



Full wwPDB NMR Structure Validation Report ⓘ

Apr 26, 2016 – 07:35 PM BST

PDB ID : 2D21
Title : NMR Structure of stereo-array isotope labelled (SAIL) maltodextrin-binding protein (MBP)
Authors : Kainosho, M.; Torizawa, T.; Iwashita, Y.; Terauchi, T.; Ono, A.M.; Guntert, P.
Deposited on : 2005-09-02

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

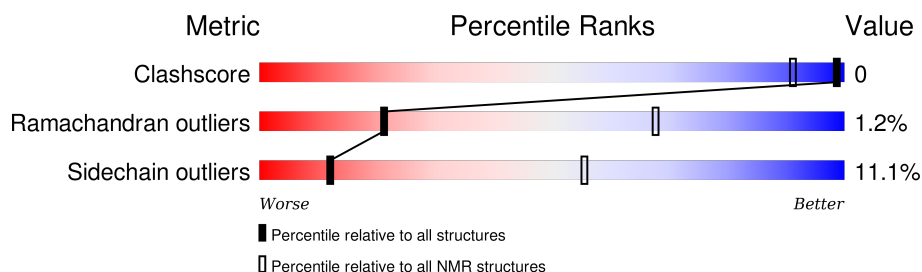
1 Overall quality at a glance

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	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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	370	 85% 11%

2 Ensemble composition and analysis

This entry contains 20 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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:3-A:111, A:260-A:324 (174)	0.48	3
2	A:116-A:117, A:126-A:226, A:247-A:258, A:331-A:370 (155)	0.54	4

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 5 clusters and 3 single-model clusters were found.

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

3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 5734 atoms, of which 2857 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Maltose-binding periplasmic protein.

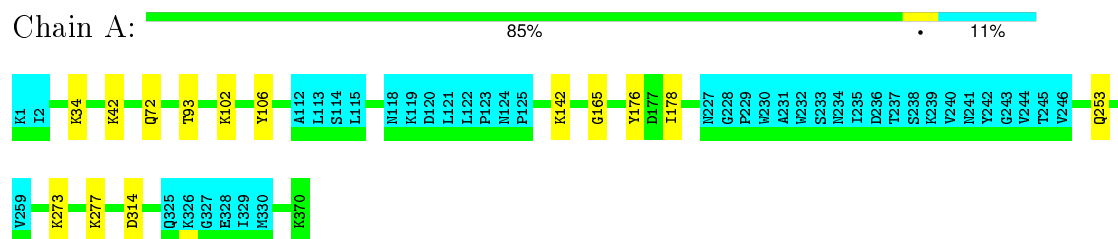
Mol	Chain	Residues	Atoms						Trace
1	A	370	Total	C	H	N	O	S	0
			5734	1853	2857	469	549	6	

4 Residue-property plots

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: Maltose-binding periplasmic protein

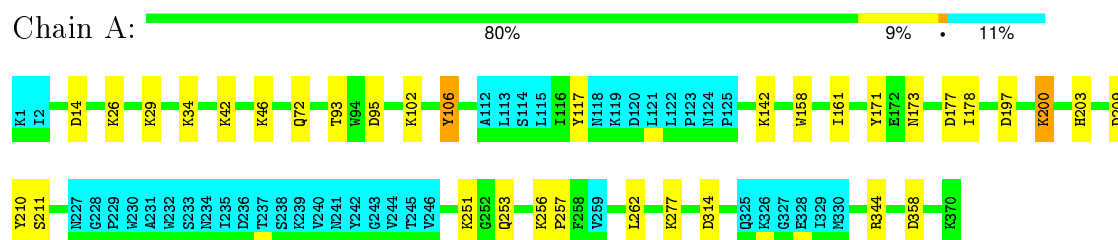


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

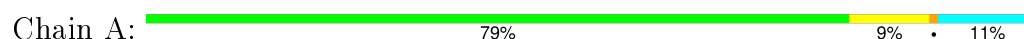
4.2.1 Score per residue for model 1

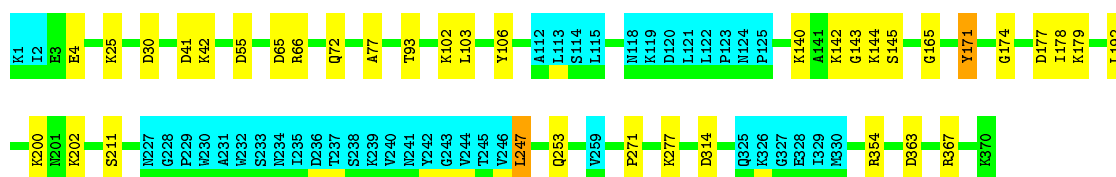
- Molecule 1: Maltose-binding periplasmic protein



