



wwPDB NMR Structure Validation Summary Report ⓘ

Mar 19, 2026 – 12:00 PM UTC

PDB ID : 9PGN / pdb_00009pgn
BMRB ID : 4986
Title : Two-conformer equilibrium of maltose-binding protein in the absence of ligand from residual dipolar coupling analysis
Authors : Shen, Y.; Bax, A.
Deposited on : 2025-07-08

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
wwPDB-RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
wwPDB-ShiftChecker : v1.2
BMRB Restraints Analysis : v1.2
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

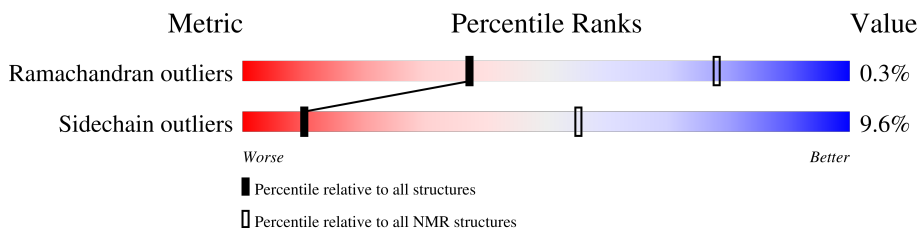
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 35%.

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)
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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	369	

2 Ensemble composition and analysis

This entry contains 10 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:5-A:105, A:266-A:312 (148)	0.29	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 3 clusters and 1 single-model cluster was found.

Cluster number	Models
1	2, 3, 5, 6
2	8, 9, 10
3	1, 7
Single-model clusters	4

3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 5095 atoms, of which 551 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Maltose/maltodextrin ABC transporter substrate-binding protein MalE.

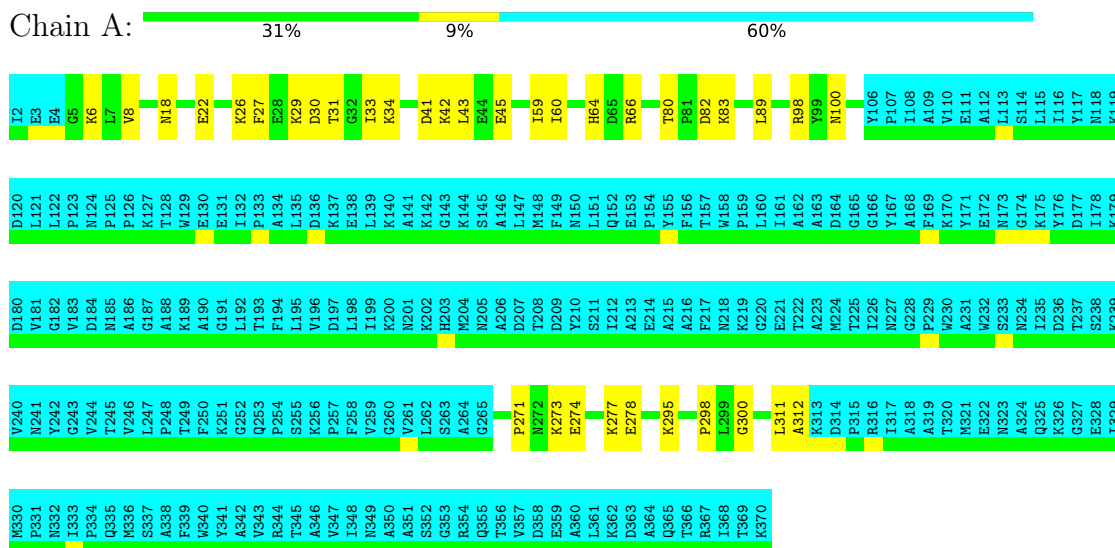
Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
1	A	369	5095	2923	551	742	867	12	0

4 Residue-property plots [i](#)

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.

