WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report ⓘ

Apr 26, 2016 – 04:39 PM BST

PDB ID	:	1QND
Title	:	STEROL CARRIER PROTEIN-2, NMR, 20 STRUCTURES
Authors	:	Lopez-Garcia, F.; Szyperski, T.; Dyer, J.H.; Choinowski, T.; Seedorf, U.; Hauser, H.; Wuthrich, K.
Deposited on	:	1999-10-14

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 ⓘ symbol.

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

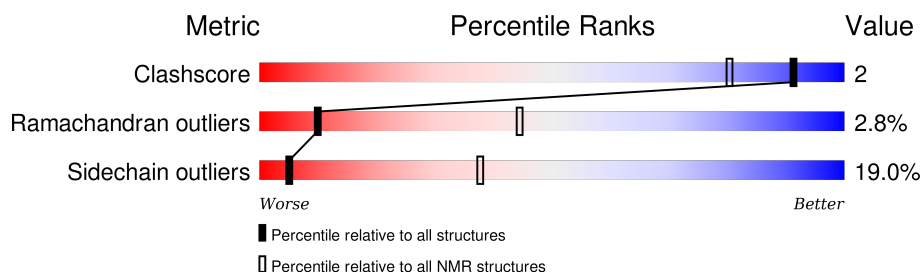|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
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 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	Percentile Ranks	Value
Clashscore		2
Ramachandran outliers		2.8%
Sidechain outliers		19.0%

Worse *Better*

▌ Percentile relative to all structures

❒ Percentile relative to all NMR structures

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 >=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	123	63%　　6% ·　　30%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 15 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:5-A:84, A:100-A:101, A:113-A:116 (86)	0.37	15

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 and 2 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1882 atoms, of which 955 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called NONSPECIFIC LIPID-TRANSFER PROTEIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	123	1882	590	955	155	176	6	0

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.

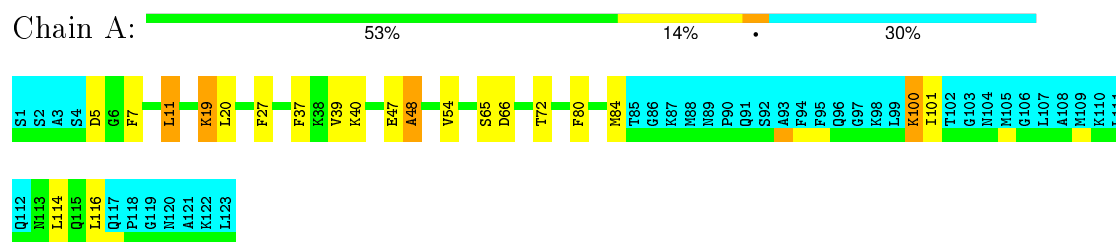• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN
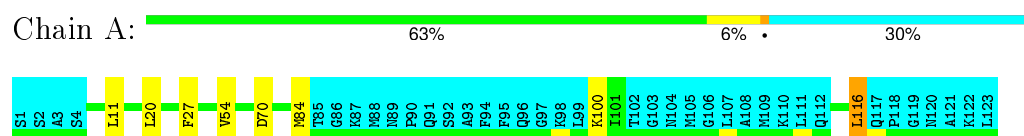
4.2 Scores per residue for each member of the ensemble

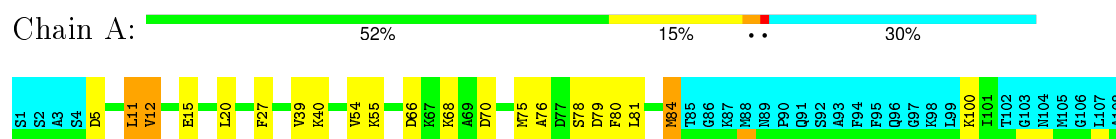Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