4.2.2 Score per residue for model 2

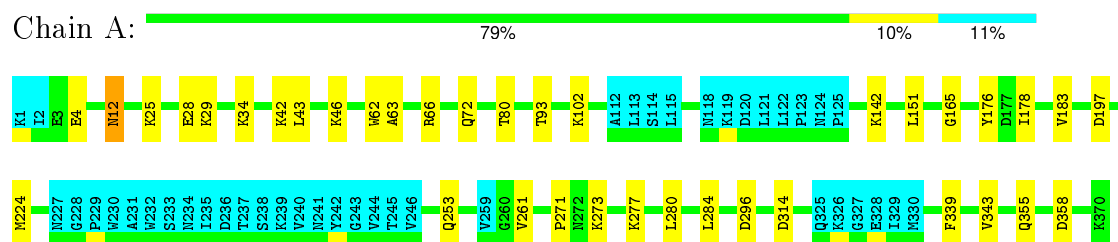
- Molecule 1: Maltose-binding periplasmic protein





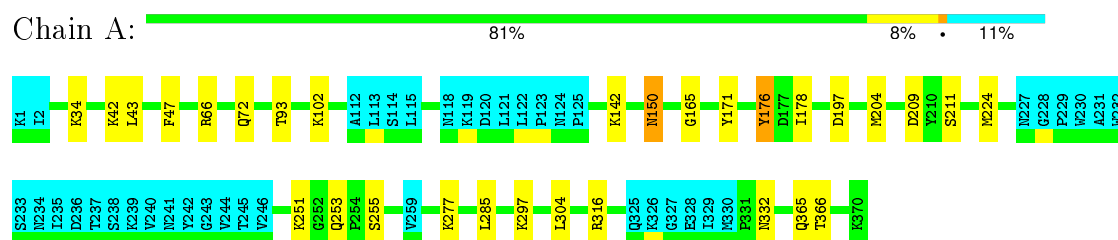
4.2.3 Score per residue for model 3 (medoid)

- Molecule 1: Maltose-binding periplasmic protein



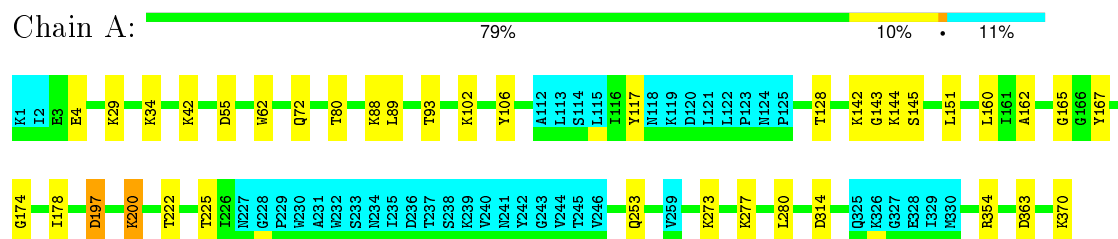
4.2.4 Score per residue for model 4

- Molecule 1: Maltose-binding periplasmic protein



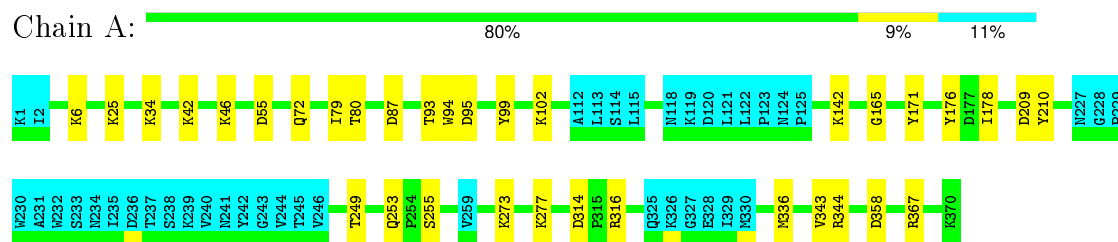
4.2.5 Score per residue for model 5

- Molecule 1: Maltose-binding periplasmic protein



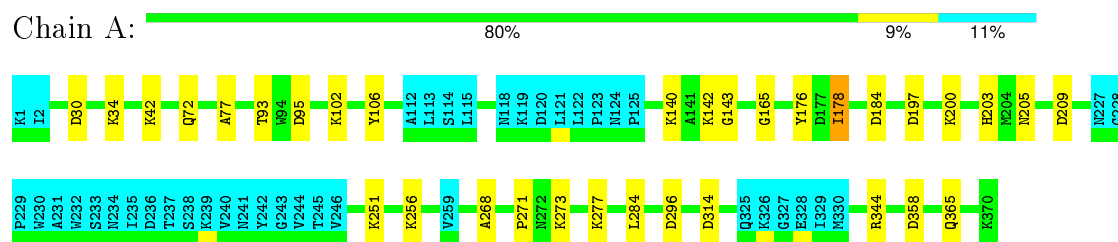
4.2.6 Score per residue for model 6

- Molecule 1: Maltose-binding periplasmic protein



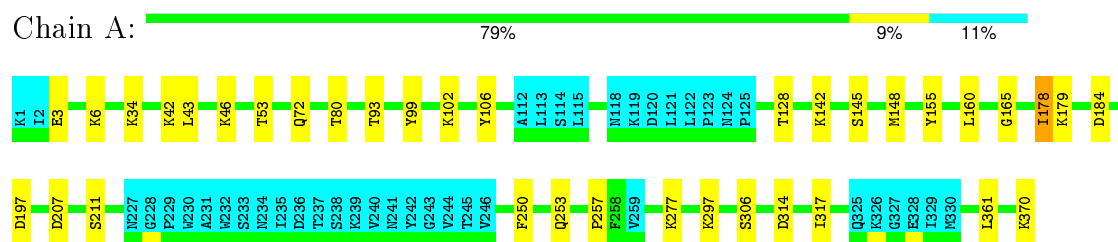
4.2.7 Score per residue for model 7

- Molecule 1: Maltose-binding periplasmic protein



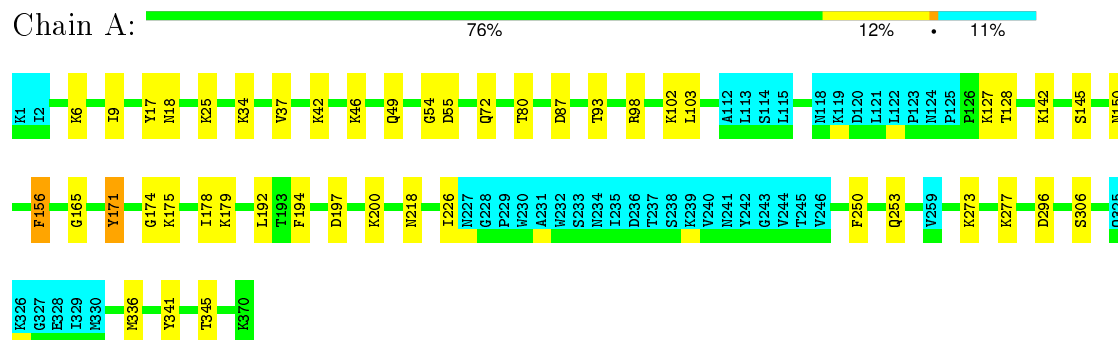
4.2.8 Score per residue for model 8

- Molecule 1: Maltose-binding periplasmic protein



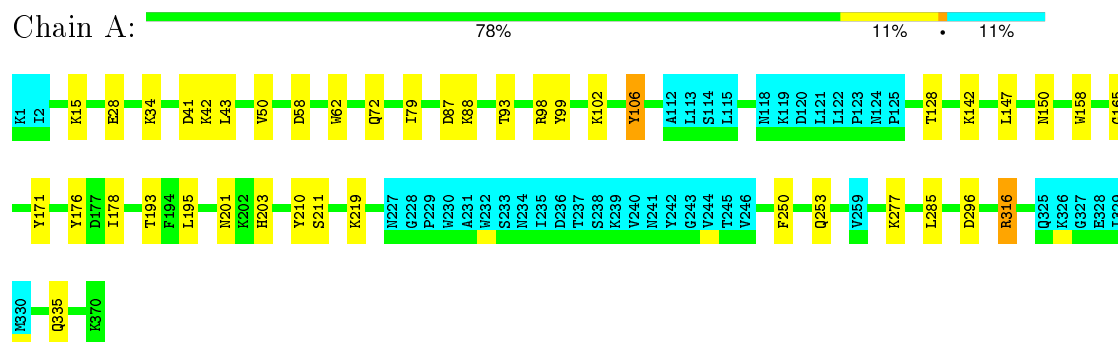
4.2.9 Score per residue for model 9

- Molecule 1: Maltose-binding periplasmic protein



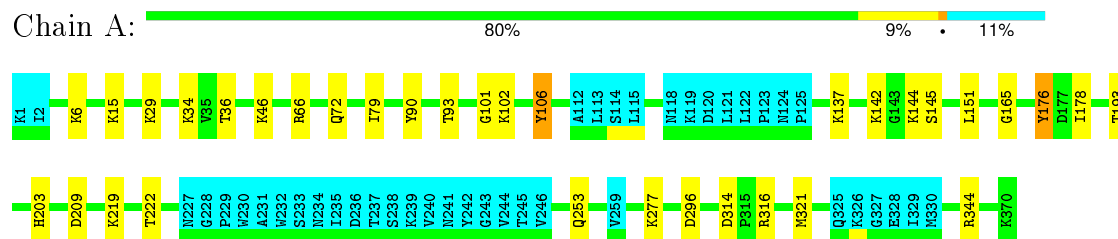
4.2.10 Score per residue for model 10

- Molecule 1: Maltose-binding periplasmic protein



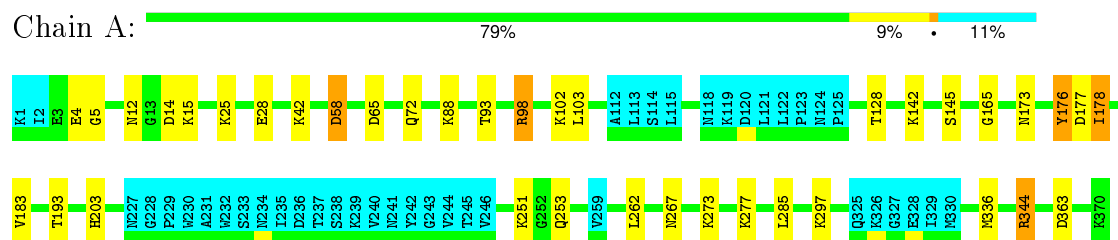
4.2.11 Score per residue for model 11

- Molecule 1: Maltose-binding periplasmic protein



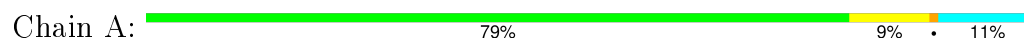
4.2.12 Score per residue for model 12

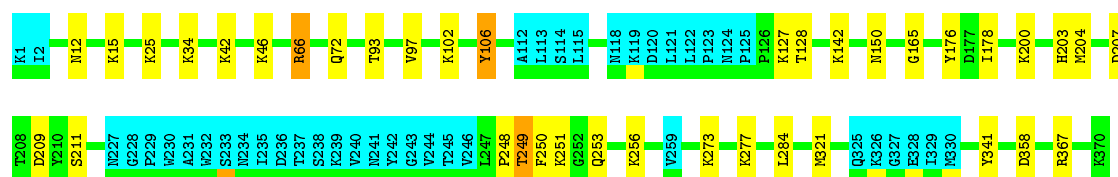
- Molecule 1: Maltose-binding periplasmic protein



4.2.13 Score per residue for model 13

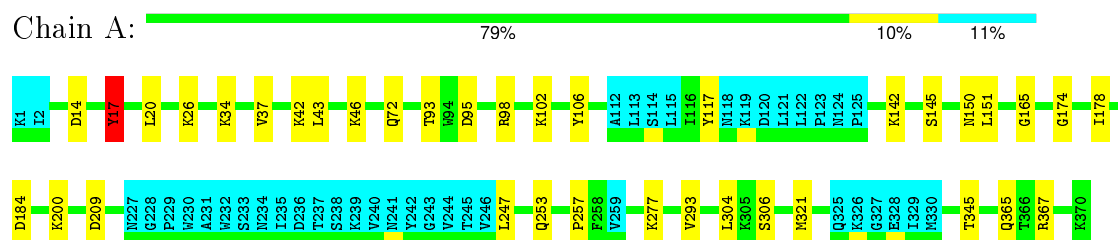
- Molecule 1: Maltose-binding periplasmic protein





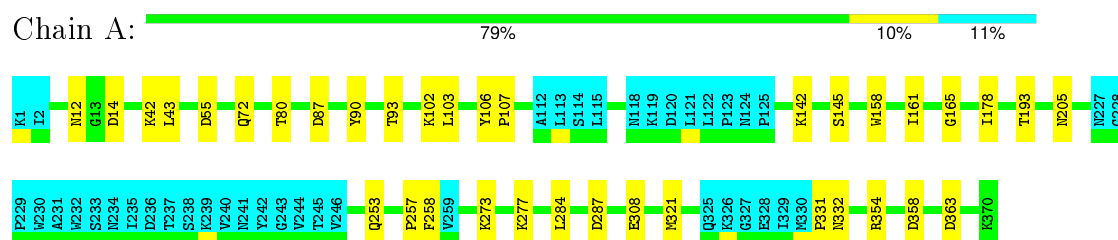
4.2.14 Score per residue for model 14

- Molecule 1: Maltose-binding periplasmic protein



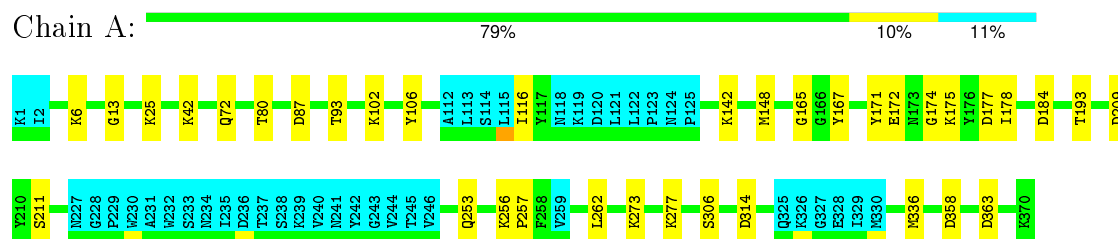
4.2.15 Score per residue for model 15

- Molecule 1: Maltose-binding periplasmic protein



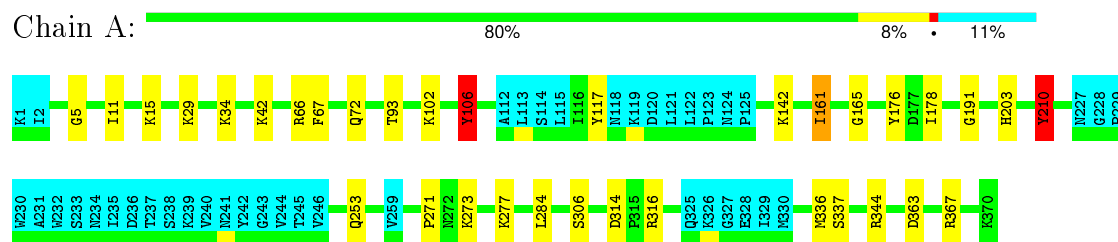
4.2.16 Score per residue for model 16

- Molecule 1: Maltose-binding periplasmic protein



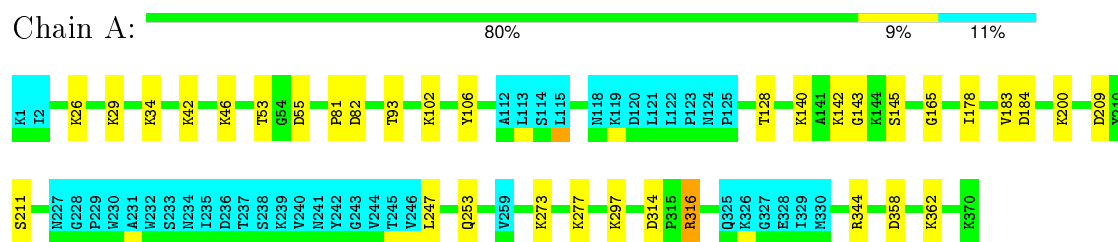
4.2.17 Score per residue for model 17

- Molecule 1: Maltose-binding periplasmic protein



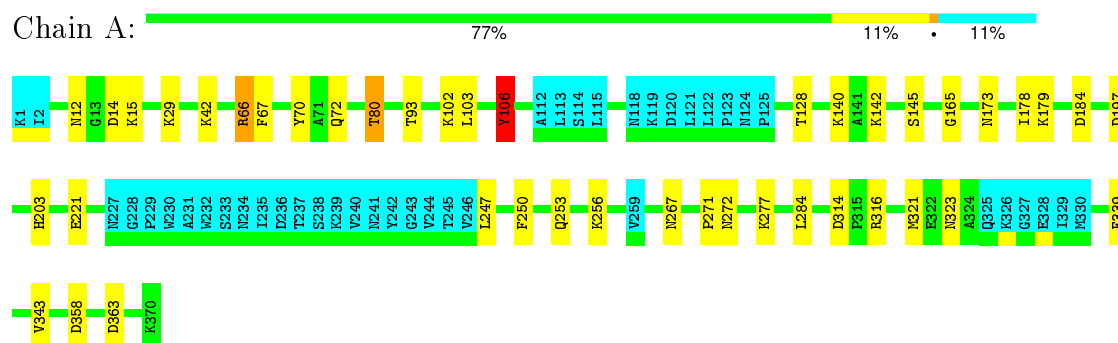
4.2.18 Score per residue for model 18

- Molecule 1: Maltose-binding periplasmic protein



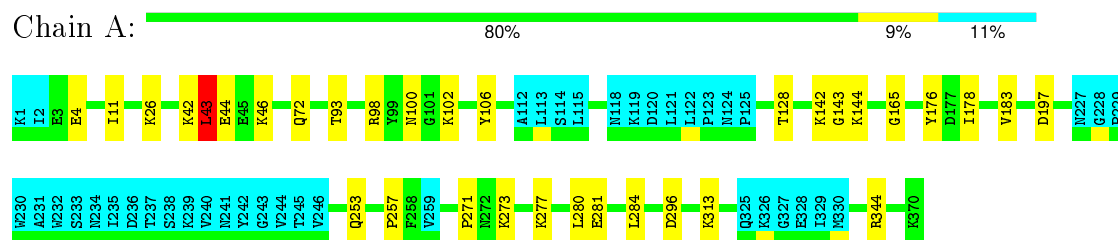
4.2.19 Score per residue for model 19

- Molecule 1: Maltose-binding periplasmic protein



4.2.20 Score per residue for model 20

- Molecule 1: Maltose-binding periplasmic protein



5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
OPALP	refinement	1.4
CYANA	structure solution	2.1

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	0.62±0.01	0±0/2621 (0.0±0.0%)	0.97±0.01	1±2/3551 (0.0±0.0%)
All	All	0.62	0/52420 (0.0%)	0.97	26/71020 (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	2.1±1.5
All	All	0	42

