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 (i)

May 29, 2020 – 08:02 am BST

PDB ID	:	5OEO
Title	:	Solution structure of the complex of TRPV5(655-725) with a Calmodulin E32Q/E68Q double mutant
Authors	:	Vuister, G.W.; Bokhovchuk, F.M.; Bate, N.; Kovalevskaya, N.; Goult, B.T.; Spronk, C.A.E.M.
Deposited on	:	2017-07-09

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

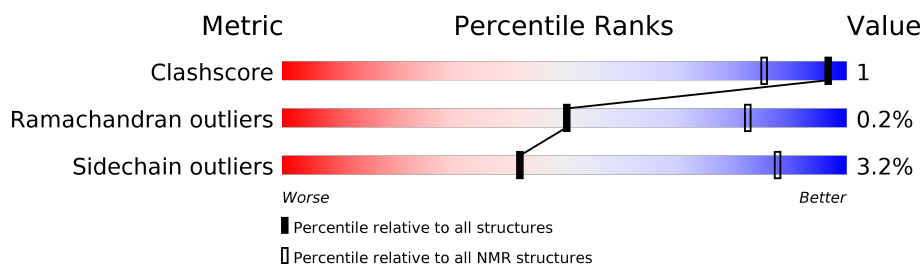|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 75%.

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		1
Ramachandran outliers		0.2%
Sidechain outliers		3.2%

Worse   *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all NMR structures

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

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	150	84% • 13%
2	C	73	16% • 82%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 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:7-A:73 (67)	0.72	7
2	A:84-A:146, C:698-C:710 (76)	0.57	1

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 2 single-model clusters were found.

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

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3347 atoms, of which 1623 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Calmodulin-1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	150	2294	722	1114	192	256	10	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	expression tag	UNP P0DP23
A	31	GLN	GLU	engineered mutation	UNP P0DP23
A	67	GLN	GLU	engineered mutation	UNP P0DP23

- Molecule 2 is a protein called Transient receptor potential cation channel subfamily V member 5.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	C	73	1051	320	509	101	121	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	653	GLY	-	expression tag	UNP Q9NQA5
C	654	ALA	-	expression tag	UNP Q9NQA5

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

| Mol	Chain	Residues	Atoms	
			Total	Ca
3	A	2	2	2

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: Calmodulin-1
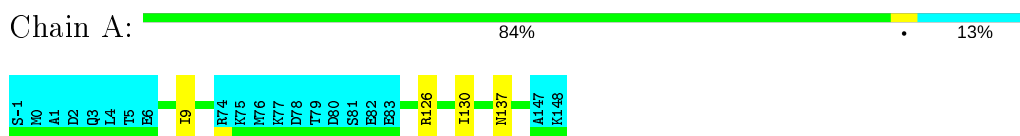
- Molecule 2: Transient receptor potential cation channel subfamily V member 5
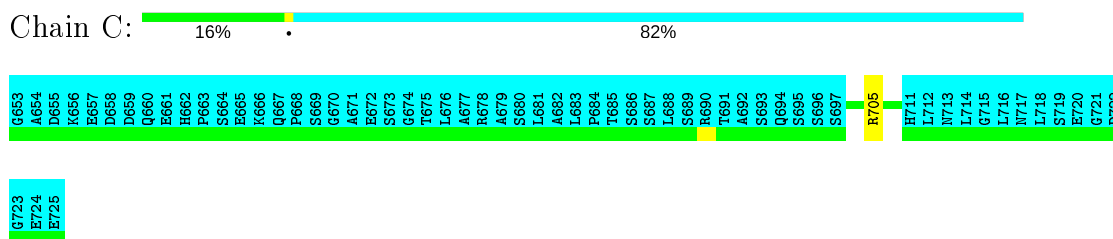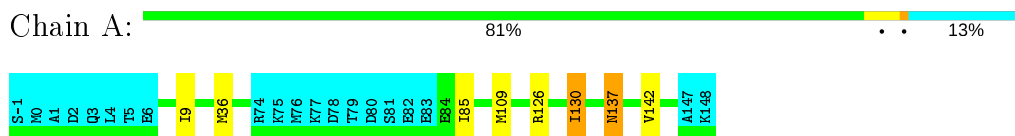