4.2.2 Score per residue for model 2

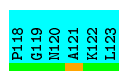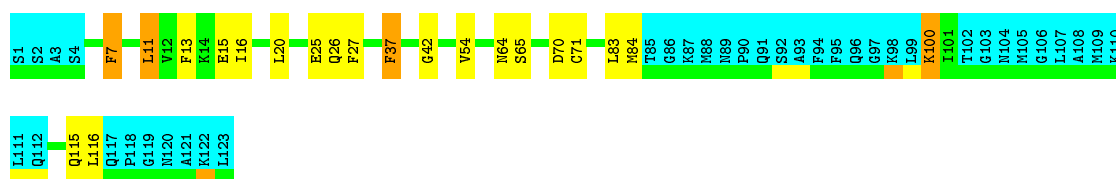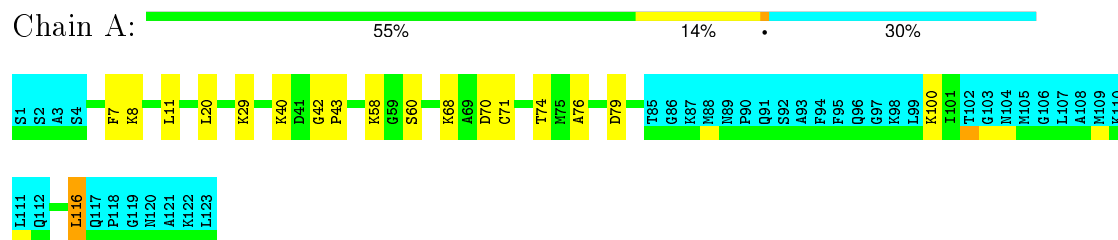• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

M109 K110 L111 Q112 L116 Q117 P118 G119 N120 A121 K122 L123

4.2.3 Score per residue for model 3

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A: | 53% | 15% | • | 30%

S1 S2 A3 S4 F7 E15 I16 E17 K18 K19 L20 E25 Q26 F27 K38 V39 A48 V54 K58 D70 M75 A76 F80 L83 M84 T85 G86 K87 M88 N89 P90 Q91 S92 A93 F94 F95 Q96 G97 K98 L99 K100 I101 T102 G103 N104 M105 G106 L107 A108

M109 K110 L111 Q112 L116 Q117 P118 G119 N120 A121 K122 L123

4.2.4 Score per residue for model 4

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A: | 51% | 19% | 30%

S1 S2 A3 S4 D5 L11 L20 F27 V28 K29 K40 P43 V54 G59 S60 V61 L62 P63 N64 D70 I73 A76 D77 S78 D79 F80 L81 A82 L83 M84 T85 G86 K87 M88 N89 P90 Q91 S92 A93 F94 F95 Q96 G97 K98 L99 K100 I101 T102 G103 N104

M105 G106 L107 A108 M109 K110 L111 Q112 N113 L114 Q115 L116 Q117 P118 G119 N120 A121 K122 L123

4.2.5 Score per residue for model 5

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A: | 56% | 11% | •• | 30%

S1 S2 A3 S4 L11 K19 L20 F27 V39 V54 K58 K68 A69 D70 M75 D79 F80 M84 T85 G86 K87 M88 N89 P90 Q91 S92 A93 F94 F95 Q96 G97 K98 L99 K100 I101 T102 G103 N104 M105 G106 L107 A108 M109 K110 L111 Q112 N113 L116 Q117

P118 G119 N120 A121 K122 L123

4.2.6 Score per residue for model 6

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A: | 53% | 14% | • | 30%

S1 S2 A3 S4 F7 L11 V12 F13 K14 E15 I16 L20 E25 Q26 F27 F37 G42 V54 N64 S65 D70 C71 L83 M84 T85 G86 K87 M88 N89 P90 Q91 S92 A93 F94 F95 Q96 G97 K98 L99 K100 I101 T102 G103 N104 M105 G106 L107 A108 M109 K110

L111 Q112 Q115 L116 Q117 P118 G119 N120 A121 K122 L123

4.2.7 Score per residue for model 7

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A: 55% 14% • 30%

S1 S2 A3 S4 F7 K8 L11 L20 K29 K40 D41 G42 P43 K58 G59 S60 K68 A69 D70 C71 T74 M75 A76 D79 T85 G86 K87 M88 N89 P90 Q91 S92 A93 F94 F95 Q96 G97 K98 L99 K100 I101 T102 G103 N104 M105 G106 L107 A108 M109 K110

