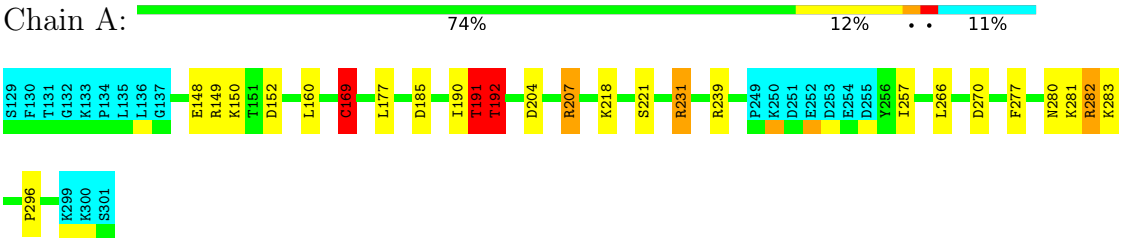
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: SCO1 protein homolog, mitochondrial

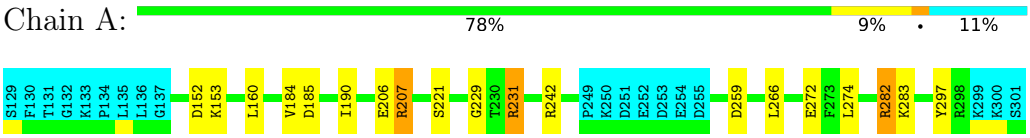
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: SCO1 protein homolog, mitochondrial

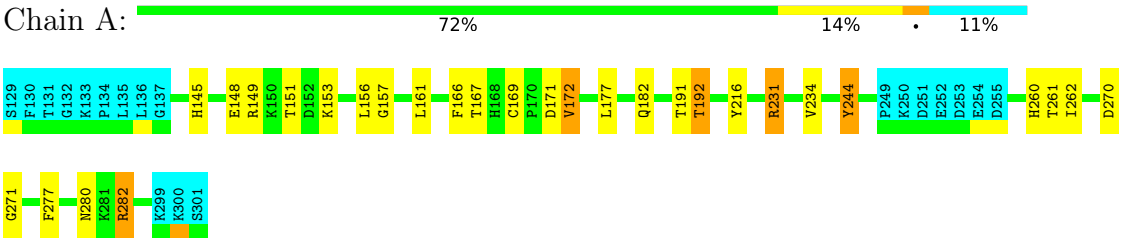
Chain A:



4.2.18 Score per residue for model 18

• Molecule 1: SCO1 protein homolog, mitochondrial

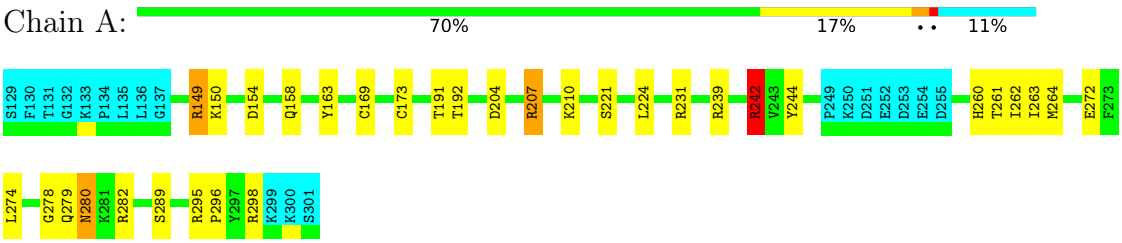
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: SCO1 protein homolog, mitochondrial

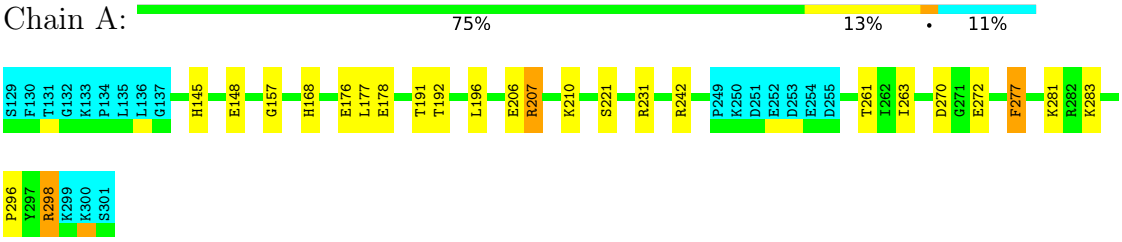
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: SCO1 protein homolog, mitochondrial

