



# wwPDB NMR Structure Validation Summary Report ⓘ

Oct 23, 2021 – 05:44 PM EDT

PDB ID : 1EZP  
Title : GLOBAL FOLD OF MALTODEXTRIN BINDING PROTEIN COMPLEXED WITH BETA-CYCLODEXTRIN USING PEPTIDE ORIENTATIONS FROM DIPOLAR COUPLINGS  
Authors : Mueller, G.A.; Choy, W.Y.; Yang, D.; Forman-Kay, J.D.; Venters, R.A.; Kay, L.E.  
Deposited on : 2000-05-11

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
ShiftChecker : 2.23.2  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

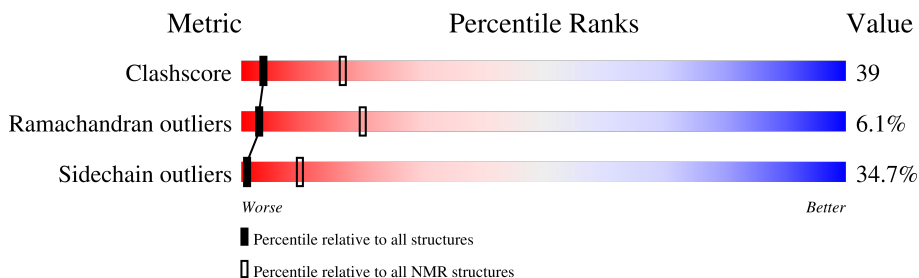
# 1 Overall quality at a glance

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	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  $\geq 3$ , 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	

## 2 Ensemble composition and analysis

This entry contains 10 models. Model 9 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:4-A:231, A:241-A:370 (358)	1.08	9

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 2 clusters and 1 single-model cluster was found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5735 atoms, of which 2858 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called MALTODEXTRIN BINDING PERIPLASMIC PROTEIN.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	370	5735	1851	2858	469	551	6	0

There is a discrepancy between the modelled and reference sequences:

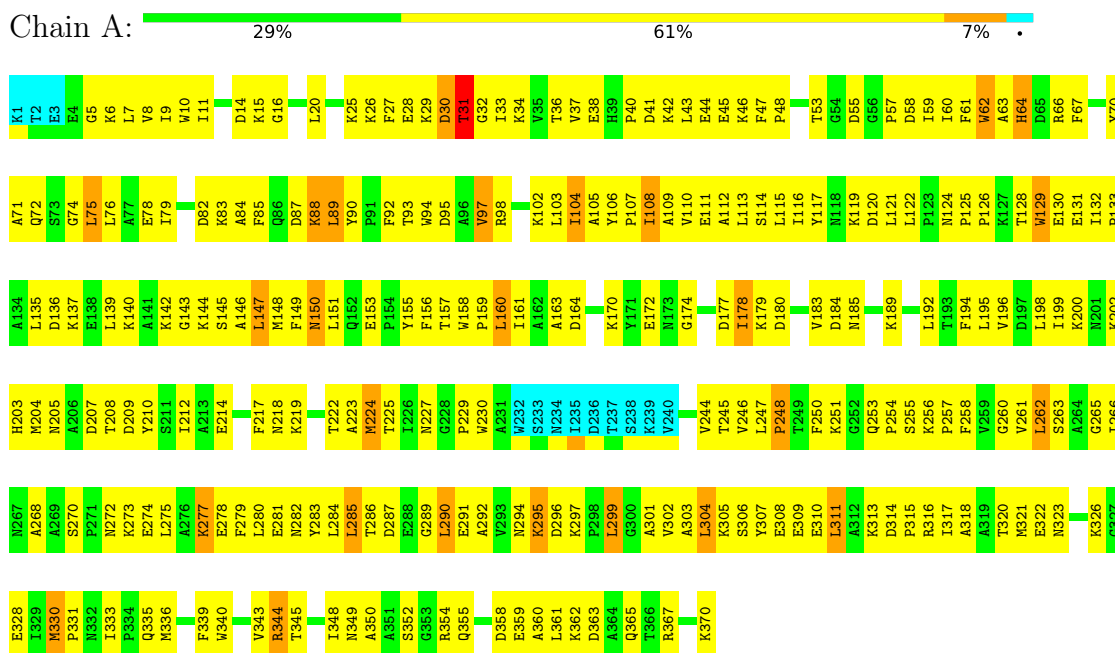
Chain	Residue	Modelled	Actual	Comment	Reference
A	2	THR	ILE	engineered mutation	UNP P02928

## 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.

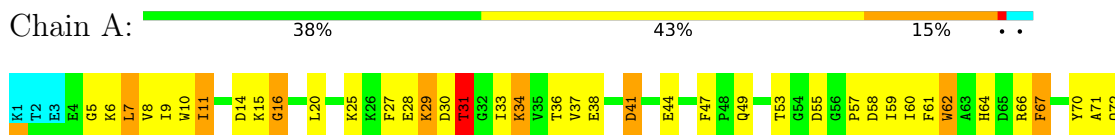
- Molecule 1: MALTODEXTRIN BINDING PERIPLASMIC PROTEIN



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

The representative model is number 9. Colouring as in section 4.1 above.

- Molecule 1: MALTODEXTRIN BINDING PERIPLASMIC PROTEIN



S73	A141	A213	D287	K362
G74	K142	E214	E288	D363
L76	G143	F217	G289	A364
A77	K144	N218	L290	D365
E78	S145	A292	E291	T366
	A146	V293	V293	B367
	L147	T222	N294	
	M148	A223	K295	
K83	F149	M224	D296	
A84	N150	T225	K297	
F85	F151	I226	P298	
Q86	L152	N227	L299	
D87	Q152	W232	G300	
K88	E153	S233	A301	
L89	F156	M234	V302	
Y90	T157	I235	A303	
	W158	D236	L304	
F92	P159	T237	K305	
T93	L160	S238	S306	
Q94	I161	K239	Y307	
D95	A162	V240	E308	
A96	A163	T245	E309	
Y97	D164	W246	E310	
R98	R167	L247	L311	
	G101	P248	A312	
	K102	T249	K313	
	L103	F250	D314	
	L104	P254	P315	
	A105	S255	R316	
Y106	N173	K256	I317	
P107	G174	F257	A318	
I108	A109	F258	E322	
	V110	V259	Q325	
	L113	G260	E328	
	S114	V261	I329	
L115	M185	L262	M330	
L116	G191	S263	I333	
	L192	A264	M336	
D120	T193	G265	F339	
L121	F194	I266	R344	
L122	L195	N267	T345	
P123	V196	A268	N349	
N124	D197	K273	A350	
P125	L198	E274	R354	
P126	I199	L275	Q355	
	M201	A276	D358	
W129	K200	K277	E359	
E130	M202	E278	A360	
E131	H203	K278	L361	
I132	M204	F279		
P133	N205	L280		
A134	A206	E281		
L135	D207	N282		
D136	T208	Y283		
K137	D209	L284		
E138	Y210	L285		
L139	S211	T286		
K140	I212			

## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing from extended coordinates torsion angle dynamics and finish with cartesian dynamics.*

Of the 243 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
CNS	structure solution	0.5
CNS	refinement	0.5

No chemical shift data was provided.

## 6 Model quality [i](#)

### 6.1 Standard geometry [i](#)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2779	2760	2757	217±16
All	All	27790	27600	27570	2171

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 39.

5 of 1469 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:HB3	1:A:317:ILE:HG21	1.10	1.23	7	1
1:A:292:ALA:HA	1:A:295:LYS:HG2	1.05	1.20	4	1
1:A:217:PHE:HB2	1:A:225:THR:HG23	1.05	1.26	8	2
1:A:77:ALA:HB2	1:A:268:ALA:HA	1.05	1.25	7	4
1:A:109:ALA:HB3	1:A:262:LEU:HB2	1.04	1.27	7	8

### 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	357/370 (96%)	280±3 (79±1%)	55±4 (15±1%)	22±3 (6±1%)	3	20
All	All	3570/3700 (96%)	2804 (79%)	547 (15%)	219 (6%)	3	20

5 of 66 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	16	GLY	10
1	A	31	THR	10
1	A	143	GLY	10
1	A	257	PRO	10
1	A	311	LEU	10

### 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	285/297 (96%)	186±6 (65±2%)	99±6 (35±2%)	1	10
All	All	2850/2970 (96%)	1861 (65%)	989 (35%)	1	10

5 of 227 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	195	LEU	10
1	A	299	LEU	10
1	A	89	LEU	9
1	A	103	LEU	9
1	A	148	MET	9

### 6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 6.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.6 Ligand geometry [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