4.2 Scores per residue for each member of the ensemble

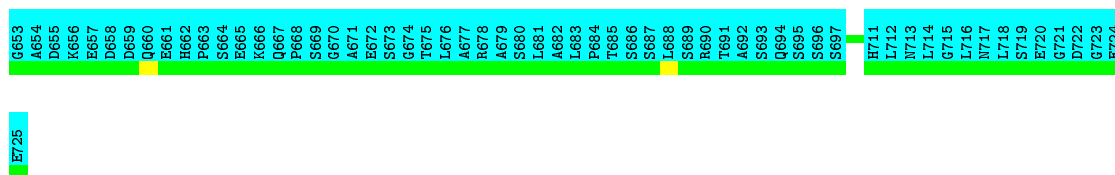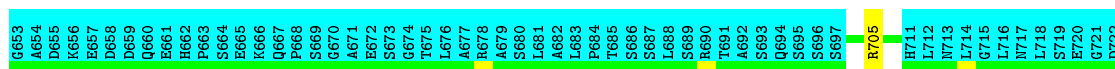Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1 (medoid)

- Molecule 1: Calmodulin-1

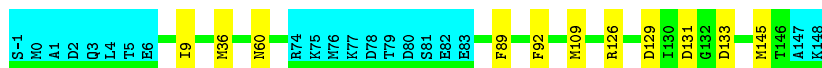- Molecule 2: Transient receptor potential cation channel subfamily V member 5

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722 G723 E724

E725

4.2.2 Score per residue for model 2

- Molecule 1: Calmodulin-1

Chain A: 85%   •   13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 M71 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R106 A147 K148

- Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 16%   •   82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722

G723 E724 E725

4.2.3 Score per residue for model 3

- Molecule 1: Calmodulin-1

Chain A: 79%   7%   13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 M36 N60 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 F89 F92 M109 R126 D129 I130 D131 G132 D133 M145 T146 A147 K148

- Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 15%   •   82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 L704 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721

D722 G723 E724 E725
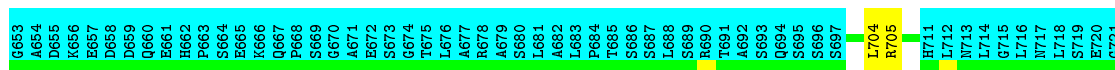
4.2.4 Score per residue for model 4

- Molecule 1: Calmodulin-1
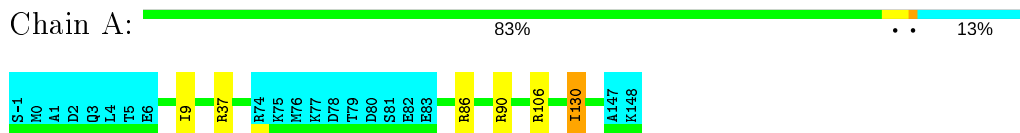
Chain A: 83% 13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 R37 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R86 R90 R106 I130 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 16% 82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 H698 R699 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722

G723 E724 E725
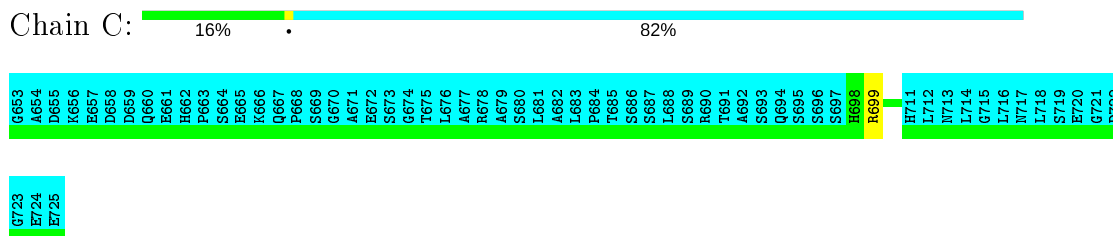
4.2.5 Score per residue for model 5

• Molecule 1: Calmodulin-1
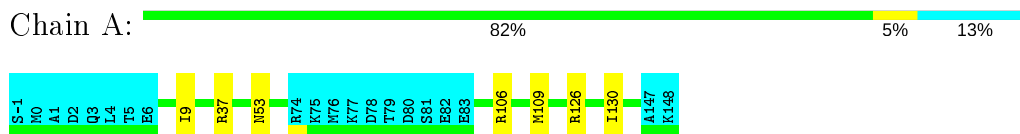
Chain A: 82% 5% 13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 R37 N53 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R106 M109 R126 I130 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 15% 82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 H698 R699 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720