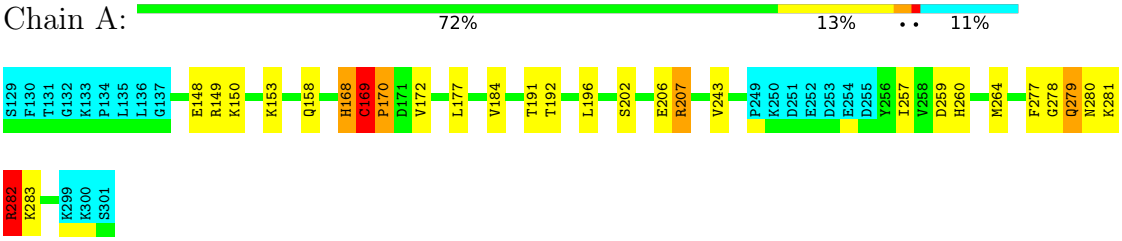
Chain A:



4.2.21 Score per residue for model 21

• Molecule 1: SCO1 protein homolog, mitochondrial

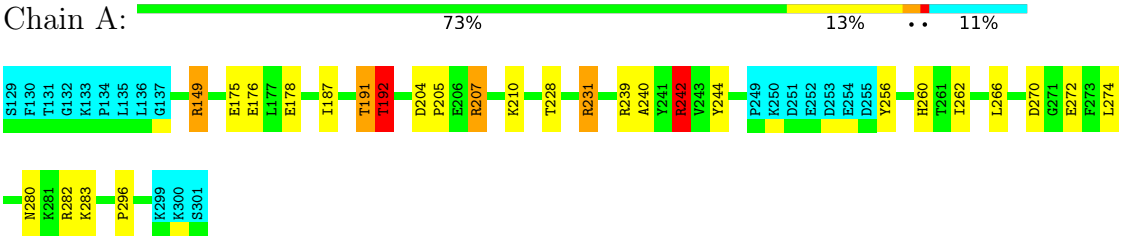
Chain A:



4.2.22 Score per residue for model 22

• Molecule 1: SCO1 protein homolog, mitochondrial

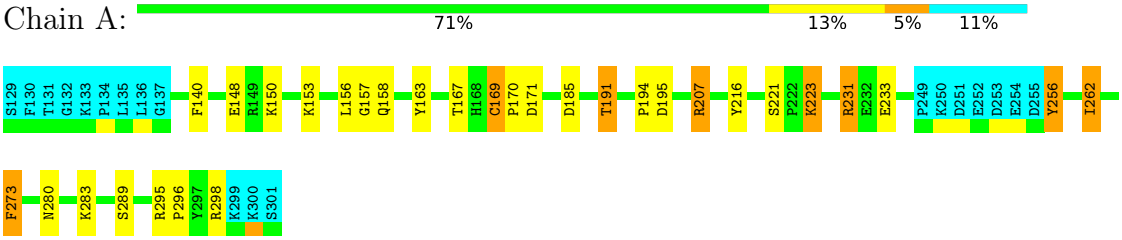
Chain A:



4.2.23 Score per residue for model 23

• Molecule 1: SCO1 protein homolog, mitochondrial

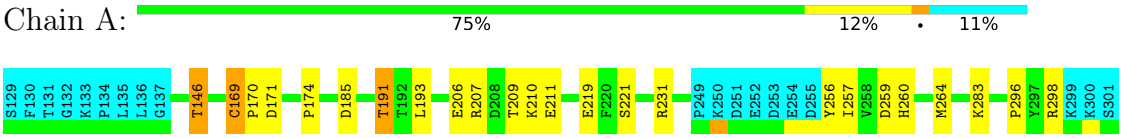
Chain A:



4.2.24 Score per residue for model 24

• Molecule 1: SCO1 protein homolog, mitochondrial

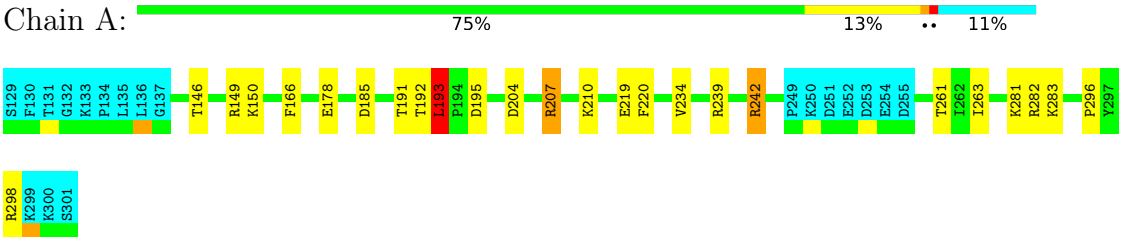
Chain A:



4.2.25 Score per residue for model 25

• Molecule 1: SCO1 protein homolog, mitochondrial

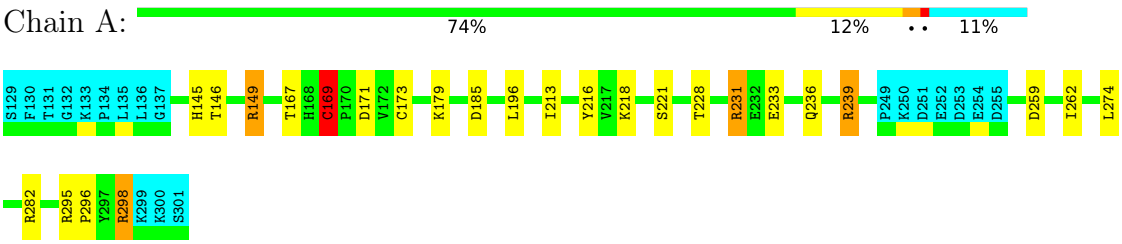
Chain A:



4.2.26 Score per residue for model 26

• Molecule 1: SCO1 protein homolog, mitochondrial

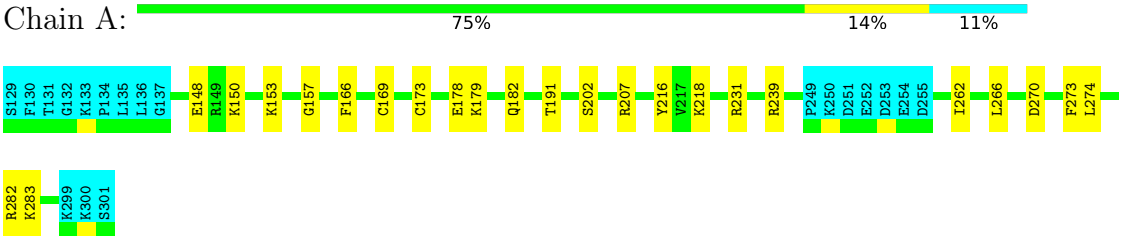
Chain A:



4.2.27 Score per residue for model 27

• Molecule 1: SCO1 protein homolog, mitochondrial

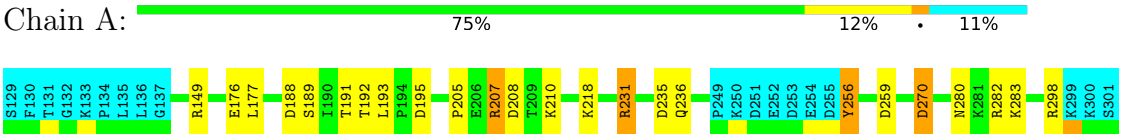
Chain A:



4.2.28 Score per residue for model 28

• Molecule 1: SCO1 protein homolog, mitochondrial

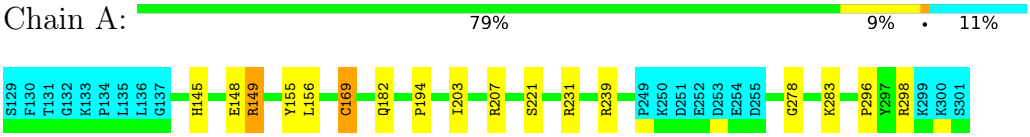
Chain A:



4.2.29 Score per residue for model 29

• Molecule 1: SCO1 protein homolog, mitochondrial

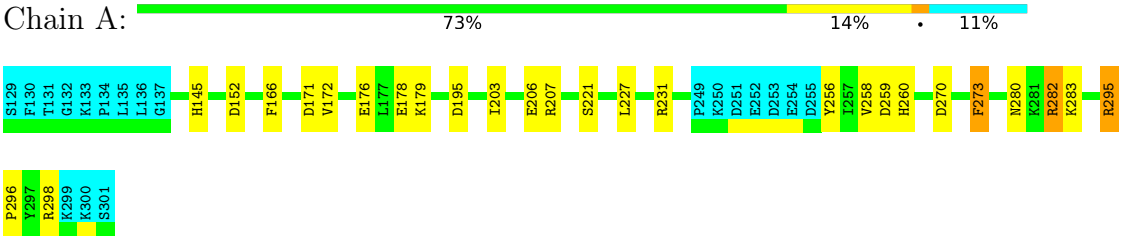
Chain A:



4.2.30 Score per residue for model 30

• Molecule 1: SCO1 protein homolog, mitochondrial

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 350 calculated structures, 30 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
DYANA	structure solution	1.5
Amber	refinement	8.0

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CU1

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	0.73±0.00	0±0/1275 (0.0± 0.0%)	1.09±0.03	4±2/1735 (0.2± 0.1%)
All	All	0.73	0/38250 (0.0%)	1.09	125/52050 (0.2%)

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	2.7±1.7
All	All	0	81

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	149	ARG	NE-CZ-NH1	9.96	125.28	120.30	22	8
1	A	231	ARG	NE-CZ-NH2	-9.84	115.38	120.30	16	7
1	A	207	ARG	NE-CZ-NH1	8.37	124.48	120.30	1	16
1	A	298	ARG	NE-CZ-NH1	8.24	124.42	120.30	26	17
1	A	282	ARG	NE-CZ-NH1	8.05	124.33	120.30	16	13
1	A	242	ARG	NE-CZ-NH1	8.02	124.31	120.30	22	1
1	A	231	ARG	NE-CZ-NH1	7.96	124.28	120.30	23	16
1	A	282	ARG	NE-CZ-NH2	-7.74	116.43	120.30	17	3
1	A	298	ARG	NE-CZ-NH2	-7.35	116.63	120.30	10	3
1	A	216	TYR	CB-CG-CD2	-7.09	116.74	121.00	23	1
1	A	239	ARG	NE-CZ-NH1	6.65	123.63	120.30	3	9
1	A	295	ARG	NE-CZ-NH2	-6.61	116.99	120.30	7	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	297	TYR	CB-CG-CD2	-6.42	117.15	121.00	13	1
1	A	295	ARG	NE-CZ-NH1	6.41	123.50	120.30	8	4
1	A	207	ARG	NE-CZ-NH2	-6.31	117.14	120.30	28	3
1	A	192	THR	CA-CB-CG2	6.28	121.19	112.40	16	2
1	A	191	THR	CA-CB-CG2	6.00	120.80	112.40	3	2
1	A	149	ARG	NE-CZ-NH2	-5.99	117.31	120.30	28	1
1	A	256	TYR	CB-CG-CD2	-5.98	117.42	121.00	5	4
1	A	231	ARG	CD-NE-CZ	5.91	131.88	123.60	14	4
1	A	282	ARG	CD-NE-CZ	5.69	131.57	123.60	17	2
1	A	241	TYR	CB-CG-CD2	-5.68	117.59	121.00	11	1
1	A	242	ARG	NE-CZ-NH2	-5.46	117.57	120.30	19	2
1	A	278	GLY	C-N-CA	5.01	134.24	121.70	19	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	CYS	Peptide	7
1	A	216	TYR	Sidechain	5
1	A	207	ARG	Sidechain	4
1	A	269	PRO	Peptide	4
1	A	149	ARG	Sidechain	4
1	A	231	ARG	Sidechain	3
1	A	244	TYR	Sidechain	3
1	A	280	ASN	Peptide	3
1	A	163	TYR	Sidechain	3
1	A	205	PRO	Peptide	3
1	A	157	GLY	Peptide	3
1	A	259	ASP	Peptide	3
1	A	170	PRO	Peptide	2
1	A	192	THR	Peptide	2
1	A	241	TYR	Sidechain	2
1	A	278	GLY	Peptide	2
1	A	140	PHE	Peptide	2
1	A	190	ILE	Peptide	2
1	A	191	THR	Peptide	2
1	A	242	ARG	Sidechain,Peptide	2
1	A	261	THR	Peptide	2
1	A	193	LEU	Peptide	2
1	A	194	PRO	Peptide	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	243	VAL	Peptide	1
1	A	245	TYR	Sidechain	1
1	A	167	THR	Peptide	1
1	A	296	PRO	Peptide	1
1	A	195	ASP	Peptide	1
1	A	298	ARG	Peptide	1
1	A	229	GLY	Peptide	1
1	A	282	ARG	Sidechain	1
1	A	172	VAL	Peptide	1
1	A	182	GLN	Peptide	1
1	A	271	GLY	Peptide	1
1	A	279	GLN	Peptide	1
1	A	202	SER	Peptide	1
1	A	208	ASP	Peptide	1
1	A	295	ARG	Sidechain	1

6.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1242	1214	1213	0±1
All	All	37290	36420	36390	13

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:184:VAL:HG23	1:A:196:LEU:H	0.50	1.66	21	1
1:A:169:CYS:HB3	1:A:171:ASP:H	0.46	1.70	26	1
1:A:182:GLN:HB3	1:A:283:LYS:HE3	0.46	1.88	29	1
1:A:277:PHE:CE1	1:A:281:LYS:HE2	0.45	2.46	21	1
1:A:182:GLN:CB	1:A:283:LYS:HE3	0.44	2.43	5	1
1:A:169:CYS:SG	1:A:170:PRO:HD2	0.44	2.53	21	1
1:A:159:TRP:CD1	1:A:269:PRO:HD3	0.43	2.48	6	1
1:A:150:LYS:HE2	1:A:155:TYR:CE2	0.43	2.48	14	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:166:PHE:CE2	1:A:258:VAL:HG11	0.43	2.48	30	1
1:A:159:TRP:CD1	1:A:193:LEU:HD23	0.43	2.48	15	1
1:A:190:ILE:O	1:A:191:THR:HG23	0.42	2.14	16	1
1:A:182:GLN:C	1:A:283:LYS:HE3	0.42	2.35	12	1
1:A:165:GLY:HA3	1:A:173:CYS:SG	0.40	2.56	9	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	154/173 (89%)	129±4 (83±2%)	18±4 (11±2%)	8±3 (5±2%)	4	24
All	All	4620/5190 (89%)	3857 (83%)	529 (11%)	234 (5%)	4	24