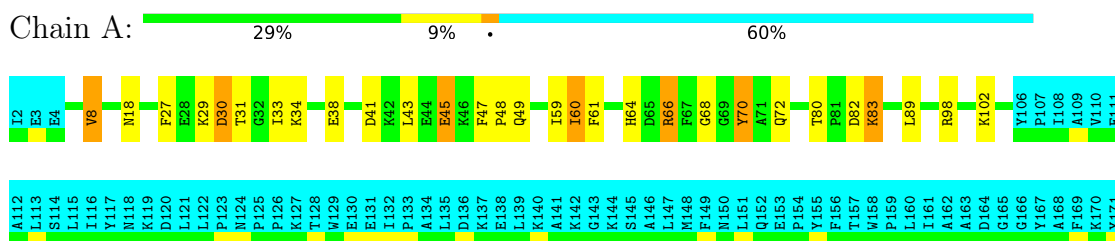
- Molecule 1: Maltose/maltodextrin ABC transporter substrate-binding protein MalE



4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: Maltose/maltodextrin ABC transporter substrate-binding protein MalE



T320	V232	E172
M321	S233	N173
E322	N234	G174
N323	I235	K175
A324	D236	Y176
Q325	T237	D177
K326	S238	I178
G327	K239	K179
E328	V240	D180
I329	N241	V181
M330	Y242	G182
P331	G243	V183
N332	V244	D184
I333	T245	N185
P334	V246	A186
Q335	L247	G187
M336	P248	A188
S337	T249	K189
A338	F250	A190
F339	K251	G191
W340	G252	L192
Y341	Q253	T193
A342	P254	F194
V343	S255	L195
K344	K256	V196
T345	P257	D197
A346	F258	L198
V347	V259	I199
I348	G260	K200
N349	V261	N201
A350	L262	K202
A351	S263	H203
S352	A264	M204
G353	G265	N205
R354		A206
Q355	K277	D207
T356	E278	T208
V357		D209
D358	E281	Y210
E359		S211
A360	V293	I212
L361	N294	A213
K362	K295	E214
D363	D296	A215
A364	K297	A216
Q365	P298	F217
T366	L299	N218
R367	G300	K219
I368		G220
T369	L304	F221
K370	E308	T222
		A223
		M224
		T225
		I226
		N227
		G228
		P229
		W230
		A231

5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 10 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
DYNAMO	structure calculation	
X-PLOR NIH	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1686
Number of shifts mapped to atoms	1683
Number of unparsed shifts	0
Number of shifts with mapping errors	3
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	35%

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	0.85±0.54	3±6/1186 (0.3± 0.5%)	1.53±0.86	34±56/1605 (2.1± 3.5%)
All	All	1.01	33/11860 (0.3%)	1.76	344/16050 (2.1%)

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	0.4±0.9
All	All	0	4

5 of 27 unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	289	GLY	CA-C	6.84	1.59	1.52	9	1
1	A	312	ALA	C-O	-6.69	1.16	1.23	8	1
1	A	22	GLU	CA-CB	-6.46	1.42	1.53	9	1
1	A	55	ASP	N-CA	6.43	1.54	1.46	8	1
1	A	300	GLY	N-CA	6.30	1.54	1.45	9	1

5 of 284 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	30	ASP	CA-CB-CG	30.69	143.29	112.60	9	3
1	A	22	GLU	CA-CB-CG	26.93	167.96	114.10	9	1
1	A	312	ALA	CA-C-N	25.94	171.08	121.54	8	2
1	A	312	ALA	C-N-CA	25.94	171.08	121.54	8	2
1	A	311	LEU	CA-C-N	-18.92	94.84	121.33	8	2

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	312	ALA	Peptide	1
1	A	66	ARG	Sidechain	1
1	A	98	ARG	Sidechain	1
1	A	102	LYS	Mainchain	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	-2191	-265	1141	15±8
All	All	-21910	-2650	11410	146

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.