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(°)	Ideal(°)	Models	
								Worst	Total
1	A	354	ARG	NE-CZ-NH1	6.29	123.45	120.30	15	1
1	A	66	ARG	NE-CZ-NH2	-6.13	117.23	120.30	3	1
1	A	66	ARG	NE-CZ-NH1	6.00	123.30	120.30	3	3
1	A	344	ARG	NE-CZ-NH2	-6.00	117.30	120.30	17	3
1	A	67	PHE	CB-CG-CD2	-5.75	116.77	120.80	19	2
1	A	80	THR	CA-CB-CG2	5.70	120.38	112.40	19	1
1	A	183	VAL	CG1-CB-CG2	-5.59	101.96	110.90	12	1
1	A	30	ASP	CB-CG-OD1	5.45	123.20	118.30	7	1
1	A	98	ARG	NE-CZ-NH1	5.30	122.95	120.30	12	1
1	A	43	LEU	CB-CG-CD2	5.26	119.94	111.00	20	1
1	A	354	ARG	NE-CZ-NH2	-5.23	117.68	120.30	5	1
1	A	106	TYR	CB-CG-CD2	-5.22	117.87	121.00	17	2
1	A	171	TYR	CB-CG-CD2	-5.19	117.89	121.00	2	1
1	A	341	TYR	CB-CG-CD1	-5.13	117.92	121.00	13	1
1	A	50	VAL	CG1-CB-CG2	-5.12	102.71	110.90	10	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	58	ASP	CB-CG-OD1	-5.11	113.70	118.30	12	1
1	A	176	TYR	CB-CG-CD2	-5.04	117.97	121.00	4	1
1	A	210	TYR	CB-CG-CD2	-5.03	117.98	121.00	17	1
1	A	4	GLU	OE1-CD-OE2	-5.02	117.27	123.30	12	1
1	A	367	ARG	NE-CZ-NH2	-5.01	117.79	120.30	17	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	66	ARG	Sidechain	5
1	A	106	TYR	Sidechain	4
1	A	98	ARG	Sidechain	4
1	A	171	TYR	Sidechain	3
1	A	176	TYR	Sidechain	3
1	A	316	ARG	Sidechain	3
1	A	367	ARG	Sidechain	3
1	A	117	TYR	Sidechain	2
1	A	17	TYR	Sidechain	2
1	A	167	TYR	Sidechain	2
1	A	90	TYR	Sidechain	2
1	A	210	TYR	Sidechain	2
1	A	156	PHE	Sidechain	1
1	A	70	TYR	Sidechain	1
1	A	47	PHE	Sidechain	1
1	A	354	ARG	Sidechain	1
1	A	155	TYR	Sidechain	1
1	A	344	ARG	Sidechain	1
1	A	341	TYR	Sidechain	1

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	2560	2533	2533	1±1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes
All	All	51200	50660	50660	22

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:43:LEU:HD22	1:A:44:GLU:N	0.54	2.16	20	1
1:A:197:ASP:HA	1:A:200:LYS:HE3	0.50	1.83	1	2
1:A:339:PHE:O	1:A:343:VAL:HG23	0.49	2.07	3	1
1:A:161:ILE:HD13	1:A:161:ILE:H	0.49	1.67	17	1
1:A:77:ALA:HB2	1:A:268:ALA:HA	0.48	1.85	7	1
1:A:150:ASN:HD22	1:A:210:TYR:HB2	0.47	1.69	10	1
1:A:161:ILE:HG22	1:A:191:GLY:HA3	0.46	1.87	17	1
1:A:17:TYR:CE1	1:A:37:VAL:HG22	0.46	2.46	14	1
1:A:20:LEU:HA	1:A:293:VAL:HG22	0.46	1.88	14	1
1:A:12:ASN:HB3	1:A:43:LEU:HD21	0.46	1.88	3	1
1:A:339:PHE:O	1:A:343:VAL:HG22	0.45	2.11	19	1
1:A:249:THR:HG22	1:A:255:SER:H	0.45	1.71	6	1
1:A:79:ILE:HD12	1:A:106:TYR:CD2	0.42	2.49	10	2
1:A:317:ILE:HD12	1:A:317:ILE:H	0.42	1.75	8	1
1:A:345:THR:HG21	1:A:367:ARG:HH22	0.41	1.75	14	1
1:A:285:LEU:HD22	1:A:304:LEU:HD21	0.41	1.92	4	1
1:A:156:PHE:CZ	1:A:226:ILE:HG21	0.40	2.51	9	1
1:A:62:TRP:CG	1:A:63:ALA:N	0.40	2.88	3	1
1:A:93:THR:HG22	1:A:107:PRO:CB	0.40	2.46	15	1
1:A:9:ILE:HD12	1:A:37:VAL:HG22	0.40	1.92	9	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	328/370 (89%)	301±4 (92±1%)	23±5 (7±1%)	4±1 (1±0%)	21	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	6560/7400 (89%)	6022 (92%)	461 (7%)	77 (1%)	21	68

All 28 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	165	GLY	19
1	A	80	THR	7
1	A	271	PRO	6
1	A	257	PRO	6
1	A	143	GLY	5
1	A	174	GLY	5
1	A	14	ASP	3
1	A	173	ASN	3
1	A	178	ILE	3
1	A	5	GLY	2
1	A	332	ASN	1
1	A	77	ALA	1
1	A	81	PRO	1
1	A	3	GLU	1
1	A	54	GLY	1
1	A	248	PRO	1
1	A	101	GLY	1
1	A	150	ASN	1
1	A	100	ASN	1
1	A	202	LYS	1
1	A	221	GLU	1
1	A	162	ALA	1
1	A	251	LYS	1
1	A	331	PRO	1
1	A	285	LEU	1
1	A	13	GLY	1
1	A	249	THR	1
1	A	145	SER	1

