wwPDB NMR Structure Validation Summary Report (i)

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*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

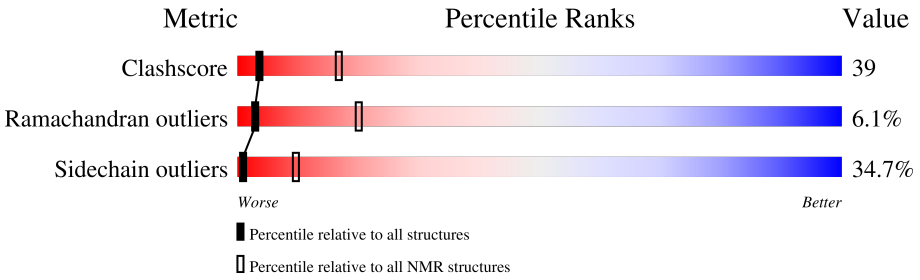
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.23.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

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 >=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	370	29% 61% 7% •

2 Ensemble composition and analysis (i)

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 (i)

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

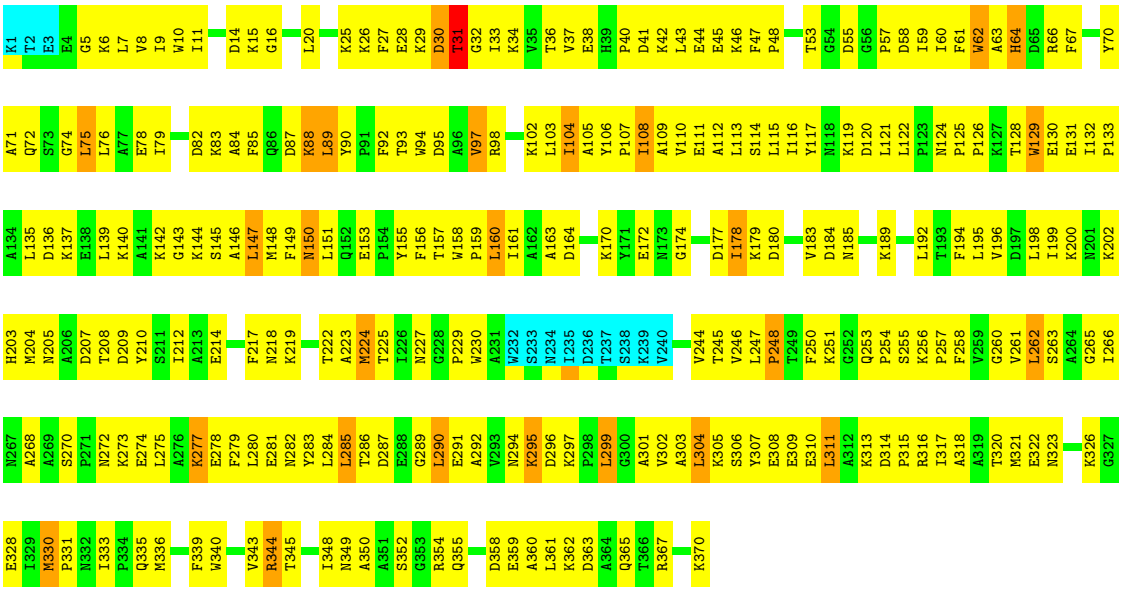
4 Residue-property plots (i)

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

Chain A: 29% 61% 7% •

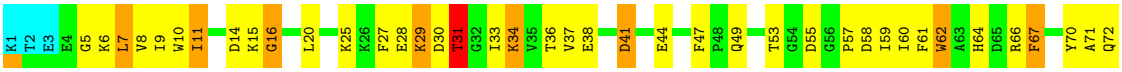


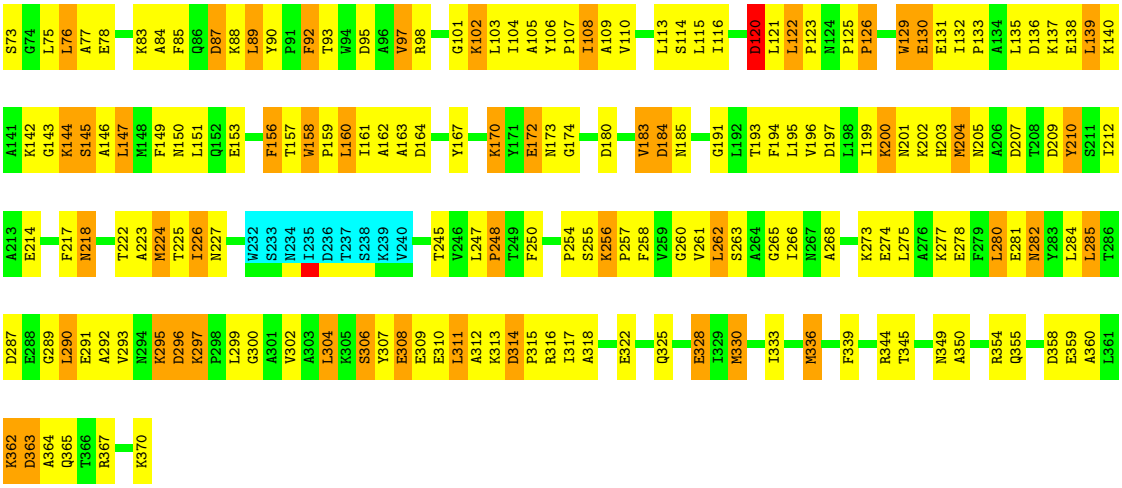
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

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

• Molecule 1: MALTODEXTRIN BINDING PERIPLASMIC PROTEIN

Chain A: 38% 43% 15% • •




S73 G74 L75 L76 A77 E78 K83 A84 F85 Q86 D87 K88 L89 Y90 P91 F92 T93 W94 D95 A96 V97 R98 G101 K102 L103 I104 A105 Y106 P107 I108 A109 V110 L113 S114 L115 I116 D120 L121 L122 P123 N124 P125 P126 W129 E130 E131 I132 P133 A134 L135 D136 K137 E138 L139 K140
A141 K142 G143 K144 S145 A146 L147 M148 F149 N150 L151 Q152 E153 F156 T157 W158 P159 L160 I161 A162 A163 D164 Y167 K170 Y171 E172 N173 G174 D180 V183 D184 N185 G191 L192 T193 F194 L195 V196 D197 L198 I199 K200 N201 K202 H203 M204 N205 A206 D207 T208 D209 Y210 S211 I212
A213 E214 F217 N218 T222 A223 M224 T225 I226 N227 W232 S233 N234 I235 D236 T237 S238 K239 V240 T245 V246 L247 P248 T249 F250 P254 S255 K256 P257 F258 V259 G260 V261 L262 S263 A264 G265 I266 N267 A268 K273 E274 L275 A276 K277 E278 F279 L280 E281 N282 Y283 L284 L285 T286
D287 E288 G289 L290 E291 A292 V293 N294 K295 D296 K297 P298 L299 G300 A301 V302 A303 L304 K305 S306 Y307 E308 E309 E310 L311 A312 K313 D314 P315 R316 I317 A318 E322 Q325 E328 I329 M330 I333 M336 F339 R344 T345 N349 A350 R354 Q355 D358 E359 A360 L361
K362 D363 A364 Q365 T366 R367 K370

5 Refinement protocol and experimental data overview (i)

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 (i)

No chemical shift data were provided