L111 Q112 L116 Q117 P118 G119 N120 A121 K122 L123

4.2.8 Score per residue for model 8

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A: 54% 13% • 30%

S1 S2 A3 S4 L11 K19 L20 F27 K30 G33 D53 V54 K55 K58 N64 D70 M75 L81 M84 T85 G86 K87 M88 N89 P90 Q91 S92 A93 F94 F95 Q96 G97 K98 L99 K100 I101 T102 G103 N104 M105 G106 L107 A108 M109 K110 L111 Q112 N113

L114 Q115 L116 Q117 P118 G119 N120 A121 K122 L123

4.2.9 Score per residue for model 9

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A: 54% 15% • 30%

S1 S2 A3 S4 D5 G6 F7 L11 K14 L20 E25 Q26 F27 K38 A48 V54 K55 K58 D70 I73 L83 M84 T85 G86 K87 M88 N89 P90 Q91 S92 A93 F94 F95 Q96 G97 K98 L99 K100 I101 T102 G103 N104 M105 G106 L107 A108 M109 K110 L111

Q112 L116 Q117 P118 G119 N120 A121 K122 L123

4.2.10 Score per residue for model 10

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A:

4.2.11 Score per residue for model 11

● Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A:

4.2.12 Score per residue for model 12
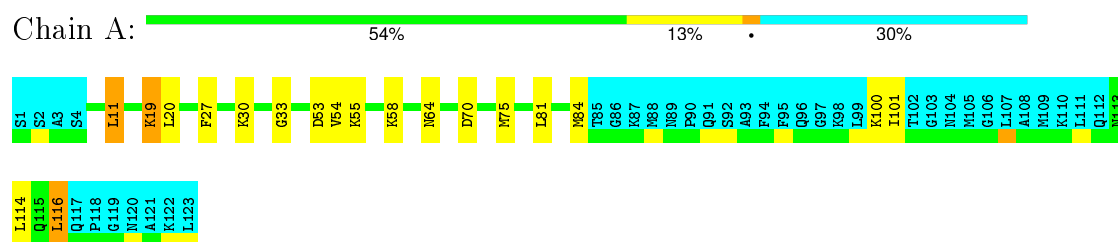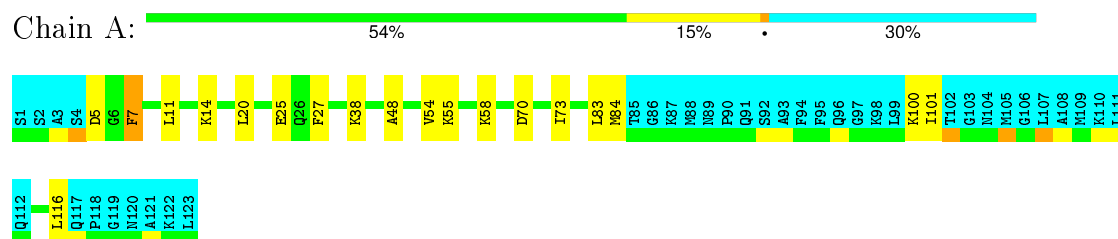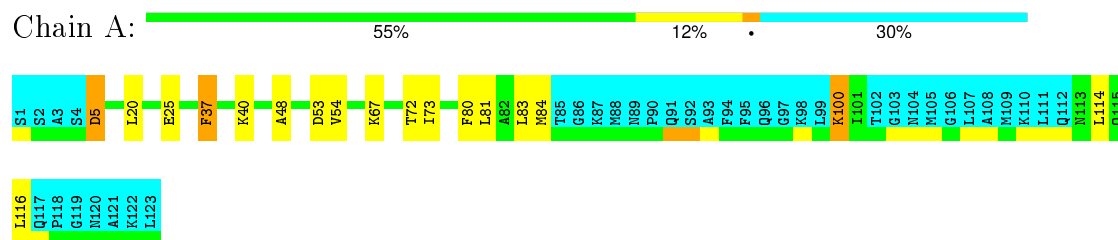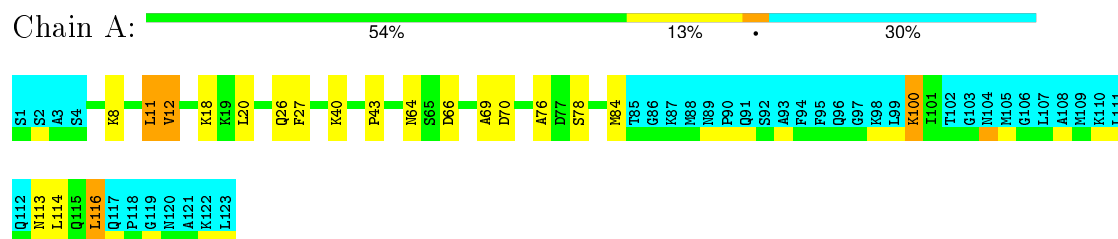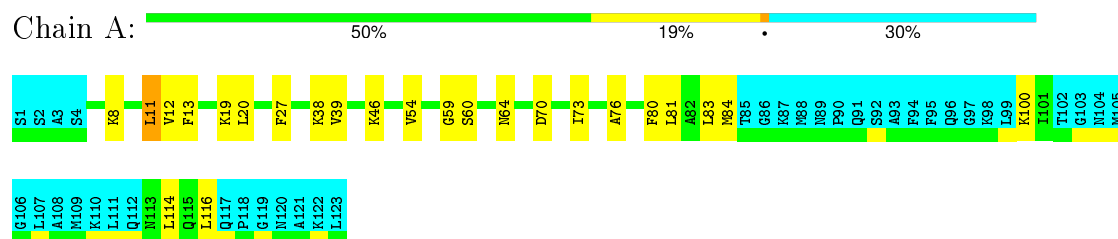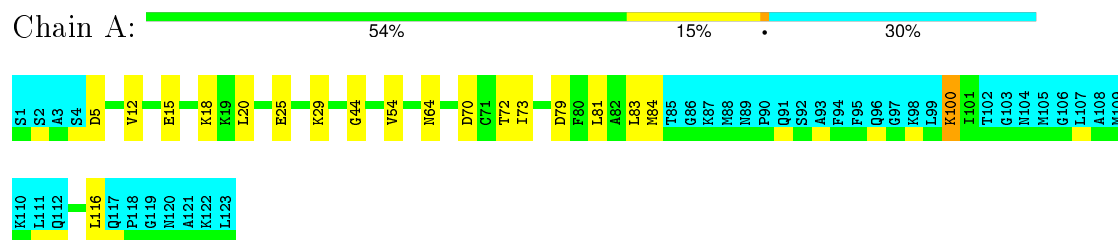
● Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A:

4.2.13 Score per residue for model 13

● Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A:

4.2.14 Score per residue for model 14

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

4.2.15 Score per residue for model 15 (medoid)

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

4.2.16 Score per residue for model 16

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

4.2.17 Score per residue for model 17

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

4.2.18 Score per residue for model 18

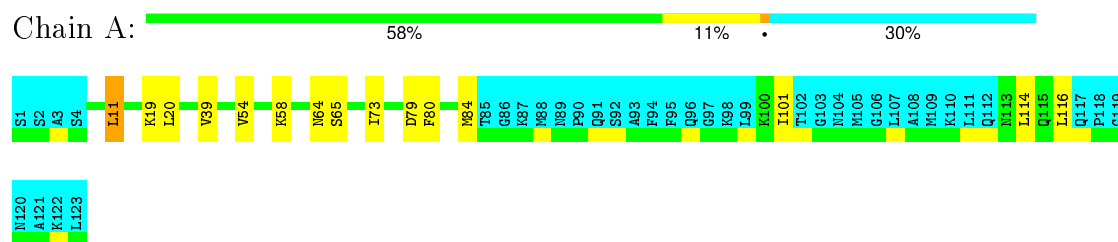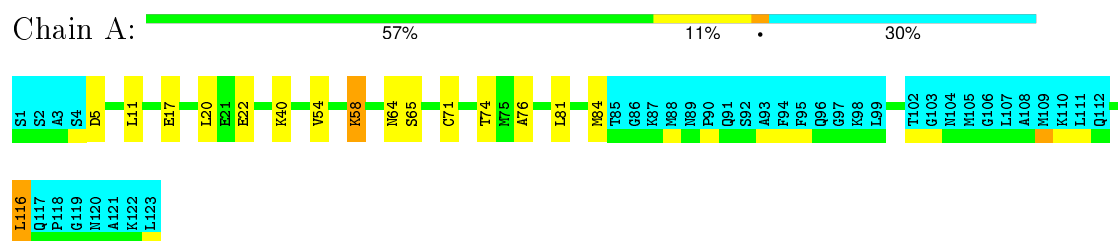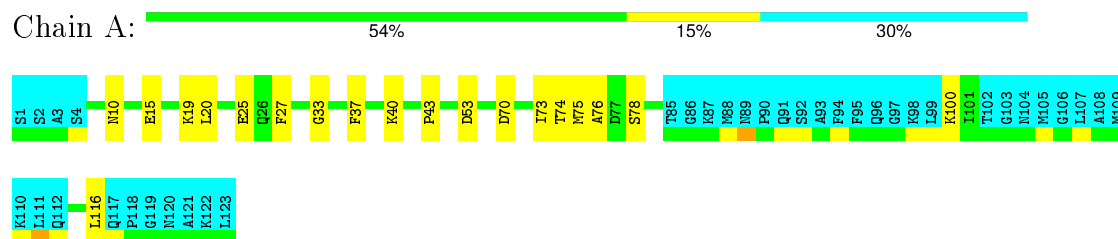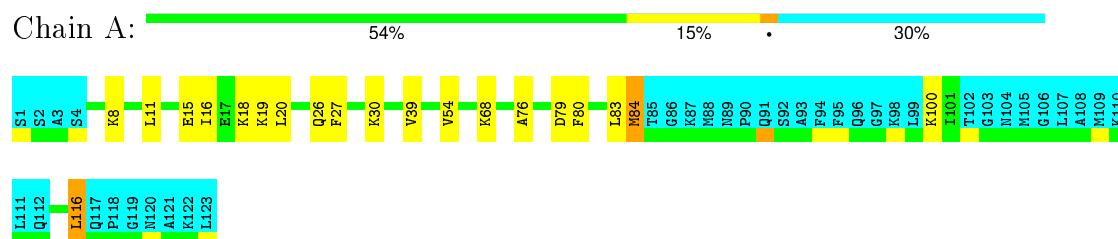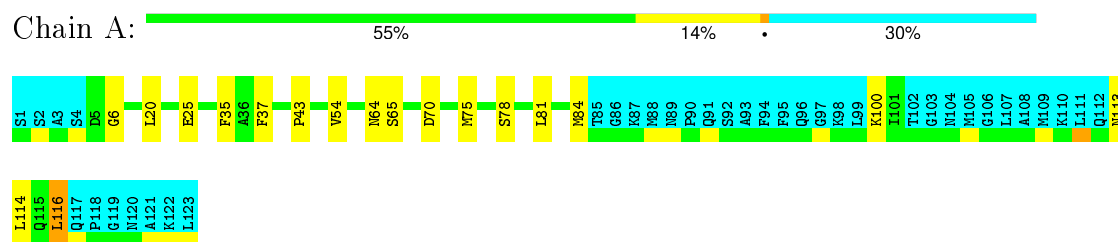• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A:

4.2.19 Score per residue for model 19

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A:
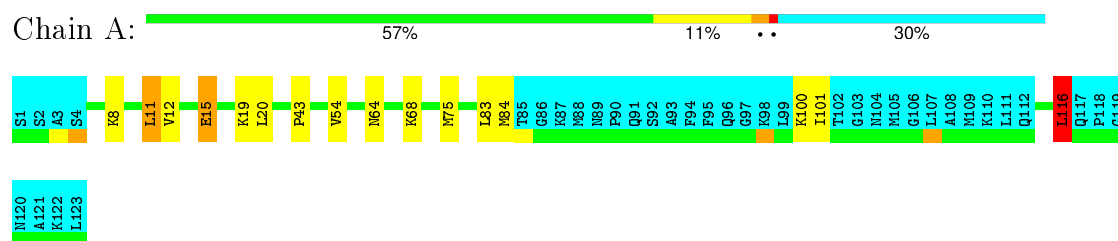
4.2.20 Score per residue for model 20

