

SUPPLEMENTARY MATERIAL FOR Predicting RNA Secondary Structures with Pseudoknots by MCMC Sampling

Dirk Metzler* Markus E. Nebel†

April 4, 2007

Contents

1	General Remarks	2
2	tmRNA	2
2.1	Sampled Structures for <i>Treponema Pallidum</i> pre-tmRNA	3
2.2	PPV and Sensitivity of Structure Estimations	4
3	tRNA from BRAlIBase	6
3.1	Posterior Probabilities of Base Pairs	6
3.2	PPV and Sensitivity of Structure Estimations	7
4	RNase P	10
4.1	Posterior Probabilities of Base Pairs	10
4.2	PPV and Sensitivity	11
5	Pseudoknot Fragments from PseudoBase	14
5.1	Posterior Probabilities of Base Pairs	14
5.2	PPV and Sensitivity	15
A	Tables	17
A.1	tmRNA	17
A.2	tRNA	22
A.3	RNase P	23
A.4	Pseudoknot Fragments	24

*Institut für Informatik, J. W. Goethe-Universität, Robert Mayer Str. 11-15, D-60325 Frankfurt am Main, Germany, metzler@cs.uni-frankfurt.de, www.cs.uni-frankfurt.de/~metzler

†Fachbereich Informatik, Technische Universität Kaiserslautern, Gottlieb-Daimler-Str., D-67663 Kaiserslautern, Germany, nebel@informatik.uni-kl.de

1 General Remarks

We use the term *structure* for “secondary structure, possibly including pseudoknots”. In other words, the structure is the set of all base pairs.

We compare the performance of the following three programs for predicting the structure of an RNA molecule from its nucleotide sequence:

McQFold, an implementation of our method, available from

<http://www.informatik.uni-frankfurt.de/~metzler/McQFold>,

RNAfold from the Vienna RNA package versions 1.4 and 1.6 [Hof03], an implementation of the classical algorithm of [ZS81] disregarding pseudoknots, and

pknotsRG version 1.2, an implementation of the method of cf. [RG04] for the prediction of secondary structures including a certain class of pseudoknots.

We run these programs with default parameter settings. The caveats discussed in [GG04] for the program performance comparisons accomplished there, do also apply to our studies.

We measure the accuracy of RNA structure predictions by the following ratios (cf. [BBC⁺00, DE04, Mat04]):

Sensitivity, the relative frequency of correctly predicted base pairs among all actual base pairs, and

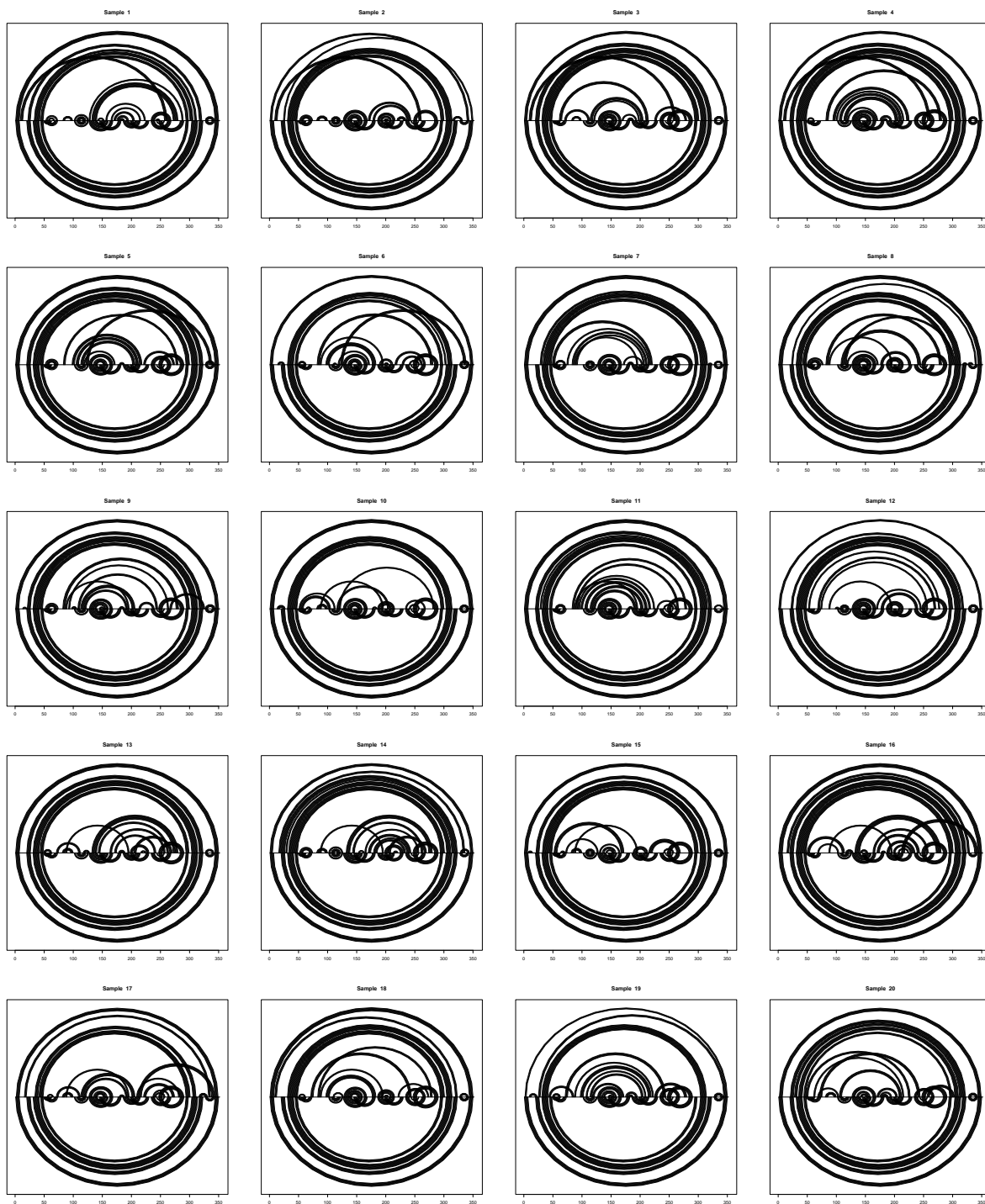
Positive predictive value (PPV), the relative frequency of correctly predicted base pairs among all predicted base pairs.

2 tmRNA

We applied the three programs to the 352 tmRNA sequences whose structures were available from the tmRNA Website <http://www.indiana.edu/~tmrna> [Wil00, WB98] in autumn 2005. Most of these tmRNA molecules have similar structures including several pseudoknots. The sequences are listed in Appendix A.1. One of them was only 99 nucleotides long. The lengths of the other sequences ranged from 211 to 435 positions, taking only the part for which the structure was given. We consider the exceptionally short sequence, for which each of the three programs RNAfold, pknotsRG, and McQFold predicted exactly the right structure, as atypical for tmRNA and therefore exclude it from the evaluations shown below.

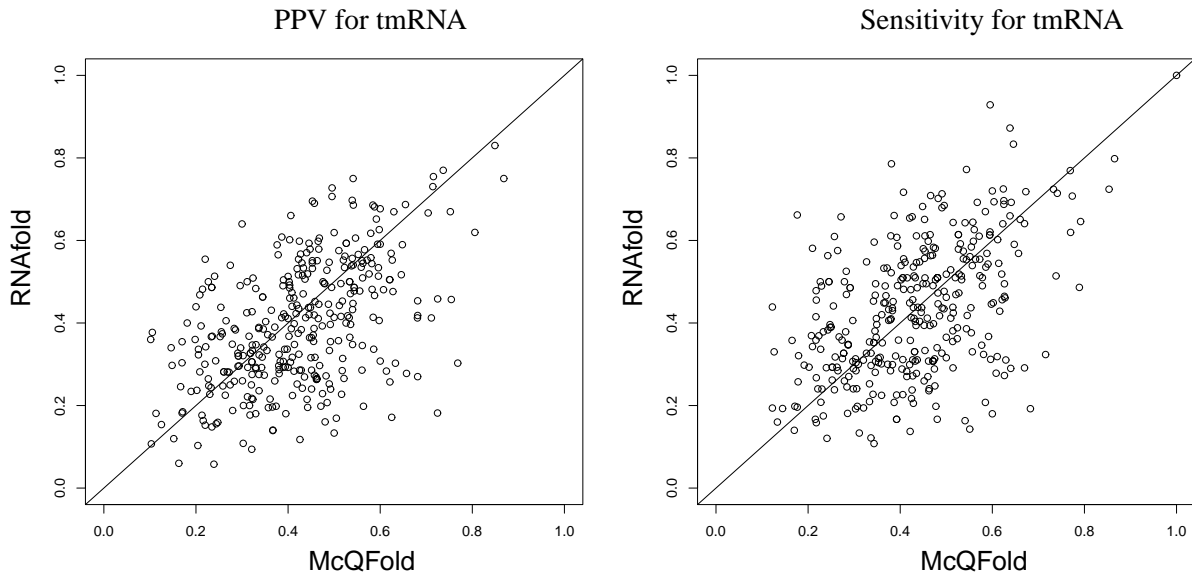
2.1 Sampled Structures for *Treponema Pallidum* pre-tmRNA

McQFold samples from an approximation of the posterior distribution of structures, given the input RNA-sequence. The upper halves of the following 20 panels show the samples for the pre-tmRNA sequence of *Treponema Pallidum*, representing each base pair by an arc connecting the corresponding positions. For comparison, the lower halves display the structure given on the tmRNA Website.