5 of 125 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:311:LEU:O	1:A:312:ALA:HB2	1.08	1.34	8	1
1:A:31:THR:HG23	1:A:33:ILE:H	1.04	1.09	8	1
1:A:311:LEU:O	1:A:312:ALA:CB	0.98	1.98	8	1
1:A:31:THR:CG2	1:A:33:ILE:H	0.94	1.75	8	1
1:A:43:LEU:CD1	1:A:60:ILE:HD11	0.91	1.95	9	3

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	148/369 (40%)	142±3 (96±2%)	5±2 (4±1%)	0±1 (0±0%)	37 78
All	All	1480/3690 (40%)	1421 (96%)	54 (4%)	5 (0%)	37 78

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	81	PRO	2
1	A	312	ALA	1
1	A	83	LYS	1
1	A	82	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	118/296 (40%)	107±6 (90±5%)	11±6 (10±5%)	10 55
All	All	1180/2960 (40%)	1067 (90%)	113 (10%)	10 55

5 of 50 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	6	LYS	7
1	A	89	LEU	5
1	A	274	GLU	5
1	A	30	ASP	4
1	A	45	GLU	4

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 oligosaccharides 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 (i)

The completeness of assignment taking into account all chemical shift lists is 35% for the well-defined parts and 34% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *shift_set_1*

7.1.1 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	1686
Number of shifts mapped to atoms	1683
Number of unparsed shifts	0
Number of shifts with mapping errors	3
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. All 3 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	.	LYS	C	176.4	0.05	1
1	A	.	LYS	CA	55.8	0.05	1
1	A	.	LYS	CB	32.2	0.05	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	359	0.36 ± 0.13	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	327	0.95 ± 0.10	Should be checked
$^{13}\text{C}'$	332	-0.02 ± 0.10	None needed (< 0.5 ppm)
^{15}N	334	1.04 ± 0.23	Should be applied

7.1.3 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 35%, i.e. 701 atoms were assigned a chemical shift out of a possible 1979. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	567/740 (77%)	140/303 (46%)	287/296 (97%)	140/141 (99%)
Sidechain	134/1070 (13%)	0/693 (0%)	134/345 (39%)	0/32 (0%)
Aromatic	0/169 (0%)	0/85 (0%)	0/84 (0%)	0/0 (—%)
Overall	701/1979 (35%)	140/1081 (13%)	421/725 (58%)	140/173 (81%)

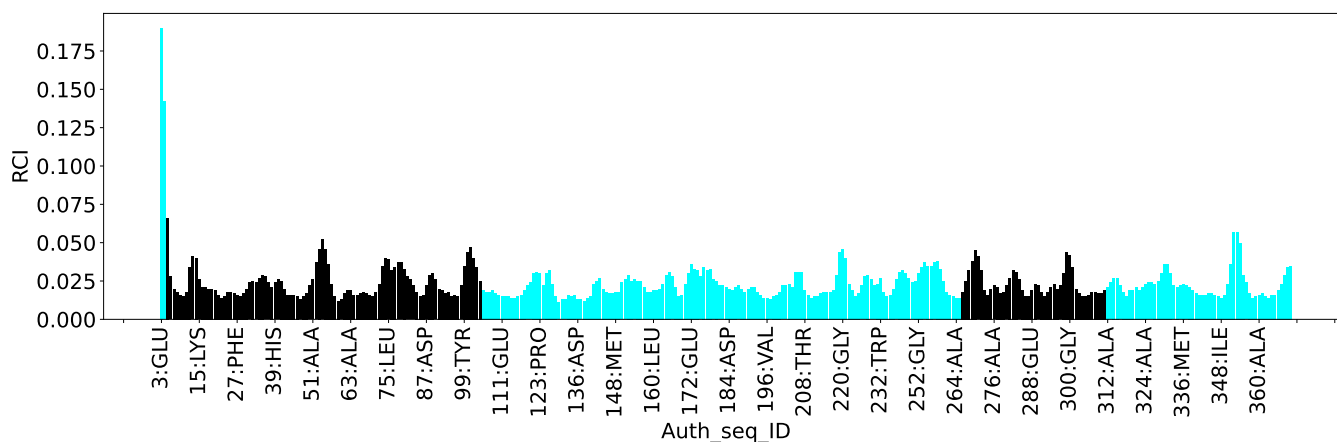
7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *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. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:



8 NMR restraints analysis

No restraints data found

9 Distance violation analysis

No distance restraints data found

10 Dihedral-angle violation analysis

No dihedral-angle restraints found