6.3.2 Protein sidechains ⓘ

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	261/297 (88%)	232±3 (89±1%)	29±3 (11±1%)	12	55
All	All	5220/5940 (88%)	4640 (89%)	580 (11%)	12	55

All 129 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	178	ILE	20
1	A	142	LYS	20
1	A	277	LYS	20
1	A	102	LYS	20
1	A	42	LYS	19
1	A	72	GLN	19
1	A	253	GLN	19
1	A	93	THR	19
1	A	106	TYR	15
1	A	34	LYS	14
1	A	314	ASP	12
1	A	273	LYS	12
1	A	46	LYS	10
1	A	209	ASP	9
1	A	176	TYR	9
1	A	128	THR	9
1	A	358	ASP	9
1	A	145	SER	9
1	A	200	LYS	8
1	A	211	SER	8
1	A	197	ASP	8
1	A	203	HIS	8
1	A	363	ASP	7
1	A	29	LYS	7
1	A	25	LYS	7
1	A	284	LEU	7
1	A	55	ASP	6
1	A	184	ASP	6
1	A	15	LYS	6
1	A	43	LEU	6
1	A	296	ASP	6
1	A	316	ARG	5
1	A	6	LYS	5
1	A	103	LEU	5
1	A	256	LYS	5
1	A	250	PHE	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	171	TYR	5
1	A	306	SER	5
1	A	336	MET	5
1	A	321	MET	5
1	A	87	ASP	5
1	A	344	ARG	5
1	A	12	ASN	5
1	A	193	THR	5
1	A	151	LEU	4
1	A	95	ASP	4
1	A	26	LYS	4
1	A	251	LYS	4
1	A	297	LYS	4
1	A	140	LYS	4
1	A	150	ASN	4
1	A	247	LEU	4
1	A	177	ASP	4
1	A	144	LYS	4
1	A	4	GLU	4
1	A	179	LYS	4
1	A	262	LEU	3
1	A	88	LYS	3
1	A	280	LEU	3
1	A	183	VAL	3
1	A	365	GLN	3
1	A	99	TYR	3
1	A	161	ILE	3
1	A	158	TRP	3
1	A	28	GLU	3
1	A	175	LYS	2
1	A	219	LYS	2
1	A	192	LEU	2
1	A	267	ASN	2
1	A	41	ASP	2
1	A	53	THR	2
1	A	127	LYS	2
1	A	222	THR	2
1	A	14	ASP	2
1	A	62	TRP	2
1	A	370	LYS	2
1	A	207	ASP	2
1	A	11	ILE	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	117	TYR	2
1	A	65	ASP	2
1	A	224	MET	2
1	A	210	TYR	2
1	A	204	MET	2
1	A	148	MET	2
1	A	58	ASP	2
1	A	205	ASN	2
1	A	160	LEU	2
1	A	261	VAL	1
1	A	137	LYS	1
1	A	194	PHE	1
1	A	308	GLU	1
1	A	30	ASP	1
1	A	17	TYR	1
1	A	201	ASN	1
1	A	287	ASP	1
1	A	98	ARG	1
1	A	97	VAL	1
1	A	18	ASN	1
1	A	366	THR	1
1	A	323	ASN	1
1	A	332	ASN	1
1	A	249	THR	1
1	A	343	VAL	1
1	A	172	GLU	1
1	A	304	LEU	1
1	A	255	SER	1
1	A	36	THR	1
1	A	195	LEU	1
1	A	218	ASN	1
1	A	94	TRP	1
1	A	79	ILE	1
1	A	362	LYS	1
1	A	337	SER	1
1	A	82	ASP	1
1	A	258	PHE	1
1	A	66	ARG	1
1	A	80	THR	1
1	A	345	THR	1
1	A	313	LYS	1
1	A	89	LEU	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	147	LEU	1
1	A	355	GLN	1
1	A	281	GLU	1
1	A	49	GLN	1
1	A	285	LEU	1
1	A	272	ASN	1
1	A	335	GLN	1
1	A	361	LEU	1
1	A	225	THR	1

6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation

No chemical shift data were provided