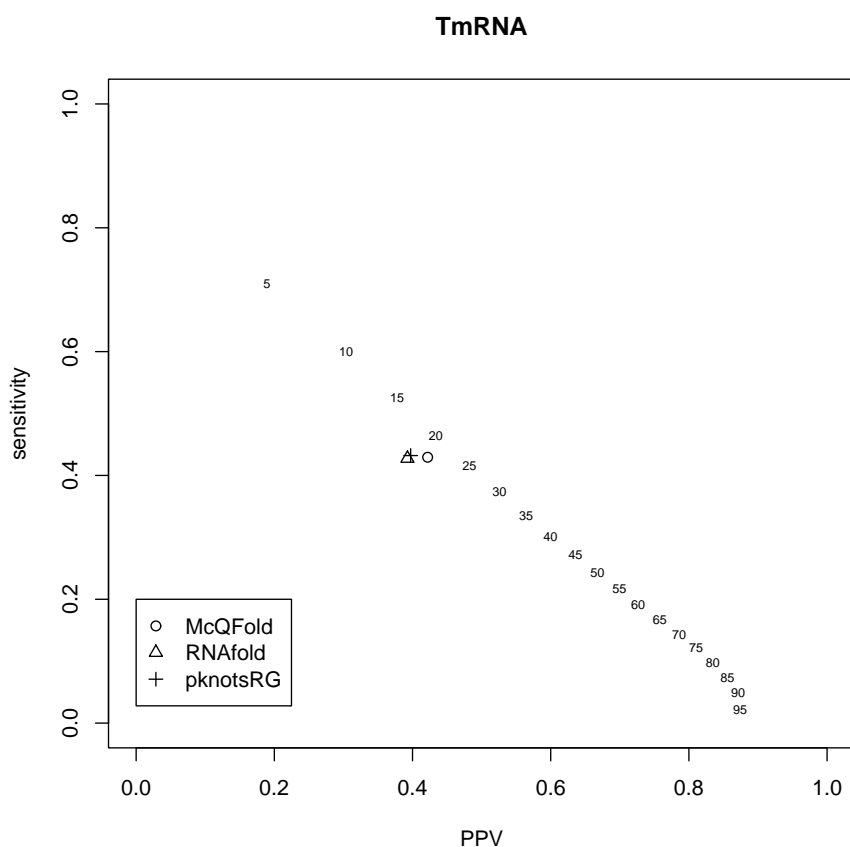


2.2 PPV and Sensitivity of Structure Estimations

Each dot in the following plots corresponds to a tmRNA sequence. Its coordinates are the PPV (left hand panel) and sensitivity (right hand panel) respectively of the structure found by the importance-sampling procedure in McQFold and by pknotsRG. For these data, RNAfold gave nearly the same results as pknotsRG (data not shown).



For the next figure we pooled the results (numbers of correctly predicted, falsely predicted, and undiscovered base pairs) for all tmRNA sequences and computed the overall PPV and sensitivity values. The open bullet refers to the importance-sampling estimation of McQFold. The numbers n in the panel refer to consensus rules which assume positions to be paired if and only if their posterior pairing probability, as computed by McQFold, exceed $n\%$. For example, one could assume that i and j are paired if and only if McQFold estimated their posterior pairing probability higher than 50%. The label 50 at $(x, y) = (0.67, 0.24)$ says that this leads to a PPV of 0.67 and a sensitivity of 0.24. (Note that using a threshold under 50% in this manner may lead to a set of contradictory base pairs in the sense that both $\{i, j\}$ and $\{i, k\}$ with $k \neq j$ can be included.)



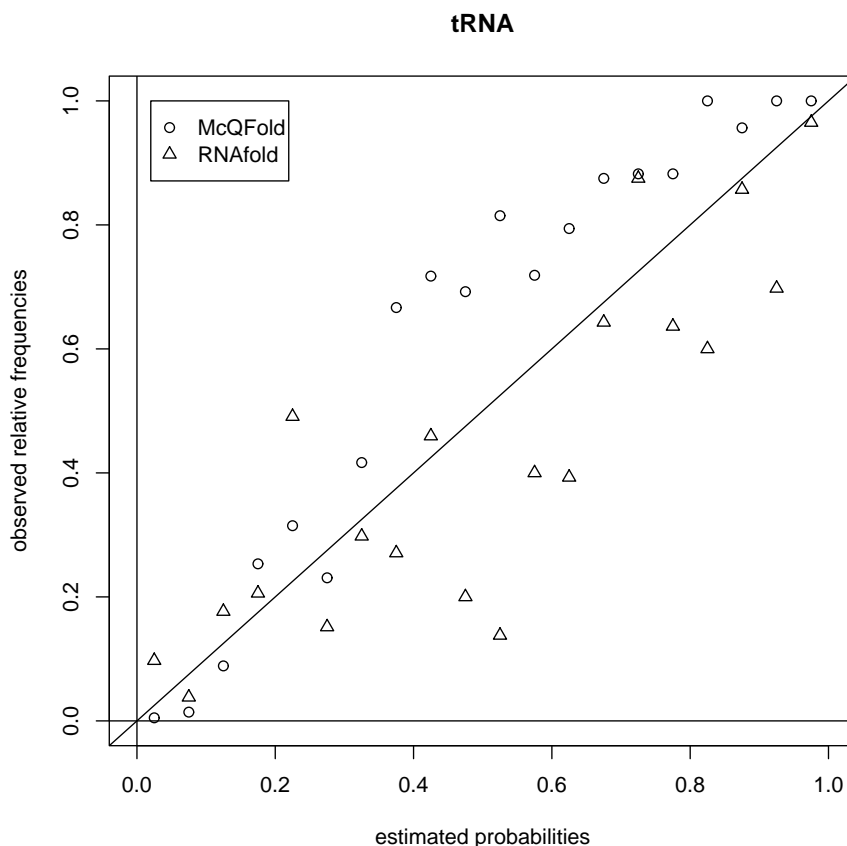
3 tRNA from BRAlibase

The structures of the 21 tRNA sequences given in the BRAlibase (cf. [GG04, GJBM⁺03, SMS⁺76], <http://www.binf.ku.dk/~pgardner/bralibase>) in March 2007 do not contain pseudo-knots. We skipped one of the sequences because it contains an undetermined nucleotide, which cannot be handled by the current version of McQFold. The 20 sequences we used are listed in Appendix A.2. The first 11 of them are highly similar to each other (more than 90% coincidence) and should therefore not be considered as a sample representative for tRNA molecules.

3.1 Posterior Probabilities of Base Pairs

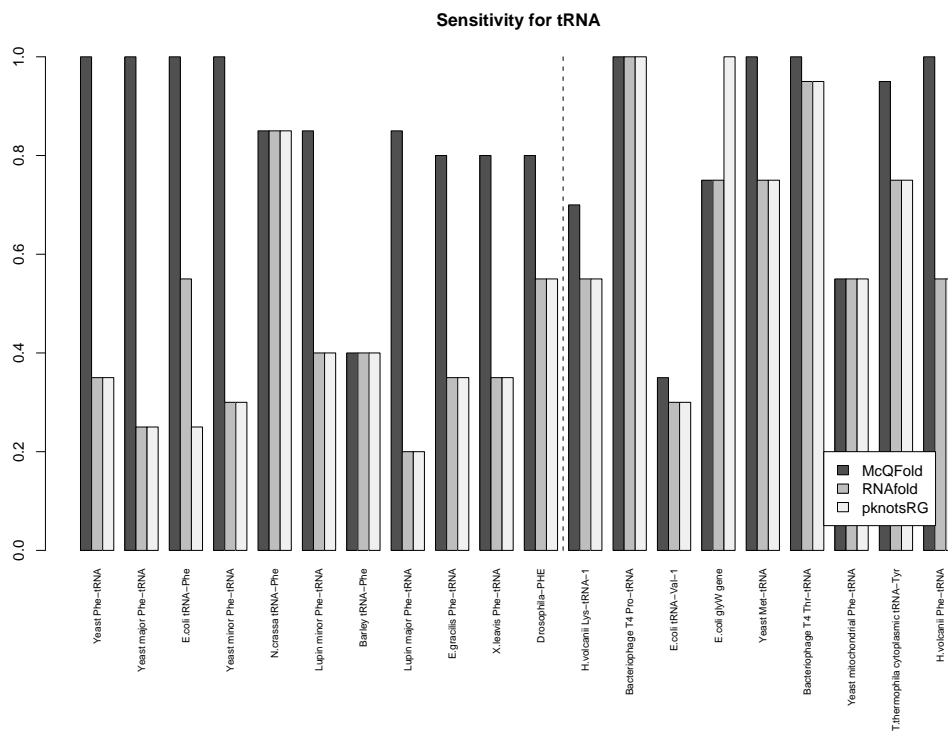
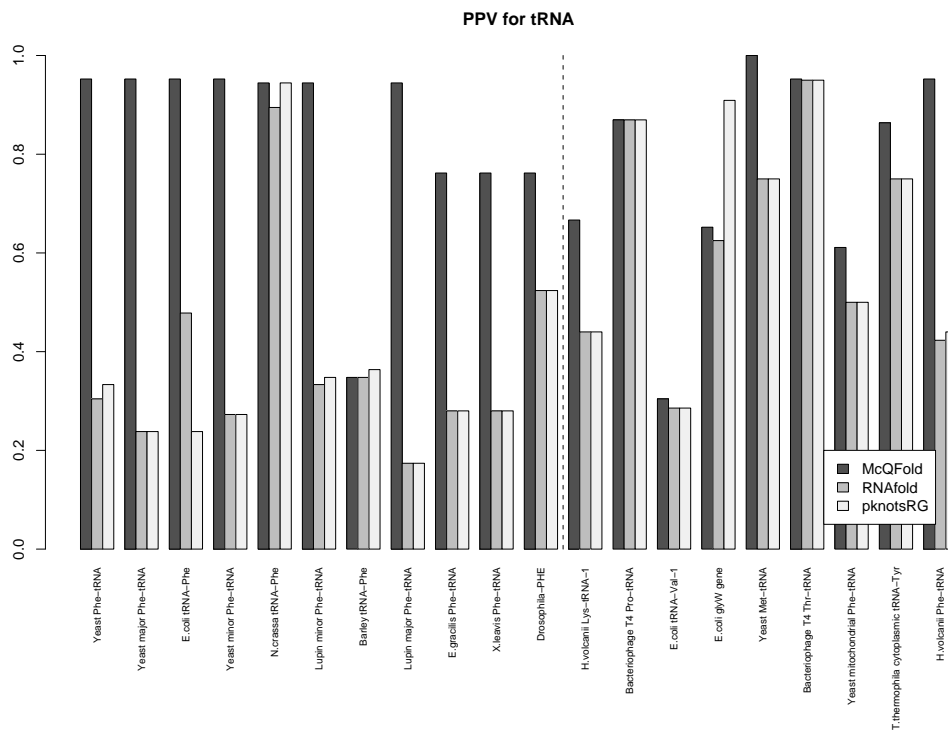
McQFold computes for each pair $\{i, j\}$ of sequence positions the posterior probability, given the entire RNA sequence, that i and j form a base pair in the actual structure of the molecule. RNAfold, when started with the option “-p”, does also compute analogous values. In general, the results of the two programs differ because they are based on different models. The analogon of the posterior structure in energy-based models is the Boltzmann-ensemble.

We use the following idea to inspect the tenability of the given values. If we gather many pairs $\{i, j\}$ of sequence positions which all have a similar posterior pairing probability of approximately p and then check how many of them are indeed paired, then their relative frequency should be close to p . For example, take from the 20 tRNA sequences all pairs $\{i, j\}$ to which McQFold assigned a posterior pairing-probability between 0.25 and 0.3. The open bullet at $(x, y) = (0.275, 0.23)$ in the next panel says that 23% of them are paired in the corresponding true structure (or at least the one given in the data base).

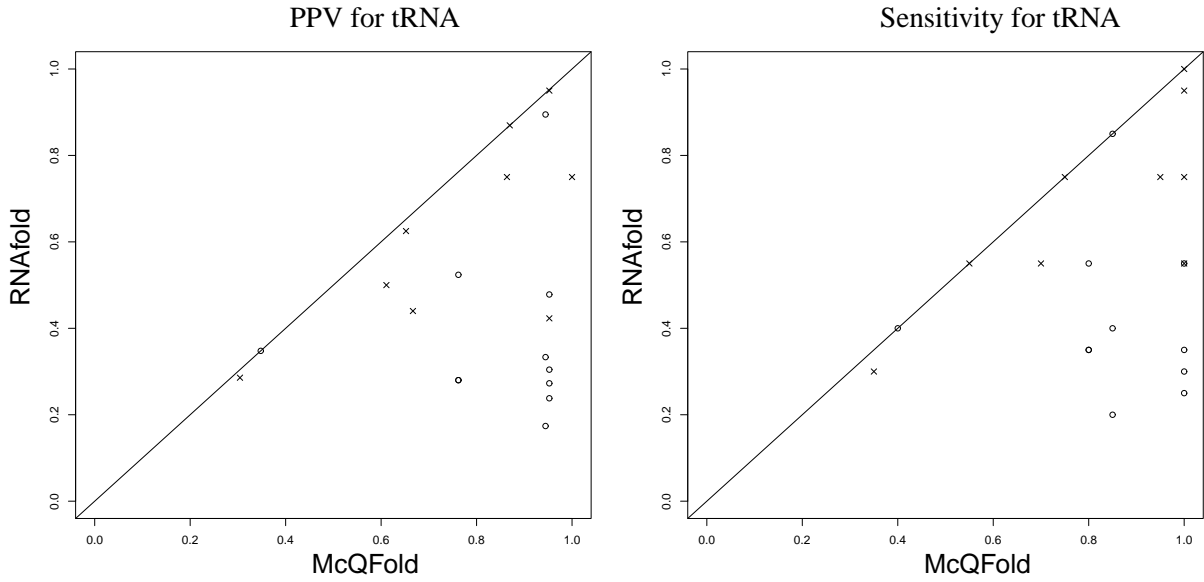


3.2 PPV and Sensitivity of Structure Estimations

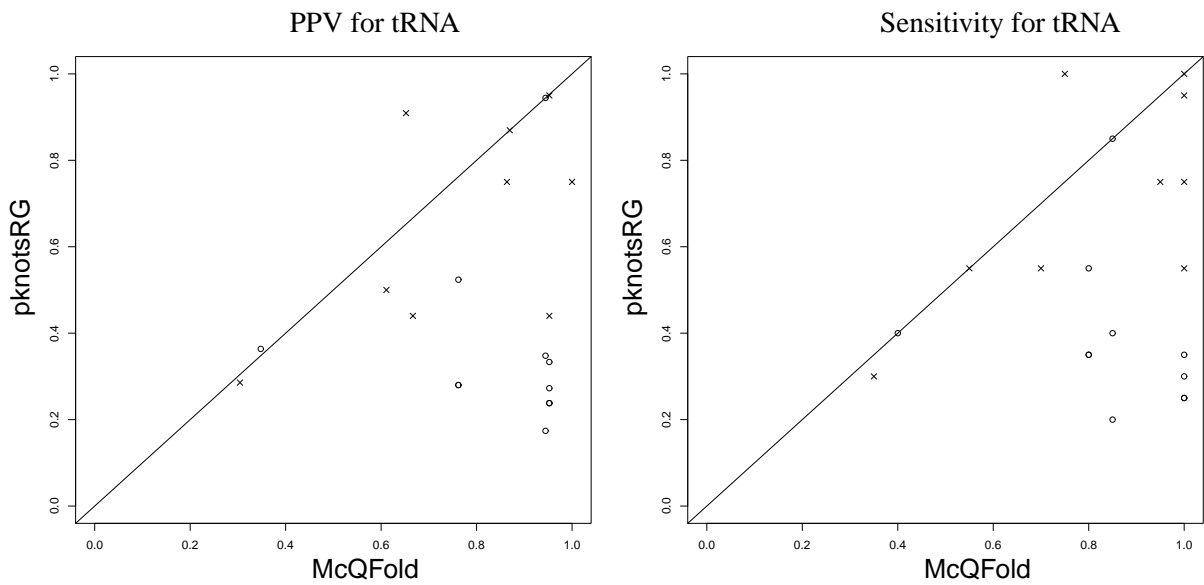
The following bar charts show the PPV and the sensitivity of the structure estimates of McQFold, RNAfold, and pknotsRG. The RNA molecules left of the dashed lines are those of high mutual sequence similarity.



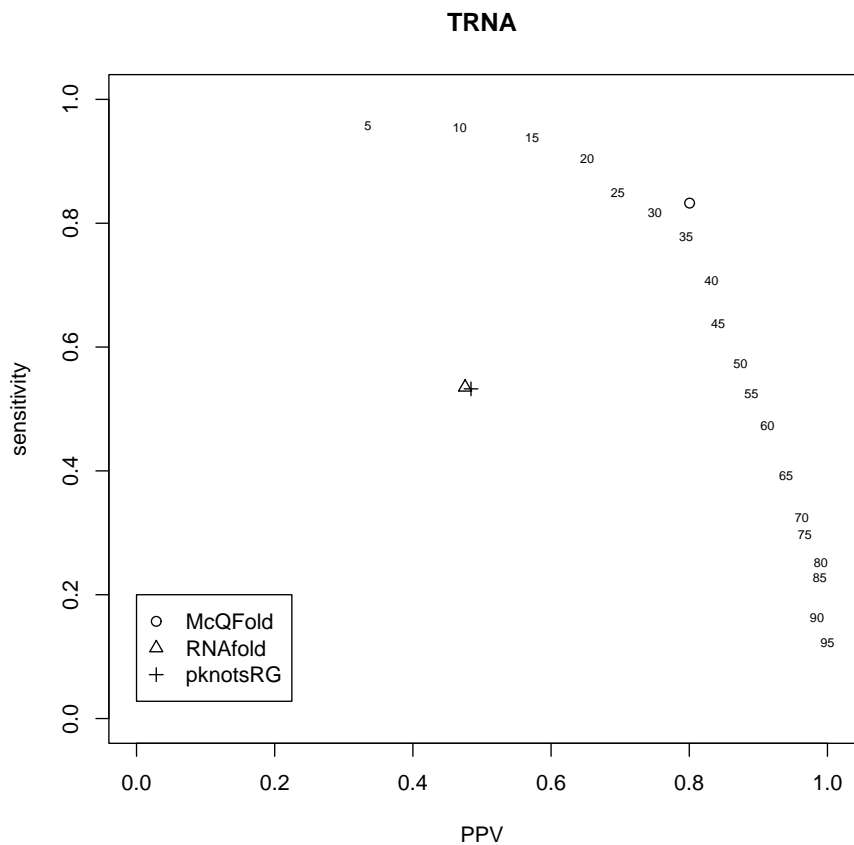
In the following plot each open bullet corresponds to one of the 11 highly similar tRNA sequences and the crosses refer to the other tRNA sequences. The coordinates are the PPV (left hand panel) and the sensitivity (right hand panel) of the structure found by the importance-sampling procedure in McQFold and by pknotsRG



The same with RNAfold instead of pknotsRG:



For the next figure we pooled the results (numbers of correctly predicted, falsely predicted, and undiscovered base pairs) for all tRNA sequences and computed the overall PPV and sensitivity values. The open bullet refers to the importance-sampling estimation of McQFold. Each number n in the panel refers to a consensus rule which assumes positions to be paired if and only if their posterior pairing probability, as computed by McQFold, exceed $n\%$.



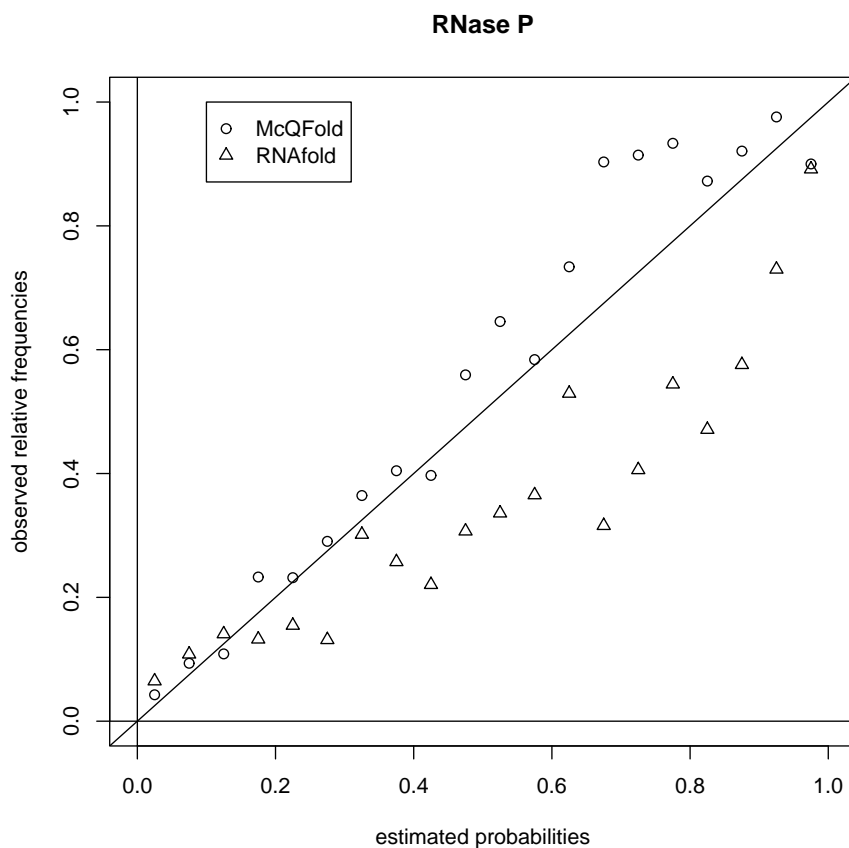
4 RNase P

The RNase P sequences and structures given in the “RNase P Database” (cf. [Bro99], <http://www.mbio.ncsu.edu/RNaseP>) are subdivided into 20 groups. For each of these groups we selected the sequence top-most in the corresponding table among all sequences completely available in March 2007, cf. Appendix A.3. We skipped the group “unknown bacterium” because it only contained incomplete sequences.

The typical RNase P structure includes pseudoknots. However, they can be eliminated by removing just two helices.

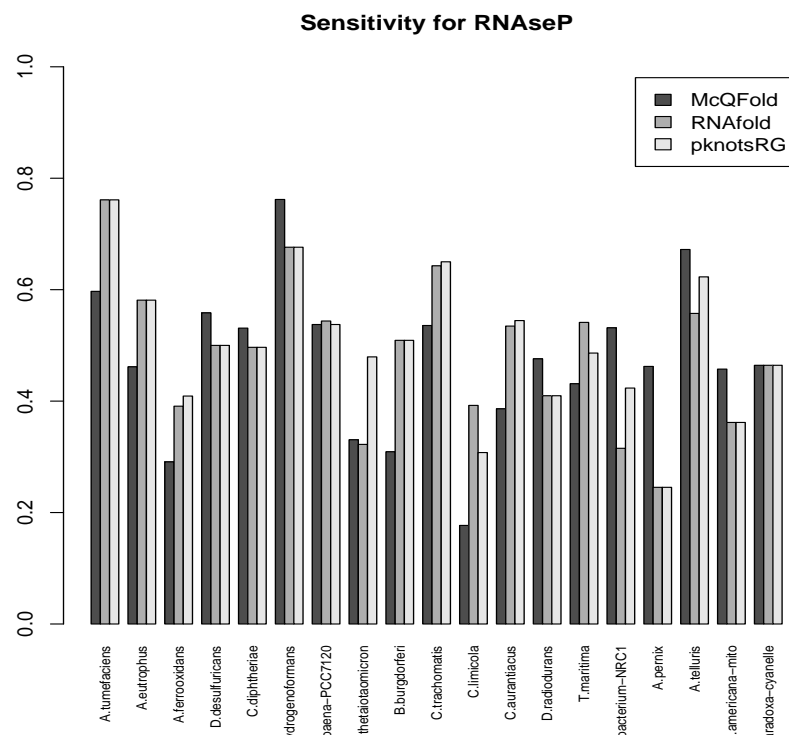
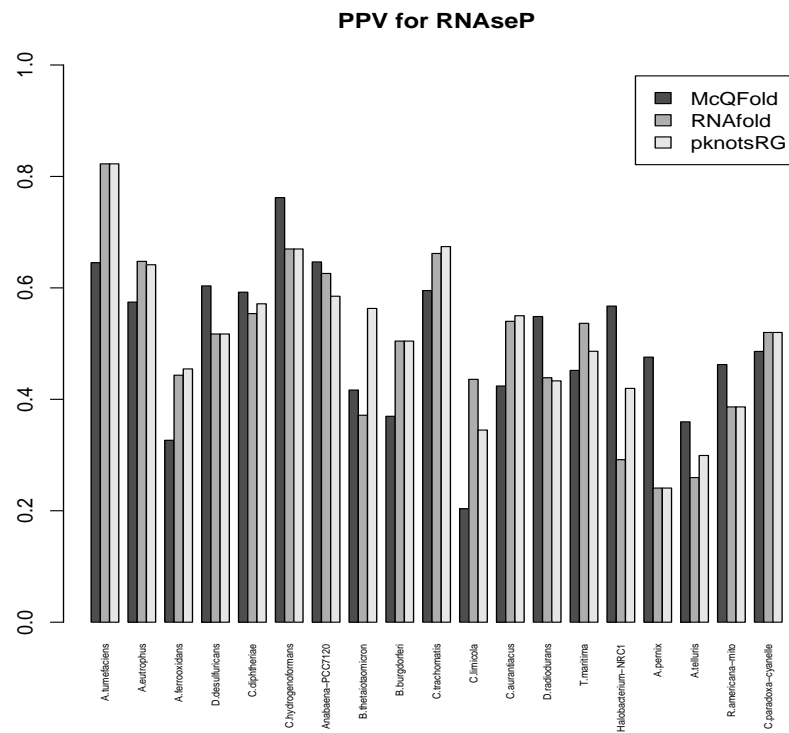
4.1 Posterior Probabilities of Base Pairs

With the next plot we inspect the tenability of base pair posterior probabilities given by McQFold and RNAfold, the latter called with option “-p”. Among all pairs of positions with posterior pairing probability of approximately p , the relative frequency of actual base pairs should be close to p . For example, the open bullet at $(x, y) = (0.275, 0.2905)$ means that 29.05% of all pairs of RNase P sequence positions for which McQFold gave a posterior pairing probability in $]0.25, 0.3]$ are indeed paired in the actual structure.

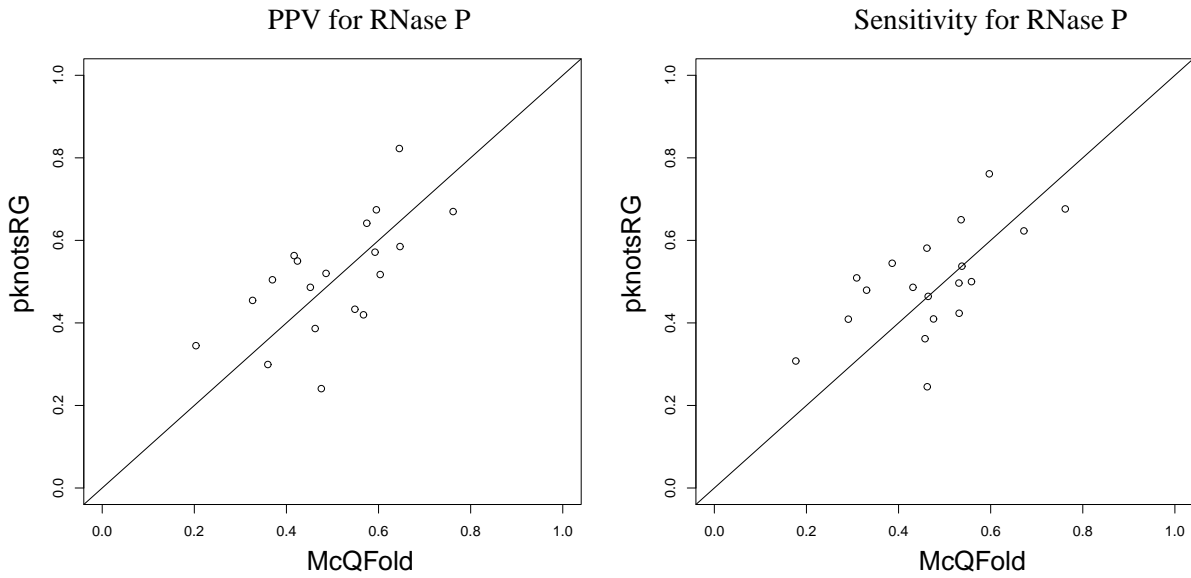


4.2 PPV and Sensitivity

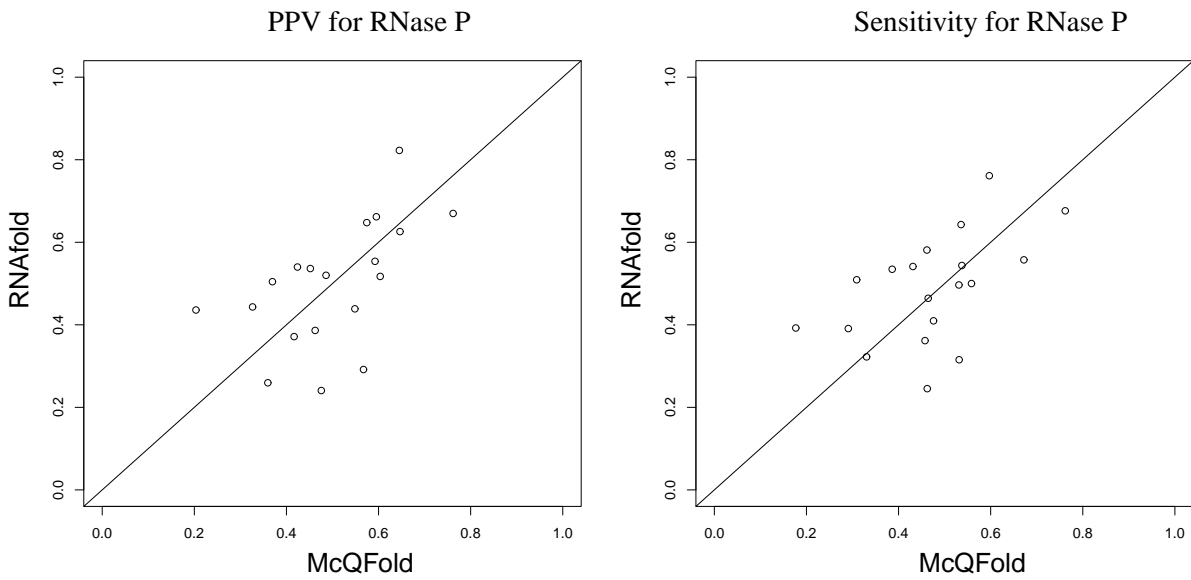
Each bar in the following charts refers to one RNase P sequence.



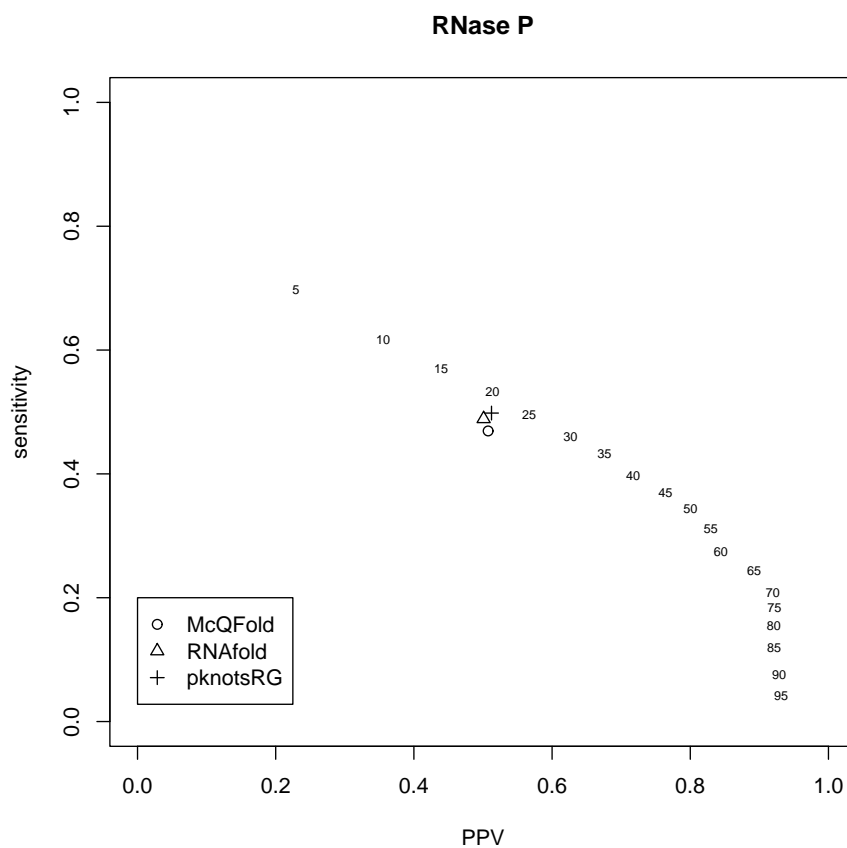
In the next plots we compare the results of McQFold's importance-sampling procedure and those of pknotsRG. Each dot refers to one RNase P sequence.



The same comparison with pknotsRG replaced by RNAfold:



For the next figure we pooled the results (numbers of correctly predicted, falsely predicted, and undiscovered base pairs) for all RNase P sequences and computed the overall PPV and sensitivity values. The open bullet refers to the importance-sampling estimation of McQFold. Each number n in the panel refers to a consensus rule which assumes positions to be paired if and only if their posterior pairing probability, as computed by McQFold, exceed $n\%$.



5 Pseudoknot Fragments from PseudoBase

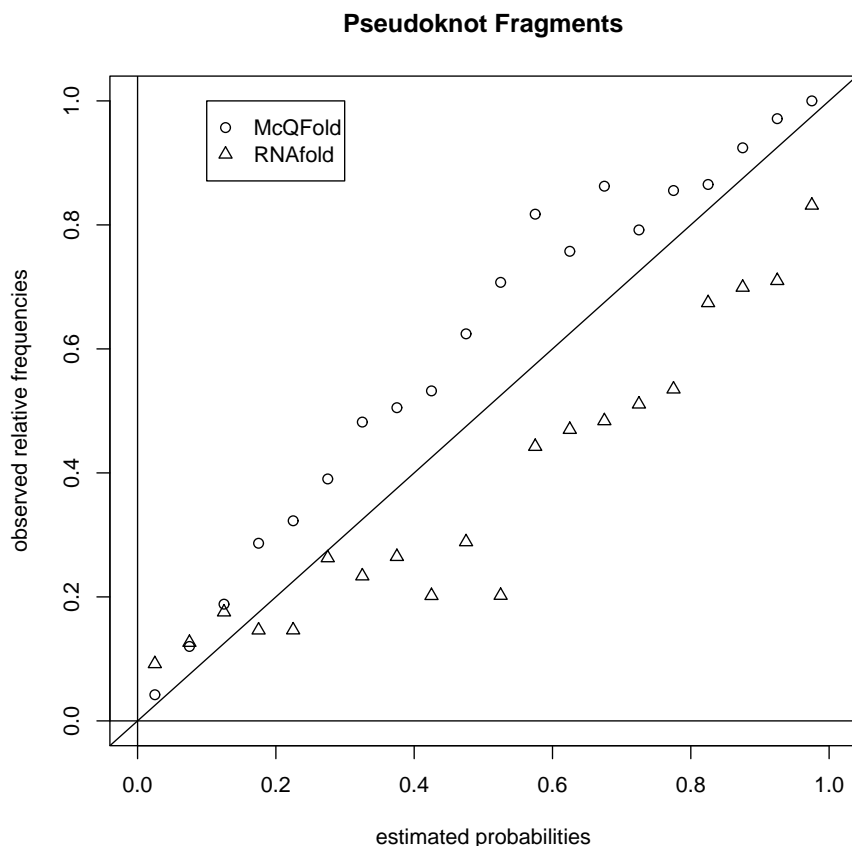
Now we turn to the sequences and structures for pseudoknot fragments given in the PseudoBase, cf. [vBGP⁺00], <http://biology.leidenuniv.nl/~batenburg/PKBGet.html>.

We excluded sequences if they were not contiguously given in the PseudoBase or if some of their nucleotides were undetermined. Since the single pseudoknot fragments do not provide much data, we pool the results according to the grouping used in the PseudoBase, see Appendix A.4.

Note that the PseudoBase does contain pseudoknot fragments only. Thus, the sequences are strongly biased for pseudoknots. By selecting contiguous fragments we generate another bias for small distances $|i - j|$ between paired positions i and j . Therefore, one should interpret the following results very cautiously.

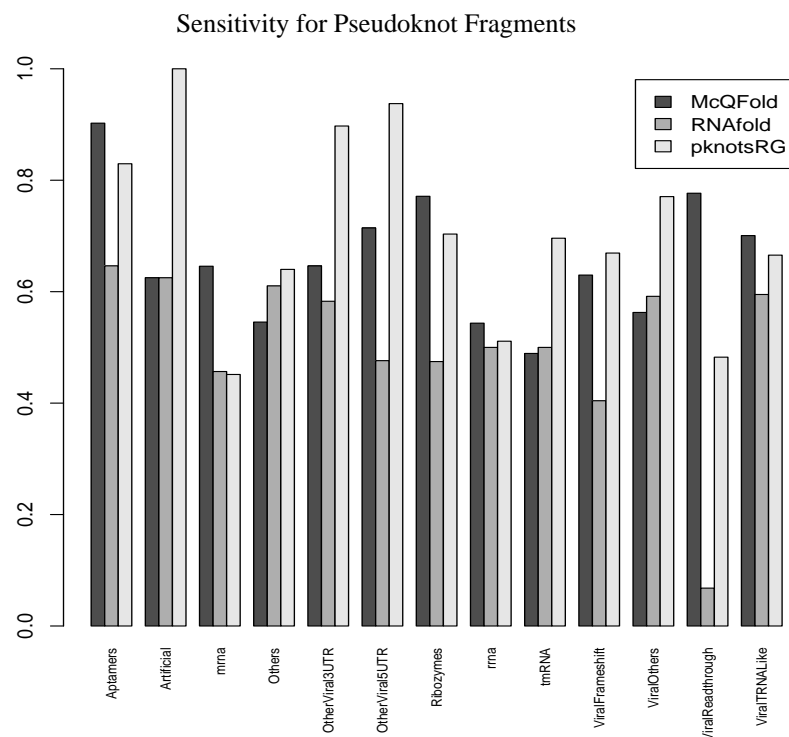
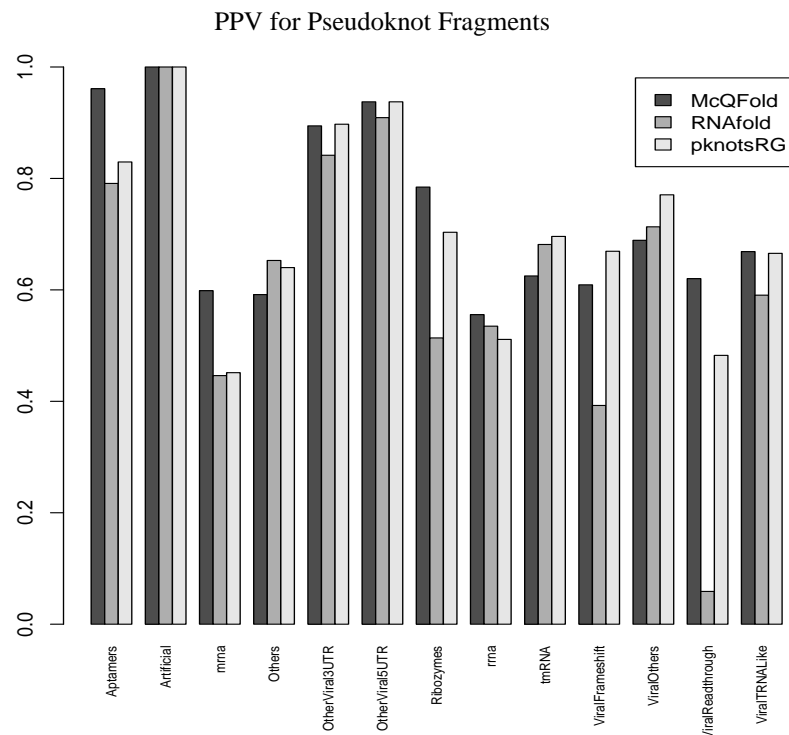
5.1 Posterior Probabilities of Base Pairs

With the next plot we inspect the tenability of posterior base-pairing probabilities given by McQFold and RNAfold, the latter called with option “-p”. More detailed explanations are given in the previous sections.

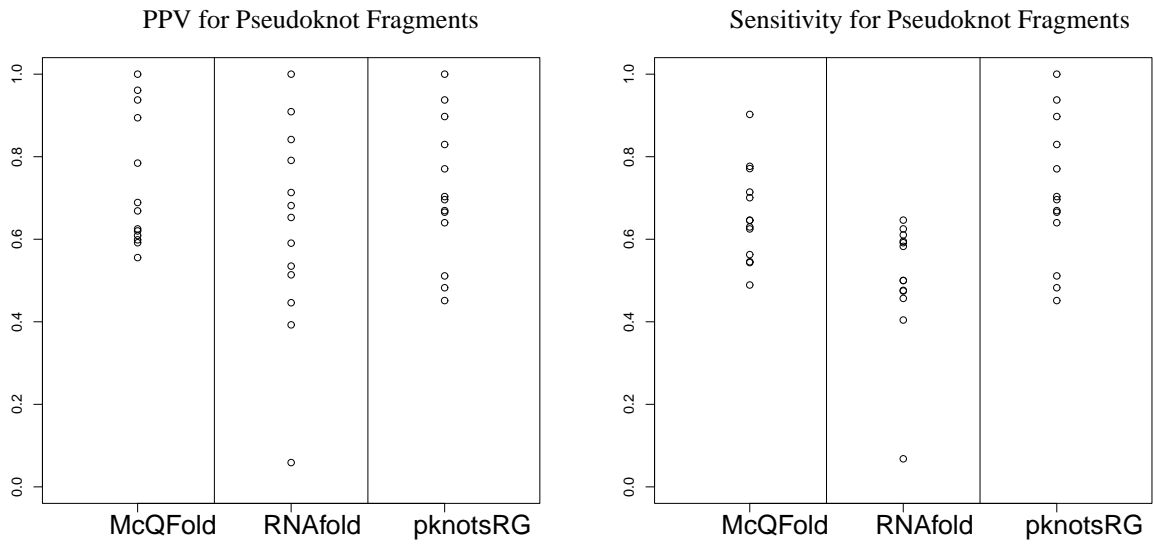


5.2 PPV and Sensitivity

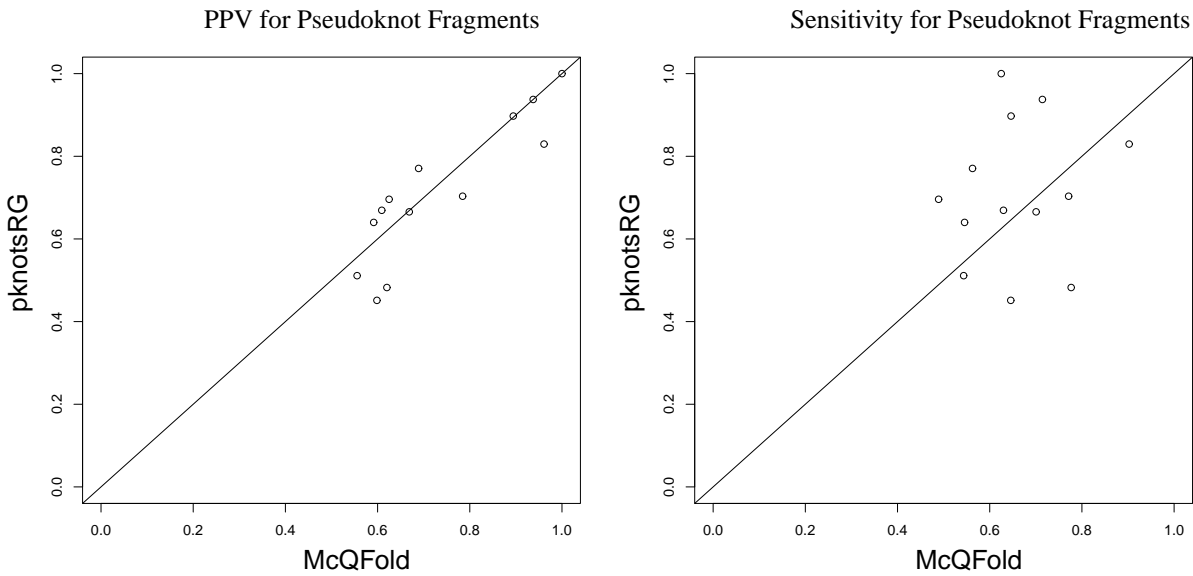
The bars in the following charts refer to groups of pseudoknot fragments.



The following stripcharts display the same results, sorted accordingly to the program used.



The next plots show a direct comparison between McQFold and pknotsRG. Each dot corresponds to one group of pseudoknot fragments.



A Tables

The following tables specify the used sequences with their lengths and give *percental* values for PPV and sensitivity. The McQFold columns refer to the McQFold's importance sampling procedure.

A.1 tmRNA

We applied McQFold, pknotsRG, and RNAfold to the following 352 tmRNA sequences whose structures were available from the tmRNA Website <http://www.indiana.edu/~tmrna> [Wil00, WB98] in autumn 2005. Most of these tmRNA molecules have a similiar structure with several pseudoknots. Sequence lengths given in the following table refer to the part of the sequence for which the structure was available from the tmRNA Website.