G721 D722 G723 E724 E725
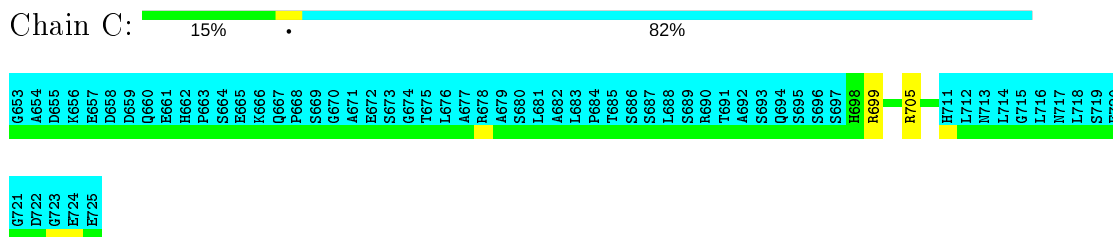
4.2.6 Score per residue for model 6

• Molecule 1: Calmodulin-1
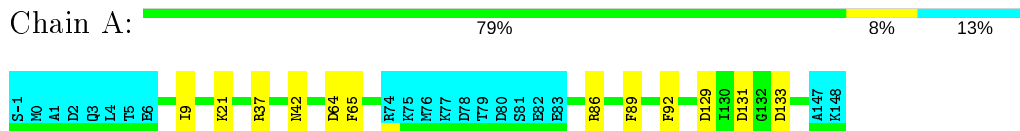
Chain A: 79% 8% 13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 K21 R37 N42 D64 F65 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R86 F89 F92 D129 I130 D131 G132 D133 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 16% 82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722

G723 E724 E725
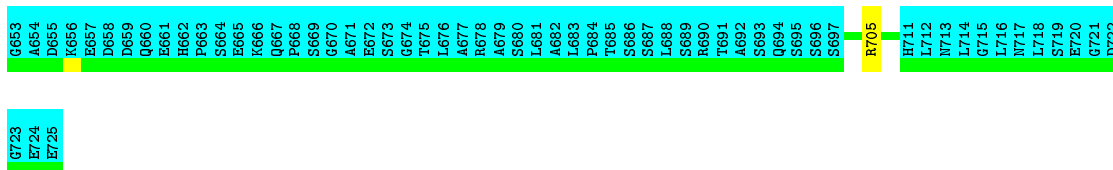
4.2.7 Score per residue for model 7

• Molecule 1: Calmodulin-1

Chain A: 81% 5% · 13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 F12 R37 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R86 R90 R106 R126 D131 N137 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 16% · 82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722

G723 E724 E725

4.2.8 Score per residue for model 8

• Molecule 1: Calmodulin-1

Chain A: 80% 7% 13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 Q41 N60 M71 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R86 F89 F92 M109 Q143 T146 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 14% ·· 82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 H698 R699 L704 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719
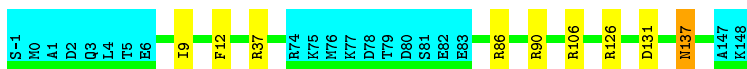
E720 G721 D722 G723 E724 E725

4.2.9 Score per residue for model 9

• Molecule 1: Calmodulin-1

Chain A:

• Molecule 2: Transient receptor potential cation channel subfamily V member 5
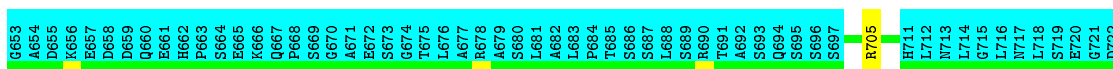
Chain C:

4.2.10 Score per residue for model 10

• Molecule 1: Calmodulin-1

Chain A:

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C:

4.2.11 Score per residue for model 11

• Molecule 1: Calmodulin-1

Chain A:

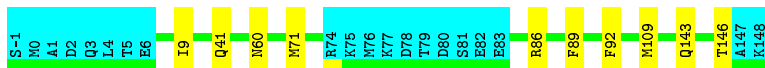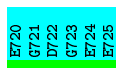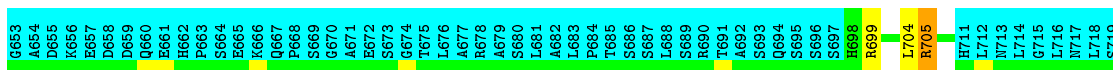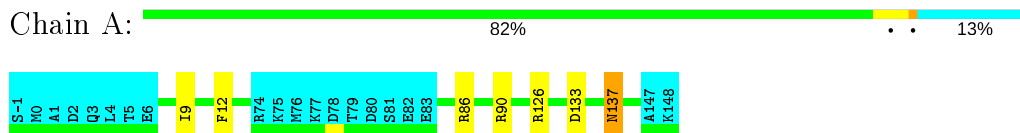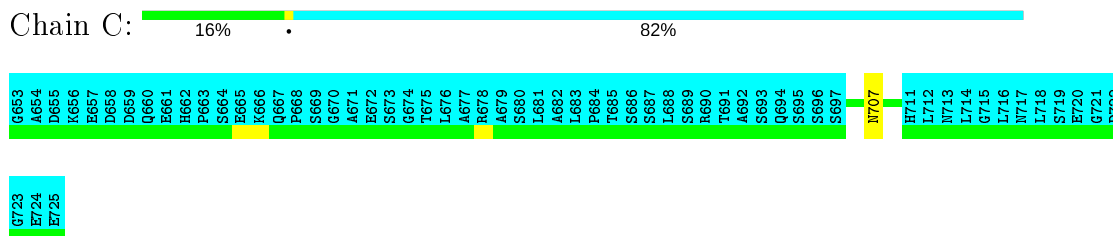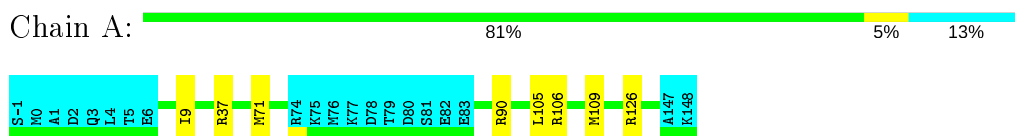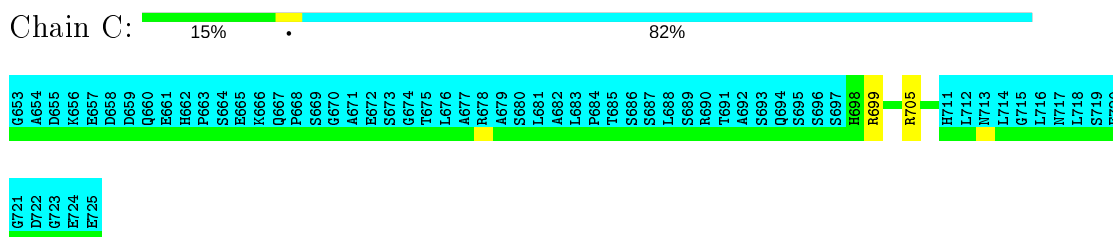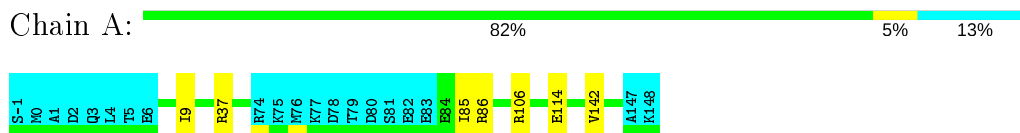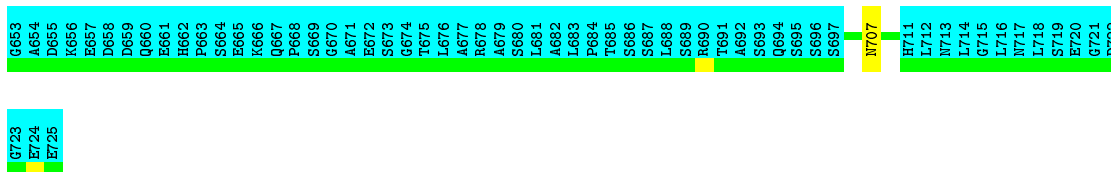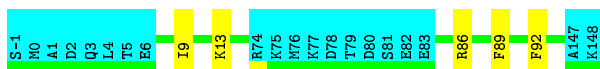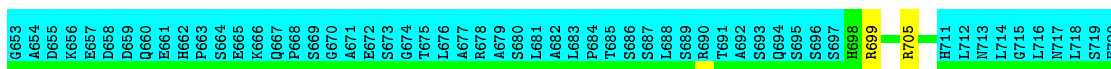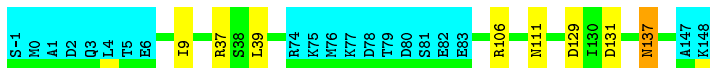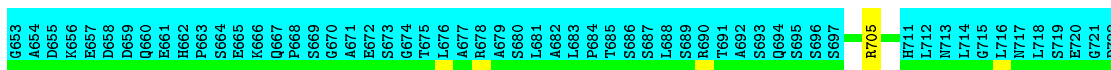• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C:

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 N707 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722

G723 E724 E725

4.2.12　Score per residue for model 12

• Molecule 1: Calmodulin-1

Chain A:　　　　　　　　　　　　　83%　　　　　　　　　　　•　　13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 K13 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R86 F89 F92 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C:　　15%　　•　　　　　　　82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 H698 R699 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720

G721 D722 G723 E724 E725

4.2.13　Score per residue for model 13

• Molecule 1: Calmodulin-1

Chain A:　　　　　　　　　　　　　81%　　　　　　　　　　5%　•　　13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 R37 S38 L39 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R106 N111 D129 I130 D131 N137 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C:　　16%　　•　　　　　　　82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722

G723 E724 E725

4.2.14　Score per residue for model 14

• Molecule 1: Calmodulin-1
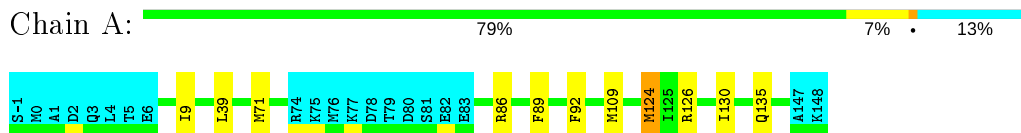
Chain A:

79%   7%   •   13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 L39 M71 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R86 F89 F92 M109 M124 I125 R126 I130 Q135 A147 K148

● Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 15%   •   82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 W701 E702 I703 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720

G721 D722 G723 E724 E725
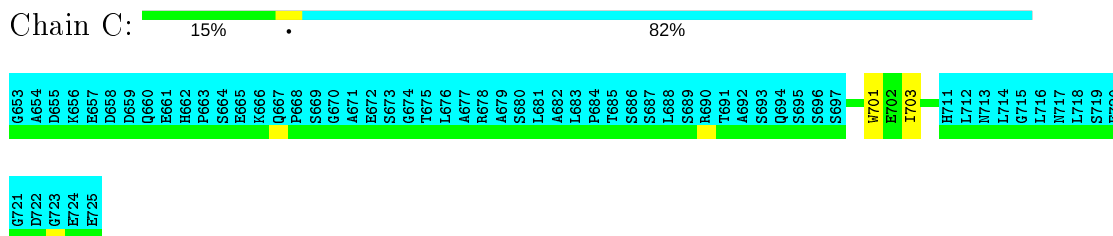
4.2.15 Score per residue for model 15

● Molecule 1: Calmodulin-1
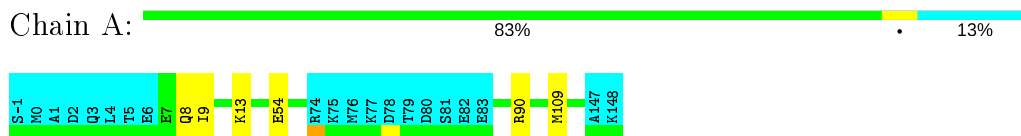
Chain A: 83%   •   13%

S-1 M0 A1 D2 Q3 L4 T5 E6 E7 Q8 I9 K13 E54 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R90 M109 A147 K148

● Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 16%   •   82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 H698 R699 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722

G723 E724 E725
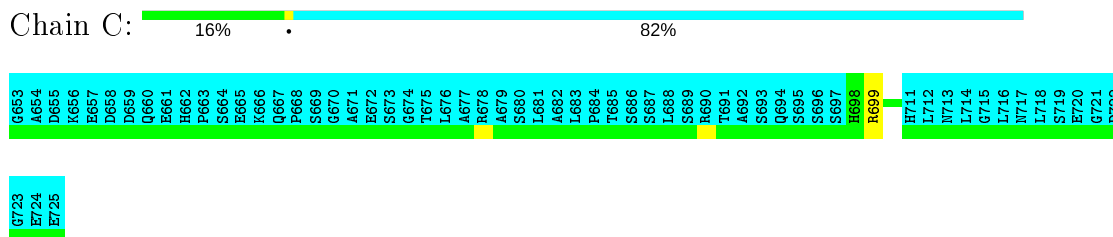
4.2.16 Score per residue for model 16