• Molecule 1: NONSPECIFIC LIPID-TRANSFER PROTEIN

Chain A:
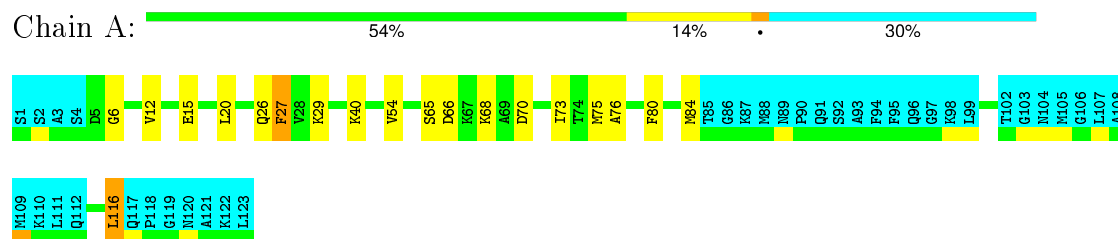
5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *TORSION ANGLE DYNAMICS*.

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
.41 A	refinement	
DYANA	structure solution	

No chemical shift data was provided. 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)

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	0.58±0.01	0±0/671 (0.0±0.0%)	1.07±0.03	1±1/897 (0.1±0.1%)
All	All	0.58	0/13420 (0.0%)	1.07	11/17940 (0.1%)

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	0.7±0.7
All	All	0	13

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	5	ASP	CB-CG-OD1	6.59	124.23	118.30	10	1
1	A	27	PHE	CB-CG-CD2	-6.24	116.43	120.80	5	2
1	A	7	PHE	CB-CG-CD2	-6.20	116.46	120.80	6	1
1	A	59	GLY	C-N-CA	5.54	135.55	121.70	4	2
1	A	7	PHE	CB-CG-CD1	5.45	124.62	120.80	6	1
1	A	37	PHE	CB-CG-CD2	-5.26	117.12	120.80	10	2
1	A	84	MET	CG-SD-CE	5.20	108.52	100.20	2	1
1	A	101	ILE	CB-CA-C	5.10	121.80	111.60	5	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	33	GLY	Peptide	2
1	A	48	ALA	Peptide	2
1	A	37	PHE	Sidechain	2
1	A	74	THR	Peptide	1
1	A	7	PHE	Sidechain	1
1	A	27	PHE	Sidechain	1
1	A	40	LYS	Peptide	1
1	A	69	ALA	Peptide	1
1	A	75	MET	Peptide	1
1	A	80	PHE	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	661	681	681	3±2
All	All	13220	13620	13620	54

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 2.

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:7:PHE:CE1	1:A:48:ALA:HB2	0.69	2.23	9	2
1:A:39:VAL:HG22	1:A:80:PHE:CE2	0.65	2.27	5	5
1:A:15:GLU:HG3	1:A:116:LEU:HD21	0.64	1.69	3	4
1:A:72:THR:HG22	1:A:73:ILE:H	0.61	1.54	13	2
1:A:12:VAL:HG13	1:A:84:MET:SD	0.60	2.36	12	4
1:A:39:VAL:HG22	1:A:80:PHE:CE1	0.60	2.32	2	1
1:A:58:LYS:H	1:A:58:LYS:HE3	0.59	1.57	15	1
1:A:16:ILE:HD11	1:A:84:MET:CE	0.56	2.31	17	2
1:A:11:LEU:HD22	1:A:11:LEU:C	0.55	2.21	1	5
1:A:84:MET:SD	1:A:116:LEU:HD21	0.54	2.42	2	2
1:A:13:PHE:HA	1:A:16:ILE:HD12	0.54	1.80	6	1
1:A:11:LEU:C	1:A:11:LEU:HD22	0.54	2.22	6	3
1:A:7:PHE:CD1	1:A:48:ALA:HB2	0.52	2.39	9	1
1:A:15:GLU:CG	1:A:116:LEU:HD21	0.50	2.36	3	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
1:A:16:ILE:HD11	1:A:84:MET:HE1	0.49	1.85	17	1
1:A:12:VAL:HG22	1:A:84:MET:SD	0.47	2.49	11	1
1:A:13:PHE:CZ	1:A:39:VAL:HG21	0.47	2.44	12	1
1:A:39:VAL:HG22	1:A:80:PHE:CD2	0.47	2.45	14	2
1:A:11:LEU:HD13	1:A:12:VAL:N	0.46	2.26	2	1
1:A:72:THR:HG22	1:A:73:ILE:N	0.45	2.26	13	1
1:A:35:PHE:CE1	1:A:37:PHE:CE1	0.44	3.05	18	1
1:A:116:LEU:N	1:A:116:LEU:HD23	0.44	2.27	20	1
1:A:7:PHE:CE2	1:A:42:GLY:HA3	0.43	2.49	7	2
1:A:15:GLU:CB	1:A:116:LEU:HD22	0.42	2.44	20	1
1:A:25:GLU:CB	1:A:58:LYS:HE3	0.42	2.45	9	1
1:A:7:PHE:CD2	1:A:48:ALA:HB2	0.42	2.50	1	1
1:A:7:PHE:CE2	1:A:48:ALA:HB2	0.42	2.49	1	1
1:A:19:LYS:HG2	1:A:27:PHE:CZ	0.41	2.51	8	2
1:A:72:THR:H	1:A:101:ILE:CG2	0.40	2.29	1	1
1:A:16:ILE:HD11	1:A:84:MET:HE2	0.40	1.94	3	1
1:A:15:GLU:HG2	1:A:116:LEU:HD22	0.40	1.94	19	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	86/123 (70%)	70±3 (82±3%)	13±3 (16±3%)	2±1 (3±1%)	10	44
All	All	1720/2460 (70%)	1404 (82%)	267 (16%)	49 (3%)	10	44

All 13 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	116	LEU	17
1	A	76	ALA	10
1	A	43	PRO	6
1	A	83	LEU	4
1	A	60	SER	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	6	GLY	2
1	A	54	VAL	2
1	A	114	LEU	1
1	A	46	LYS	1
1	A	113	ASN	1
1	A	101	ILE	1
1	A	68	LYS	1
1	A	44	GLY	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	71/99 (72%)	58±2 (81±3%)	14±2 (19±3%)	5	38
All	All	1420/1980 (72%)	1150 (81%)	270 (19%)	5	38

All 49 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	20	LEU	20
1	A	100	LYS	15
1	A	54	VAL	15
1	A	70	ASP	14
1	A	11	LEU	14
1	A	84	MET	12
1	A	27	PHE	11
1	A	64	ASN	10
1	A	116	LEU	9
1	A	81	LEU	8
1	A	19	LYS	8
1	A	40	LYS	8
1	A	79	ASP	7
1	A	114	LEU	7
1	A	75	MET	7
1	A	5	ASP	7
1	A	73	ILE	6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	65	SER	6
1	A	25	GLU	6
1	A	58	LYS	6
1	A	8	LYS	5
1	A	68	LYS	5
1	A	26	GLN	5
1	A	78	SER	5
1	A	83	LEU	5
1	A	66	ASP	4
1	A	29	LYS	4
1	A	18	LYS	4
1	A	53	ASP	3
1	A	38	LYS	3
1	A	55	LYS	3
1	A	15	GLU	3
1	A	71	CYS	3
1	A	12	VAL	3
1	A	113	ASN	2
1	A	37	PHE	2
1	A	74	THR	2
1	A	80	PHE	2
1	A	17	GLU	1
1	A	30	LYS	1
1	A	67	LYS	1
1	A	47	GLU	1
1	A	101	ILE	1
1	A	62	LEU	1
1	A	60	SER	1
1	A	22	GLU	1
1	A	14	LYS	1
1	A	115	GLN	1
1	A	10	ASN	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 ⓘ

There are no carbohydrates 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