tmRNA	McQFold		RNAFold		pknotsRG		sequence length
	PPV	Sens.	PPV	Sens.	PPV	Sens.	
Acidithiobacillus ferrooxidans tmRNA, v1 ATCC 23270	49	48	42	45	50	52	364
Acidithiobacillus ferrooxidans tmRNA, v2	58	62	41	43	50	52	364
Xylella fastidiosa tmRNA, v1 9a5c	17	18	18	20	18	20	355
Xylella fastidiosa tmRNA, v2 Dixon	23	22	23	24	22	24	355
Xylella fastidiosa tmRNA, v3 Ann-1	32	28	37	38	37	38	355
Xylella fastidiosa tmRNA, v4 Temecula1	31	28	22	25	22	25	355
Xanthomonas campestris tmRNA	24	29	16	18	16	18	397
Xanthomonas axonopodis tmRNA	32	35	43	47	42	47	397
Xanthomonas oryzae tmRNA	30	34	11	12	16	19	397
Uncultured rice4 tmRNA	39	38	41	43	31	32	354
Stenotrophomonas maltophilia tmRNA	41	43	32	31	37	36	354
Dichelobacter nodosus tmRNA	26	27	37	37	38	37	352
Francisella tularensis tmRNA, v1 Schu 4	27	31	25	30	24	29	421
Francisella tularensis tmRNA, v2 LVS	32	37	25	30	24	29	421
Nitrosococcus oceani tmRNA	32	32	18	19	18	19	361
Uncultured whale1 tmRNA	24	23	31	34	31	34	415
Legionella pneumophila tmRNA	29	24	38	38	29	31	406
Shewanella SAR-2 pre-tmRNA, v2	54	41	75	72	78	75	353
Coxiella burnetii pre-tmRNA	53	54	50	55	50	55	357
Methylococcus capsulatus tmRNA	34	34	28	30	26	28	356
Shewanella SAR-1 tmRNA, v2	63	58	48	49	48	49	348
Uncultured marineEBAC20E09 tmRNA	62	58	50	51	50	50	347
Azotobacter vinelandii tmRNA	43	42	38	41	38	41	352
Pseudomonas aeruginosa tmRNA	60	60	41	45	41	45	353
Pseudomonas syringae tmRNA, v1 DC3000	36	35	37	40	38	41	388
Pseudomonas syringae tmRNA, v2 B728a	48	45	50	51	50	51	392
Pseudomonas syringae tmRNA, v3 1448A	54	49	69	68	70	69	388
Pseudomonas fluorescens tmRNA, v1 PFO-1	54	50	45	48	44	47	390
Pseudomonas fluorescens tmRNA, v2 SBW25	45	47	53	56	62	66	397
Pseudomonas fluorescens tmRNA, v3 Pf-5	54	50	70	68	69	68	394
Pseudomonas putida tmRNA, v1 KT2440	53	50	50	52	51	53	399
Moraxella catarrhalis tmRNA	56	53	51	58	51	58	359
Acinetobacter ADP1 tmRNA	37	38	39	41	39	41	360
Uncultured c. remanei1 tmRNA	37	38	26	28	26	28	359
Uncultured bone3 tmRNA	32	31	20	20	20	20	359
Psychrobacter 273-4 tmRNA	39	39	29	32	29	32	359
Chromohalobacter salexigens pre-tmRNA	32	33	51	55	50	55	380
Saccharophagus degradans tmRNA	33	38	35	44	35	44	389
Gammaproteobacteria SAR-1 tmRNA	21	22	47	48	47	48	345
Colwellia psychrerythraea tmRNA	48	41	27	28	28	29	350
Shewanella oneidensis tmRNA	11	12	38	44	30	32	355
Shewanella putrefaciens tmRNA	22	22	34	37	33	36	356
Shewanella PV-4 tmRNA	43	42	12	14	12	14	356
Shewanella denitrificans tmRNA	22	22	16	17	16	17	355
Shewanella SAR-1 tmRNA, v1	45	46	36	39	37	39	355
Shewanella SAR-2 tmRNA, v1	11	12	18	19	18	19	355
Idiomarina loihiensis tmRNA	43	40	57	54	58	56	359
Photobacterium profundum tmRNA	32	29	35	35	35	35	368
Vibrio fischeri tmRNA	48	48	16	17	32	34	367
Vibrio cholerae tmRNA	44	42	21	22	41	41	367
Vibrio parahaemolyticus tmRNA	39	36	29	30	30	30	368
Vibrio vulnificus tmRNA	36	33	22	21	36	33	367
Aeromonas hydrophila tmRNA, v1 PPD134/91	62	60	17	18	34	35	360
Aeromonas hydrophila tmRNA, v2 ATCC 7966	59	59	34	37	34	37	360
Wigglesworthia glossinidia pre-tmRNA	54	45	56	56	57	57	362
Blochmannia floridanus pre-tmRNA	45	43	44	44	45	44	390
Buchnera aphidicola pre-tmRNA, v1 APS	74	62	77	69	80	73	366

tmRNA	McQFold		RNAFold		pknotsRG		sequence length
	PPV	Sens.	PPV	Sens.	PPV	Sens.	
Buchnera aphidicola pre-tmRNA, v2 Sg	58	57	69	69	71	74	373
Buchnera aphidicola pre-tmRNA, v3 Bp	62	47	50	49	51	49	364
Photorhabdus asymbiotica tmRNA	15	13	34	33	34	33	364
Photorhabdus luminescens tmRNA	53	45	42	43	47	51	364
Shigella dysenteriae tmRNA	56	48	48	51	48	51	363
Shigella flexneri tmRNA	49	43	49	51	49	51	363
Shigella sonnei tmRNA	54	48	49	51	49	51	363
Escherichia coli tmRNA, v1 K12, DH10B, K1	54	48	49	51	49	51	363
Escherichia coli tmRNA, v3 O157:H7, 042	54	48	48	51	48	51	363
Enterobacter sakazakii tmRNA	38	34	32	33	33	34	364
Serratia marcescens tmRNA, v2 Db11	26	25	38	39	38	39	364
Pantoea stewartii tmRNA	35	30	27	28	27	28	363
Proteus mirabilis tmRNA	42	43	44	46	44	45	367
Klebsiella pneumoniae tmRNA	58	52	48	51	48	51	363
Erwinia carotovora tmRNA	52	48	23	25	23	25	363
Erwinia amylovora tmRNA	28	30	39	41	39	41	381
Erwinia chrysanthemi tmRNA	46	45	69	65	41	41	381
Salmonella bongori tmRNA	33	30	29	30	29	30	363
Salmonella enterica tmRNA, v1 Typhi, Typhimurium	25	25	37	40	38	41	363
Salmonella enterica tmRNA, v2 Paratyphi	19	18	23	26	29	31	363
Salmonella enterica tmRNA, v3 Enteritidis, Dublin, Pullorum, Gallinarum	38	35	29	31	29	31	363
Salmonella enterica tmRNA, v5 diarizonae	29	30	30	31	30	31	363
Salmonella enterica tmRNA, v6 Choleraesuis	31	30	29	31	29	31	363
Citrobacter rodentium tmRNA	39	35	49	50	46	48	365
Yersinia pestis tmRNA	32	28	33	33	34	35	364
Yersinia pseudotuberculosis tmRNA	32	28	33	33	34	35	364
Yersinia enterocolitica tmRNA	29	33	33	35	33	35	381
Mannheimia haemolytica tmRNA	51	51	42	41	42	41	367
Mannheimia succiniciproducens tmRNA	25	23	16	17	16	17	367
Haemophilus ducreyi tmRNA, v1 35000HP	29	26	30	28	29	28	367
Haemophilus influenzae tmRNA, v1 Rd, 10810	23	22	26	27	24	24	366
Haemophilus influenzae tmRNA, v2 R2866	22	22	15	16	15	16	366
Haemophilus influenzae tmRNA, v3 R2846	23	22	50	46	23	23	366
Haemophilus influenzae tmRNA, v4 86-028NP	34	33	32	33	32	33	366
Haemophilus somnus tmRNA	26	23	25	24	25	24	367
Pasteurella multocida tmRNA	34	30	24	24	24	24	367
Actinobacillus actinomycetemcomitans tmRNA	33	32	29	31	11	12	366
Actinobacillus pleuropneumoniae tmRNA, v1 sv.1, sv.7	38	35	31	30	35	35	366
Actinobacillus pleuropneumoniae tmRNA, v2 sv.5	41	36	35	32	36	37	366
Jakoba libera mitochondrion pre-tmRNA homolog	100	100	100	100	100	100	** 99 **
Tremblaya princeps pre-tmRNA, v1 Dysmicoccus	35	64	34	69	34	69	261
Neisseria gonorrhoeae tmRNA, v2 FA 1090	17	18	30	32	30	32	363
Neisseria lactamica tmRNA	34	35	29	31	30	32	362
Neisseria meningitidis tmRNA, v1 Z2491	32	35	28	28	40	39	363
Neisseria meningitidis tmRNA, v2 MC58	39	40	34	36	35	37	363
Neisseria meningitidis tmRNA, v3 FAM18	21	22	34	36	35	37	363
Chromobacterium violaceum tmRNA, v2 ATCC 12472	39	38	42	41	41	41	365
Nitrosomonas europaea tmRNA	45	42	36	36	63	58	360
Nitrospirillum multiformis tmRNA	41	39	66	66	66	66	362
Methylobacillus flagellatus tmRNA	38	38	30	30	30	30	361
Uncultured farm7 tmRNA	53	52	40	42	40	42	364
Azoarcus BH72 tmRNA	28	29	26	27	25	26	378
Thiobacillus denitrificans tmRNA	28	27	35	36	35	36	367
Uncultured stronglyoides1 tmRNA	40	43	49	54	49	54	378
Uncultured rice2 tmRNA	23	25	37	39	37	39	379
Uncultured aedes1 tmRNA	34	35	39	41	59	59	379
Polaromonas JS666 tmRNA	58	56	54	54	39	39	387
Rubrivivax gelatinosus tmRNA	42	44	53	55	54	57	372
Ralstonia solanacearum tmRNA	21	22	39	42	39	40	360
Ralstonia eutropha tmRNA	39	38	42	43	43	43	360
Ralstonia metallidurans tmRNA	46	45	39	39	39	39	360
Burkholderia thailandensis tmRNA	39	36	50	53	50	51	370
Burkholderia pseudomallei tmRNA	45	44	58	61	57	59	370
Burkholderia mallei tmRNA	47	45	58	61	57	59	370
Burkholderia fungorum tmRNA	32	31	21	24	21	24	370
Burkholderia cenocepacia tmRNA	40	41	51	56	54	56	370
Burkholderia cepacia tmRNA, v1 R1808	40	41	46	48	46	48	370
Burkholderia cepacia tmRNA, v2 R18194	45	45	32	33	31	33	370
Bordetella pertussis tmRNA	53	53	59	61	60	61	387
Bordetella parapertussis tmRNA	53	53	59	61	60	61	387
Bordetella bronchiseptica tmRNA	45	46	59	61	60	61	387
Bordetella avium tmRNA	36	34	43	45	43	45	406
Magnetococcus sp. tmRNA	18	27	40	66	39	66	366
Proteobacteria SAR-1 tmRNA, v1	49	45	55	58	55	58	355
Proteobacteria SAR-1 tmRNA, v1	40	38	50	52	48	50	345
Proteobacteria SAR-1 pre-tmRNA, v1	51	44	44	49	37	37	348
Lawsonia intracellularis tmRNA	63	61	30	32	52	58	376
Desulfovibrio desulfuricans tmRNA, v2 G20	44	45	41	44	65	64	360