● Molecule 1: Calmodulin-1

Chain A: 77%   9%   •   13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 L32 I63 M71 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 E84 I85 R90 V91 F92 I100 L105 R126 D131 M137 V142 Q143 A147 K148

● Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 16%   •   82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 N707 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722

G723 E724 E725

4.2.17 Score per residue for model 17

• Molecule 1: Calmodulin-1
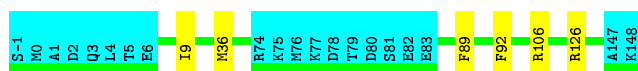
Chain A: 83% · 13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 M36 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 F89 F92 R106 R126 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 15% · 82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 H698 R699 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720

G721 D722 G723 E724 E725
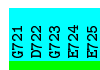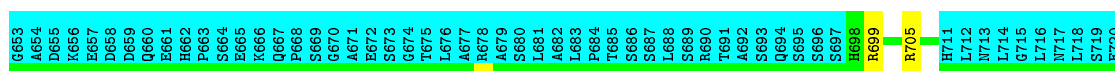
4.2.18 Score per residue for model 18

• Molecule 1: Calmodulin-1

Chain A: 77% 8% · 13%

S-1 M0 A1 D2 Q3 L4 T5 E6 I9 K21 L32 L39 N42 D50 I63 D64 F65 P66 Q67 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R106 M109 E127 I130 M145 T146 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C: 16% · 82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720 G721 D722

G723 E724 E725
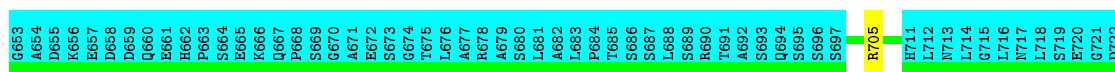
4.2.19 Score per residue for model 19

• Molecule 1: Calmodulin-1
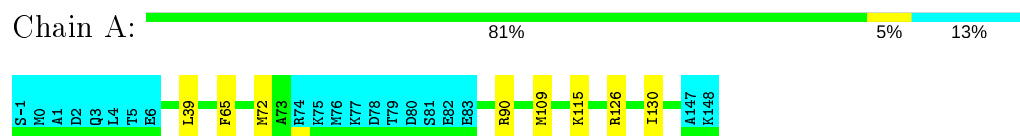
Chain A:

81%　　5%　13%

S-1 M0 A1 D2 Q3 L4 T5 E6 L39 F65 M72 A73 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 R90 M109 K115 R126 I130 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C:

15%　·　82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 H698 R699 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720

G721 D722 G723 E724 E725
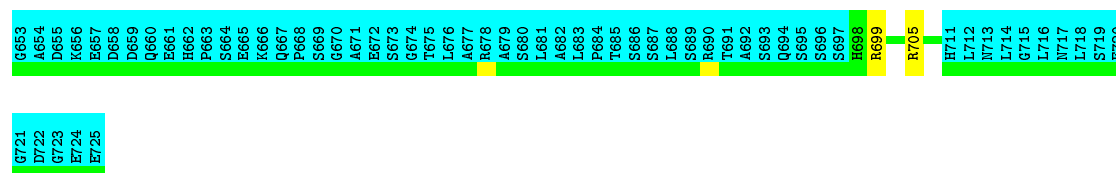
4.2.20 Score per residue for model 20

• Molecule 1: Calmodulin-1
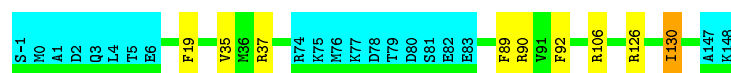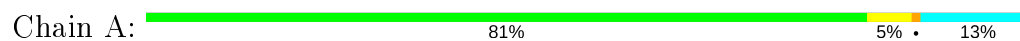
Chain A:

81%　　5% · 13%

S-1 M0 A1 D2 Q3 L4 T5 E6 F19 V35 M36 R37 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 F89 R90 V91 F92 R106 R126 I130 A147 K148

• Molecule 2: Transient receptor potential cation channel subfamily V member 5

Chain C:

15%　·　82%

G653 A654 D655 K656 E657 D658 D659 Q660 E661 H662 P663 S664 E665 K666 Q667 P668 S669 G670 A671 E672 S673 G674 T675 L676 A677 R678 A679 S680 L681 A682 L683 P684 T685 S686 S687 L688 S689 R690 T691 A692 S693 Q694 S695 S696 S697 H698 R699 R705 H711 L712 N713 L714 G715 L716 N717 L718 S719 E720

G721 D722 G723 E724 E725
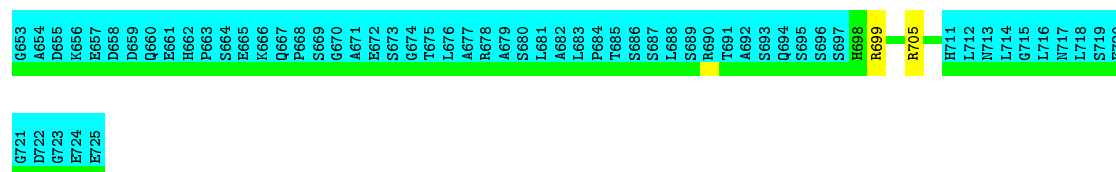
5 Refinement protocol and experimental data overview (i)

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

Of the 120 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
CYANA	structure calculation	3.97
YASARA	refinement	15.6

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	1766
Number of shifts mapped to atoms	1766
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	75%

No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1 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	1021	959	959	2±1
2	C	111	109	109	0±0
All	All	22680	21360	21360	32

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 1.

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:131:ASP:OD2	1:A:133:ASP:OD2	0.73	2.07	6	1
1:A:129:ASP:OD1	1:A:131:ASP:OD2	0.62	2.17	13	1
1:A:85:ILE:HG21	1:A:142:VAL:HG22	0.56	1.78	16	3
1:A:19:PHE:CD1	1:A:35:VAL:HG22	0.54	2.38	20	1
1:A:92:PHE:CE2	1:A:100:ILE:HG23	0.51	2.40	16	1
1:A:131:ASP:OD2	1:A:133:ASP:OD1	0.49	2.30	3	1
1:A:92:PHE:CD1	2:C:704:LEU:HD11	0.48	2.43	8	1
1:A:89:PHE:HA	1:A:92:PHE:CE2	0.47	2.45	20	4
1:A:92:PHE:CZ	1:A:100:ILE:HG23	0.46	2.46	16	1
1:A:89:PHE:HA	1:A:92:PHE:CE1	0.46	2.46	3	3
1:A:137:ASN:C	1:A:137:ASN:HD22	0.46	2.13	16	4
1:A:137:ASN:HD22	1:A:137:ASN:C	0.45	2.14	7	1
1:A:130:ILE:H	1:A:130:ILE:HD13	0.45	1.72	20	1
1:A:130:ILE:HD13	1:A:130:ILE:H	0.45	1.72	1	3
1:A:143:GLN:HA	1:A:146:THR:HG22	0.43	1.90	8	1
1:A:124:MET:SD	2:C:701:TRP:CD1	0.43	3.12	14	1
1:A:9:ILE:HD12	1:A:12:PHE:CZ	0.42	2.48	7	1
1:A:9:ILE:HA	1:A:12:PHE:CE1	0.42	2.50	9	1
1:A:32:LEU:HD22	1:A:63:ILE:HD13	0.41	1.91	16	1
1:A:32:LEU:HD12	1:A:63:ILE:HD11	0.40	1.93	18	1

5.2　Torsion angles ⓘ

5.2.1　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	130/150 (87%)	128±1 (99±1%)	1±1 (1±1%)	0±1 (0±1%)	50	82
2	C	13/73 (18%)	13±0 (100±2%)	0±0 (0±2%)	0±0 (0±0%)	100	100
All	All	2860/4460 (64%)	2825 (99%)	29 (1%)	6 (0%)	50	82

All 4 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	42	ASN	2
1	A	65	PHE	2
1	A	64	ASP	1
1	A	130	ILE	1

5.2.2   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	110/128 (86%)	106±2 (97±2%)	4±2 (3±2%)	42	88
2	C	11/59 (19%)	11±0 (97±4%)	0±0 (3±4%)	42	88
All	All	2420/3740 (65%)	2342 (97%)	78 (3%)	42	88