All 53 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	191	THR	18
1	A	270	ASP	17
1	A	296	PRO	16
1	A	221	SER	14
1	A	262	ILE	12
1	A	206	GLU	12
1	A	192	THR	11
1	A	168	HIS	9
1	A	169	CYS	9
1	A	170	PRO	8
1	A	280	ASN	8
1	A	207	ARG	7
1	A	279	GLN	6
1	A	195	ASP	6
1	A	194	PRO	5
1	A	204	ASP	5
1	A	171	ASP	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	203	ILE	4
1	A	167	THR	4
1	A	263	ILE	4
1	A	283	LYS	4
1	A	298	ARG	3
1	A	172	VAL	3
1	A	244	TYR	3
1	A	282	ARG	3
1	A	273	PHE	3
1	A	242	ARG	2
1	A	223	LYS	2
1	A	147	GLY	2
1	A	281	LYS	2
1	A	158	GLN	2
1	A	256	TYR	2
1	A	261	THR	2
1	A	272	GLU	2
1	A	243	VAL	1
1	A	155	TYR	1
1	A	220	PHE	1
1	A	208	ASP	1
1	A	258	VAL	1
1	A	271	GLY	1
1	A	209	THR	1
1	A	159	TRP	1
1	A	257	ILE	1
1	A	196	LEU	1
1	A	277	PHE	1
1	A	278	GLY	1
1	A	240	ALA	1
1	A	157	GLY	1
1	A	146	THR	1
1	A	174	PRO	1
1	A	193	LEU	1
1	A	228	THR	1
1	A	152	ASP	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	136/153 (89%)	121±3 (89±2%)	15±3 (11±2%)	9 53
All	All	4080/4590 (89%)	3624 (89%)	456 (11%)	9 53

All 103 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	207	ARG	21
1	A	148	GLU	18
1	A	178	GLU	15
1	A	283	LYS	15
1	A	150	LYS	12
1	A	231	ARG	11
1	A	260	HIS	11
1	A	262	ILE	11
1	A	153	LYS	11
1	A	256	TYR	10
1	A	149	ARG	10
1	A	185	ASP	10
1	A	218	LYS	10
1	A	274	LEU	9
1	A	210	LYS	9
1	A	298	ARG	8
1	A	145	HIS	8
1	A	156	LEU	8
1	A	177	LEU	8
1	A	169	CYS	7
1	A	282	ARG	7
1	A	259	ASP	7
1	A	166	PHE	7
1	A	211	GLU	6
1	A	277	PHE	6
1	A	179	LYS	6
1	A	281	LYS	6
1	A	242	ARG	6
1	A	280	ASN	6
1	A	219	GLU	6
1	A	294	MET	5
1	A	270	ASP	5
1	A	266	LEU	5
1	A	257	ILE	5
1	A	191	THR	5
1	A	173	CYS	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	239	ARG	5
1	A	224	LEU	4
1	A	193	LEU	4
1	A	295	ARG	4
1	A	152	ASP	4
1	A	202	SER	4
1	A	208	ASP	4
1	A	160	LEU	4
1	A	158	GLN	4
1	A	192	THR	4
1	A	264	MET	4
1	A	176	GLU	4
1	A	187	ILE	3
1	A	297	TYR	3
1	A	245	TYR	3
1	A	233	GLU	3
1	A	196	LEU	3
1	A	155	TYR	3
1	A	272	GLU	3
1	A	146	THR	3
1	A	144	THR	2
1	A	188	ASP	2
1	A	244	TYR	2
1	A	195	ASP	2
1	A	175	GLU	2
1	A	223	LYS	2
1	A	230	THR	2
1	A	261	THR	2
1	A	154	ASP	2
1	A	209	THR	2
1	A	258	VAL	2
1	A	168	HIS	2
1	A	151	THR	2
1	A	171	ASP	2
1	A	172	VAL	2
1	A	184	VAL	2
1	A	204	ASP	2
1	A	234	VAL	2
1	A	289	SER	2
1	A	273	PHE	2
1	A	236	GLN	2
1	A	143	THR	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	183	VAL	1
1	A	246	SER	1
1	A	279	GLN	1
1	A	167	THR	1
1	A	232	GLU	1
1	A	237	VAL	1
1	A	142	LEU	1
1	A	180	MET	1
1	A	221	SER	1
1	A	276	TYR	1
1	A	141	SER	1
1	A	181	ILE	1
1	A	215	ASN	1
1	A	200	PHE	1
1	A	216	TYR	1
1	A	285	GLU	1
1	A	161	LEU	1
1	A	243	VAL	1
1	A	228	THR	1
1	A	220	PHE	1
1	A	213	ILE	1
1	A	182	GLN	1
1	A	189	SER	1
1	A	235	ASP	1
1	A	227	LEU	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 monosaccharides in this entry.

6.6 Ligand geometry

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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