



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 06:01 PM BST

PDB ID : 1XS9
Title : A MODEL OF THE TERNARY COMPLEX FORMED BETWEEN MARA,
THE ALPHA-CTD OF RNA POLYMERASE AND DNA
Authors : Dangi, B.; Gronenborn, A.M.; Rosner, J.L.; Martin, R.G.
Deposited on : 2004-10-18

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
<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

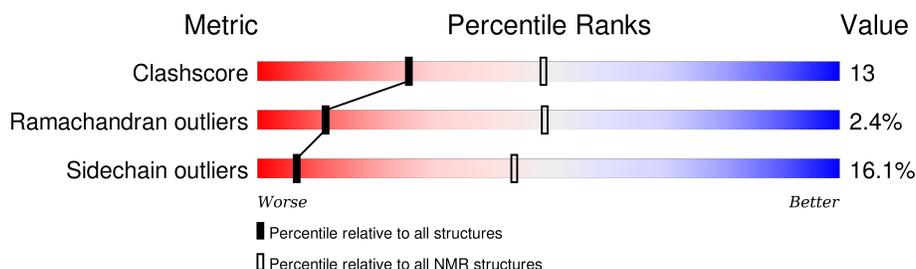
1 Overall quality at a glance i

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	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	B	20	
2	C	20	
3	A	132	
4	D	84	

2 Ensemble composition and analysis

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4720 atoms, of which 2181 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called 5'-D(P*GP*AP*TP*TP*TP*AP*GP*CP*AP*AP*AP*AP*CP*GP*TP*GP*GP*CP*AP*T)-3'.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
1	B	20	642	197	227	79	119	20	0

- Molecule 2 is a DNA chain called 5'-D(P*AP*TP*GP*CP*CP*AP*CP*GP*TP*TP*TP*TP*GP*CP*TP*AP*AP*AP*TP*C)-3'.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	C	20	636	195	229	69	123	20	0

- Molecule 3 is a protein called Multiple antibiotic resistance protein marA.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
3	A	129	2154	682	1069	195	202	6	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	CLONING ARTIFACT	UNP P0ACH5
A	-1	SER	-	CLONING ARTIFACT	UNP P0ACH5
A	0	HIS	-	CLONING ARTIFACT	UNP P0ACH5

- Molecule 4 is a protein called DNA-directed RNA polymerase alpha chain.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
4	D	81	1288	400	656	108	122	2	0

There are 3 discrepancies between the modelled and reference sequences:

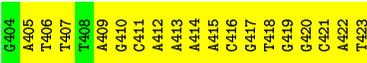
Chain	Residue	Modelled	Actual	Comment	Reference
D	246	GLY	-	CLONING ARTIFACT	UNP P0A7Z4
D	247	SER	-	CLONING ARTIFACT	UNP P0A7Z4
D	248	HIS	-	CLONING ARTIFACT	UNP P0A7Z4

4 Residue-property plots [i](#)

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.

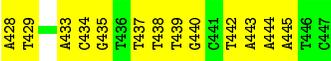
- Molecule 1: 5'-D(P*GP*AP*TP*TP*TP*AP*GP*CP*AP*AP*AP*AP*CP*GP*TP*GP*GP*CP*AP*T)-3'

Chain B: 


 A404
 A405
 T406
 T407
 T408
 A409
 G410
 C411
 A412
 A413
 A414
 A415
 C416
 G417
 T418
 G419
 G420
 C421
 A422
 T423

- Molecule 2: 5'-D(P*AP*TP*GP*CP*CP*AP*CP*GP*TP*TP*TP*TP*GP*CP*TP*AP*AP*AP*TP*C)-3'

Chain C: 


 A428
 T429
 A433
 C434
 G435
 T436
 T437
 T438
 T439
 G440
 C441
 T442
 A443
 A444
 A445
 T446
 C447

- Molecule 3: Multiple antibiotic resistance protein marA

Chain A: 


 GLY
 SER
 HIS
 M1
 T8
 T13
 W19
 I20
 E21
 L24
 E25
 S26
 P27
 L30
 E31
 K32
 R36
 S40
 Q45
 R46
 M47
 E51
 L56
 Y59
 I60
 T66
 E74
 S75
 L80
 R85
 K99
 D103
 V104
 P105
 R110
 Q115
 G116
 E117
 P123
 L124


 N125
 H126
 Y127
 N128
 S129

- Molecule 4: DNA-directed RNA polymerase alpha chain

Chain D: 


 GLY
 SER
 HIS
 F249
 I252
 R255
 P256
 V257
 L260
 E261
 L262
 T263
 V264
 R265
 S266
 C269
 L270
 K271
 A272
 E273
 A274
 L275
 H276
 Y277
 I278
 G279
 D280
 Q283
 R284
 T285
 E286
 N284
 T301
 R310
 S313
 L314
 G315
 M316
 R317
 R321
 P322
 P323
 T326
 A327
 D328
 E329

5 Refinement protocol and experimental data overview

The models were refined using the following method: *MODEL BASED ON CHEMICAL SHIFT PERTURBATION MAPPING*.

Of the 200 calculated structures, 1 were deposited, based on the following criterion: ?.

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

Software name	Classification	Version
HADDOCK	refinement	
NMRPIPE	structure solution	
XWINNMR	structure solution	
HADDOCK	structure solution	

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](#)

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	B	415	227	226	24
2	C	407	229	228	11
3	A	1085	1069	1065	18
4	D	632	656	654	18
All	All	2539	2181	2173	63

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 13.

5 of 63 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:B:414:DA:H2'	1:B:415:DA:O4'	0.71	1.86
1:B:419:DG:H2'	1:B:420:DG:O4'	0.70	1.86
1:B:415:DA:H2'	1:B:416:DC:O4'	0.70	1.85
2:C:437:DT:H2'	2:C:438:DT:O4'	0.65	1.91
2:C:442:DT:H2'	2:C:443:DA:O4'	0.61	1.95

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	
3	A	127/132 (96%)	112 (88%)	14 (11%)	1 (1%)	29	74
4	D	79/84 (94%)	69 (87%)	6 (8%)	4 (5%)	5	26
All	All	206/216 (95%)	181 (88%)	20 (10%)	5 (2%)	12	49

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

Mol	Chain	Res	Type
4	D	321	TRP
4	D	323	PRO
3	A	26	SER
4	D	266	SER
4	D	294	ASN

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	
3	A	121/123 (98%)	105 (87%)	16 (13%)	9	50
4	D	71/73 (97%)	56 (79%)	15 (21%)	4	33
All	All	192/196 (98%)	161 (84%)	31 (16%)	7	44

5 of 31 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
3	A	85	ARG
3	A	36	ARG
4	D	263	THR
4	D	277	TYR
3	A	56	LEU

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