All 29 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	9	ILE	16
1	A	109	MET	9
1	A	130	ILE	6
1	A	71	MET	5
1	A	137	ASN	5
1	A	39	LEU	4
1	A	36	MET	3
2	C	707	ASN	3
1	A	105	LEU	2
1	A	21	LYS	2
1	A	13	LYS	2
1	A	145	MET	2
2	C	705	ARG	2
1	A	60	ASN	2
1	A	143	GLN	1
1	A	54	GLU	1
1	A	115	LYS	1
1	A	124	MET	1
1	A	41	GLN	1
1	A	72	MET	1
1	A	53	ASN	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	114	GLU	1
2	C	703	ILE	1
1	A	111	ASN	1
1	A	8	GLN	1
1	A	135	GLN	1
2	C	704	LEU	1
1	A	127	GLU	1
1	A	67	GLN	1

5.2.3 RNA (i)

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.4 Carbohydrates (i)

There are no carbohydrates in this entry.

5.5 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

5.6 Other polymers (i)

There are no such molecules in this entry.

5.7 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 75% for the well-defined parts and 59% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *nef_ chemical_ shift_ list_ ShiftList_ 1-final.prot*

6.1.1 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	1766
Number of shifts mapped to atoms	1766
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

6.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	153	-1.22 ± 0.12	Should be applied
$^{13}C_\beta$	152	-0.74 ± 0.07	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	170	0.21 ± 0.22	None needed (< 0.5 ppm)

6.1.3 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 75%, i.e. 1287 atoms were assigned a chemical shift out of a possible 1725. 15 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	532/711 (75%)	269/284 (95%)	126/286 (44%)	137/141 (97%)
Sidechain	693/900 (77%)	436/520 (84%)	247/339 (73%)	10/41 (24%)

Continued on next page...

Continued from previous page...

	Total	**¹H**	**¹³C**	**¹⁵N**
Aromatic	62/114 (54%)	44/62 (71%)	17/49 (35%)	1/3 (33%)
Overall	1287/1725 (75%)	749/866 (86%)	390/674 (58%)	148/185 (80%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 59%, i.e. 1550 atoms were assigned a chemical shift out of a possible 2612. 16 out of 26 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**¹H**	**¹³C**	**¹⁵N**
Backbone	653/1105 (59%)	330/441 (75%)	153/446 (34%)	170/218 (78%)
Sidechain	835/1379 (61%)	528/802 (66%)	297/516 (58%)	10/61 (16%)
Aromatic	62/128 (48%)	44/70 (63%)	17/53 (32%)	1/5 (20%)
Overall	1550/2612 (59%)	902/1313 (69%)	467/1015 (46%)	181/284 (64%)

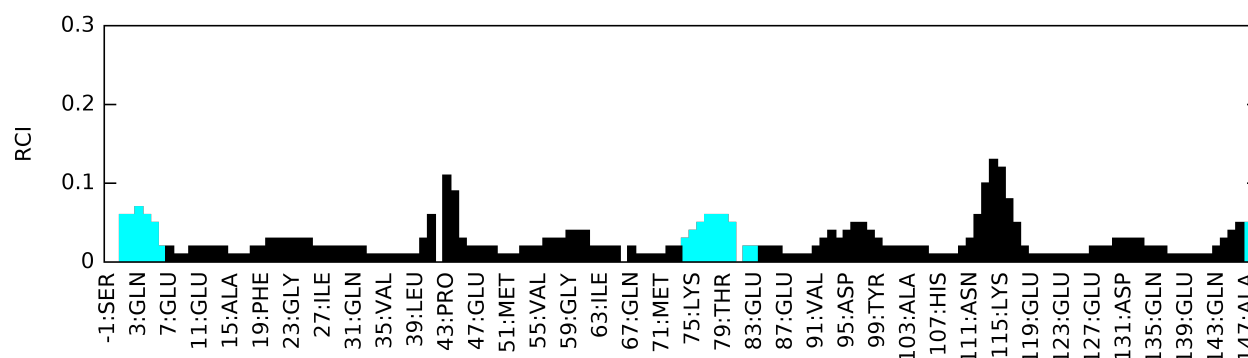### 6.1.4 Statistically unusual chemical shifts ⓘ

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	39	LEU	HA	1.96	6.66 − 1.96	-5.0

6.1.5 Random Coil Index (RCI) plots ⓘ

The images below report *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A:

Random coil index (RCI) for chain C: