wwPDB NMR Structure Validation Summary Report (i)

May 7, 2024 – 01:00 pm BST

PDB ID	:	2XDI
Title	:	Tryptophan repressor with L75F mutation in its apo form (no L- tryptophan bound)
Authors	:	Tyler, R.; Copie, V.
Deposited on	:	2010-05-03

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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

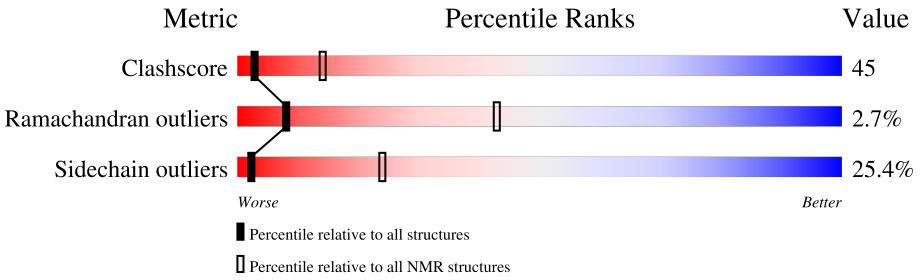
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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	107	10% 43% 9% 36% .
1	B	107	9% 44% 9% 36% .

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:15-A:63, A:89-A:106, B:15-B:63, B:89-B:106 (134)	1.53	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 3 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3408 atoms, of which 1712 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called TRP OPERON REPRESSOR.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	105	1704	531	856	153	161	3	0
1	B	105	1704	531	856	153	161	3	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	75	PHE	LEU	engineered mutation	UNP P0A882
B	75	PHE	LEU	engineered mutation	UNP P0A882

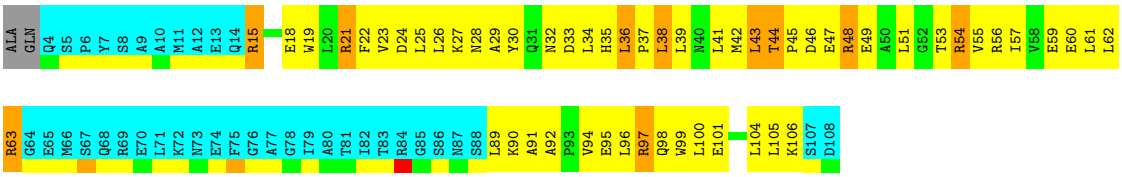
4 Residue-property plots (i)

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.

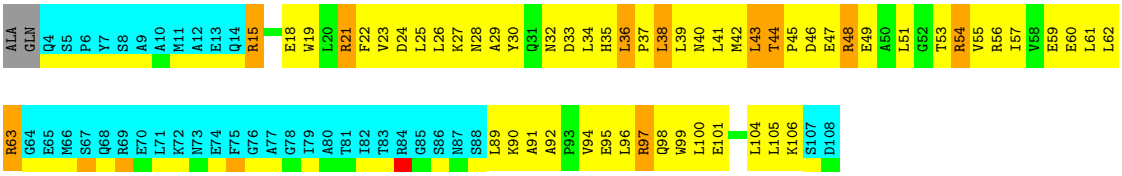
• Molecule 1: TRP OPERON REPRESSOR

Chain A: 10% 43% 9% 36%



• Molecule 1: TRP OPERON REPRESSOR

Chain B: 9% 44% 9% 36%

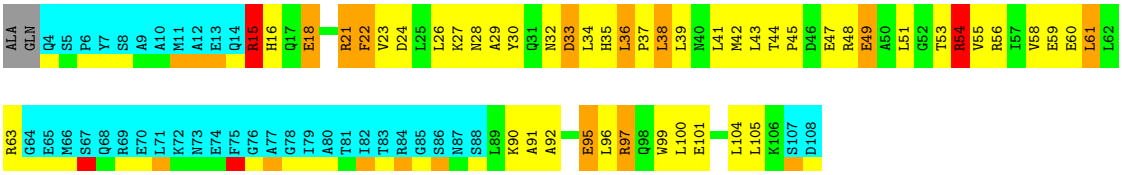


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

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

• Molecule 1: TRP OPERON REPRESSOR

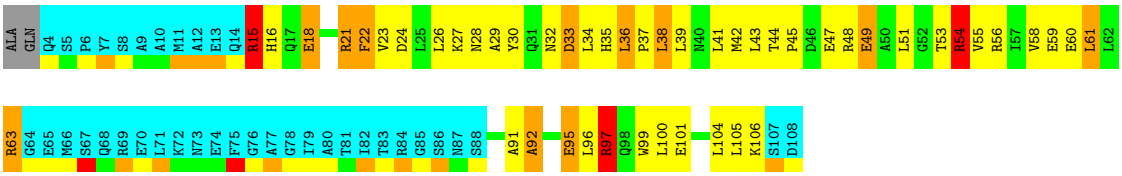
Chain A: 17% 35% 9% • 36% •



• Molecule 1: TRP OPERON REPRESSOR

Chain B:

17%
33%
10%
36%


ALA
GLN
Q4
S5
P6
Y7
S8
A9
A10
M11
A12
E13
Q14
R15
H16
Q17
E18
R21
F22
V23
D24
L25
L26
K27
N28
A29
Y30
Q31
N32
D33
L34
H35
L36
P37
L38
L39
N40
L41
M42
L43
T44
P45
D46
E47
R48
E49
A50
L51
G52
T53
R54
V55
R56
I57
V58
E59
E60
L61
L62

R63
G64
E65
M66
S67
Q68
R69
E70
L71
K72
N73
E74
F75
G76
A77
G78
I79
A80
T81
I82
T83
R84
G85
S86
N87
S88
A91
A92
E95
L96
R97
Q98
W99
L100
E101
L104
L105
K106
S107
D108

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *CNS-XPLOR*.

Of the 40 calculated structures, 20 were deposited, based on the following criterion: *LEAST RESTRAINT VIOLATION*.

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

Software name	Classification	Version
CNS	refinement	
NMRView	structure solution	

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.

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	6.5±0.8
1	B	0.0±0.0	6.3±0.7
All	All	0	256

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

5 of 14 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	21	ARG	Sidechain	20
1	A	54	ARG	Sidechain	20
1	B	21	ARG	Sidechain	20
1	A	48	ARG	Sidechain	19
1	B	48	ARG	Sidechain	19

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	568	594	594	60±13
1	B	568	594	594	60±14
All	All	22720	23760	23760	2098

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 45.

5 of 905 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:18:GLU:O	1:A:22:PHE:CD1	1.33	1.80	19	4
1:A:18:GLU:O	1:A:22:PHE:HD1	1.22	1.08	19	1
1:A:34:LEU:HD13	1:B:99:TRP:CE3	1.04	1.88	20	2
1:B:53:THR:O	1:B:57:ILE:HD12	0.94	1.62	17	1
1:A:29:ALA:HB2	1:B:99:TRP:CH2	0.94	1.98	14	7

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	67/107 (63%)	60±2 (89±3%)	6±2 (9±2%)	2±1 (3±2%)	9	45
1	B	67/107 (63%)	60±2 (89±3%)	5±2 (8±2%)	2±1 (3±2%)	8	42
All	All	2680/4280 (63%)	2384 (89%)	224 (8%)	72 (3%)	8	43

5 of 14 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	B	44	THR	12
1	A	44	THR	11
1	A	45	PRO	10
1	B	45	PRO	10
1	B	106	LYS	9

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	62/92 (67%)	46±3 (75±5%)	16±3 (25±5%)	2 24
1	B	62/92 (67%)	46±3 (75±5%)	16±3 (25±5%)	2 24
All	All	2480/3680 (67%)	1850 (75%)	630 (25%)	2 24

5 of 96 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	36	LEU	16
1	A	38	LEU	16
1	B	36	LEU	16
1	B	38	LEU	16
1	B	63	ARG	14

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