tmRNA	McQFold		RNAFold		pknotsRG		sequence length
	PPV	Sens.	PPV	Sens.	PPV	Sens.	
Desulfovibrio vulgaris tmRNA	46	47	27	31	27	31	352
Desulfotalea psychrophila tmRNA	51	51	36	38	36	39	366
Uncultured ciona tmRNA	32	36	33	35	33	35	408
Geobacter sulfurreducens tmRNA	40	43	37	39	29	30	356
Geobacter metallireducens tmRNA	62	61	26	28	26	28	356
Desulfuromonas acetoxidans tmRNA	23	24	25	29	25	29	353
Bacteriovorax marinus tmRNA	24	55	6	14	6	14	385
Bdellovibrio bacteriovorus tmRNA	27	27	28	31	28	31	346
Myxococcus xanthus pre-tmRNA	32	34	9	11	21	25	364
Wolinella succinogenes tmRNA	60	61	53	54	59	60	369
Campylobacter lari tmRNA	54	49	43	44	43	44	359
Campylobacter upsaliensis tmRNA	30	28	49	50	30	34	359
Campylobacter fetus tmRNA, v2 23D	27	25	41	44	42	45	361
Campylobacter coli tmRNA, v1 RM2228	27	26	54	58	55	59	359
Campylobacter jejuni tmRNA	24	24	51	56	52	58	359
Helicobacter pylori tmRNA, v2 26695	41	45	47	50	48	50	386
Helicobacter pylori tmRNA, v3 J99	42	45	52	57	55	62	386
Helicobacter mustelae tmRNA	42	39	55	59	55	59	366
Helicobacter hepaticus tmRNA	51	48	58	58	59	60	369
Clostridium acetobutylicum pre-tmRNA	72	68	18	19	51	49	357
Clostridium perfringens tmRNA, v1 ATCC 13124	44	43	27	27	26	26	361
Clostridium perfringens tmRNA, v2 13	46	43	32	31	32	31	361
Clostridium thermocellum tmRNA	49	54	35	39	39	44	387
Clostridium botulinum pre-tmRNA (with Group I intron)	77	72	30	32	30	32	355
Clostridium tetani pre-tmRNA (with Group I intron)	70	67	67	64	64	63	355
Clostridium difficile tmRNA	48	46	50	50	39	37	350
Ruminococcus albus tmRNA	45	47	29	29	28	29	354
Carboxydotherrhus hydrogenoformans tmRNA	59	62	65	64	62	61	359
Thermoanaerobacter tengcongensis tmRNA	59	60	68	69	52	52	351
Coprothermobacter proteolyticus tmRNA	75	77	67	71	67	70	353
Moorella thermoacetica tmRNA	55	54	57	56	53	54	358
Mesoplasma florum tmRNA	52	59	45	54	47	56	408
Spiroplasma kunkelii tmRNA	44	45	57	68	57	68	424
Mycoplasma capricolum tmRNA	49	57	45	51	45	51	411
Mycoplasma mycoides tmRNA	47	50	30	37	42	52	411
Ureaplasma parvum tmRNA	44	46	20	24	10	13	413
Mycoplasma pulmonis tmRNA	31	37	42	49	54	65	387
Mycoplasma mobile tmRNA	56	56	54	60	54	60	376
Mycoplasma penetrans tmRNA	20	26	24	32	24	32	405
Mycoplasma arthritis tmRNA	50	54	13	16	9	10	394
Mycoplasma genitalium tmRNA, v1 G-37	42	42	43	48	44	48	388
Mycoplasma pneumoniae tmRNA	41	52	40	47	40	47	387
Mycoplasma hyopneumoniae tmRNA	15	38	30	79	30	79	424
Mycoplasma gallisepticum tmRNA	34	37	39	45	39	45	408
Phytoplasma asteris tmRNA	35	36	40	45	40	45	423
Paenibacillus larvae pre-tmRNA	41	38	46	45	47	46	364
Bacillus anthracis tmRNA	21	20	28	29	28	29	355
Bacillus thuringiensis tmRNA, v2 97-27	26	25	28	29	28	29	355
Bacillus cereus tmRNA, v1 ATCC 14579, G9241	27	26	28	29	28	29	355
Bacillus cereus tmRNA, v2 ATCC 10987	49	48	33	34	34	34	355
Bacillus cereus tmRNA, v3 ZK	26	25	28	29	28	29	355
Bacillus halodurans pre-tmRNA	42	44	38	40	37	39	358
Bacillus clausii pre-tmRNA	45	47	70	71	71	70	352
Bacillus subtilis pre-tmRNA	52	53	55	59	57	59	360
Bacillus licheniformis pre-tmRNA	55	54	61	64	62	65	353
Geobacillus stearothermophilus pre-tmRNA	63	66	55	57	55	57	349
Geobacillus kaustophilus pre-tmRNA	57	59	55	57	55	57	349
Exiguobacterium sp. pre-tmRNA	49	46	25	27	9	9	356
Oceanobacillus ihyensensis tmRNA	50	49	71	71	70	68	364
Staphylococcus aureus pre-tmRNA, v1 COL, NCTC 8325, NCTC6571	56	48	55	50	55	50	359
Staphylococcus aureus pre-tmRNA, v2 MRSA252	45	40	55	50	55	50	359
Staphylococcus aureus pre-tmRNA, v3 MSSA476	43	39	52	49	52	49	359
Staphylococcus aureus pre-tmRNA, v4 Mu50	46	42	55	51	55	51	359
Staphylococcus aureus pre-tmRNA, v5 MW2	43	41	55	51	55	51	359
Staphylococcus aureus pre-tmRNA, v6 N315	45	41	55	51	55	51	359
Staphylococcus epidermidis pre-tmRNA, v1 ATCC 14990	43	42	47	45	52	51	358
Staphylococcus epidermidis pre-tmRNA, v2 SR1, RP62A	54	53	54	52	55	56	359
Staphylococcus epidermidis pre-tmRNA, v3 (ends unsequenced) DSM20044, ATCC 12228	47	47	61	58	61	63	358
Listeria monocytogenes pre-tmRNA, v1 ATCC 15313, 4b F2365, 4b H7858, 1/2a F6854	21	23	48	50	48	50	367
Listeria monocytogenes pre-tmRNA, v2 EGD-e	23	25	49	50	49	50	367
L. innocua pre-tmRNA, v2 Clip11262, food 1, food 2, food 3, NCTC 11288, NCTC 12210	54	55	54	56	49	50	367
Listeria innocua pre-tmRNA, v3 (ends unsequenced) ATCC 12210	23	25	49	50	49	50	367
Lactobacillus gasserii pre-tmRNA	60	59	51	54	50	53	363
Lactobacillus johnsonii pre-tmRNA	58	58	62	67	61	66	364
Lactobacillus sakei pre-tmRNA	40	43	60	66	55	58	367
Lactobacillus delbrueckii pre-tmRNA, v2 ATCC BAA-365	52	55	56	58	54	56	359
Lactobacillus casei pre-tmRNA	59	62	47	52	45	49	361
Lactobacillus acidophilus pre-tmRNA	47	53	45	50	44	50	366

tmRNA	McQFold		RNAFold		pknotsRG		sequence length
	PPV	Sens.	PPV	Sens.	PPV	Sens.	
Lactobacillus plantarum pre-tmRNA, v2 WCFS1	20	21	36	34	28	29	369
Lactobacillus brevis pre-tmRNA	51	51	50	53	54	55	369
Pediococcus pentosaceus pre-tmRNA	62	61	57	60	55	58	366
Leuconostoc mesenteroides pre-tmRNA	60	56	68	63	69	66	355
Oenococcus oeni pre-tmRNA	39	34	61	60	61	60	347
Carnobacterium maltaromaticum pseudo-tmRNA (defective, missing 5-prime end)	17	19	25	30	15	19	211
Enterococcus faecium pre-tmRNA	38	38	59	61	57	59	364
Enterococcus faecalis pre-tmRNA, v1 V583	42	42	60	65	57	63	364
Streptococcus equi pre-tmRNA	44	43	52	51	52	51	346
Streptococcus zooepidemicus pre-tmRNA	58	53	52	51	52	51	346
Streptococcus suis pre-tmRNA, v1 P1/7	22	21	55	58	33	35	351
Streptococcus suis pre-tmRNA, v2 89/1591	56	52	35	36	34	35	345
Streptococcus uberis pre-tmRNA	46	40	26	29	27	29	347
Streptococcus pyogenes pre-tmRNA, v1 M1 GAS, Manfredo, M49	59	59	59	61	59	61	345
Streptococcus pyogenes pre-tmRNA, v2 MGAS8232, MGAS315, SSI-1, MGAS10394	60	59	59	61	59	61	345
Streptococcus pyogenes pre-tmRNA, v3	34	29	46	49	44	47	347
Streptococcus agalactiae pre-tmRNA, v1 2603V/R, A909	34	29	46	49	44	47	347
Streptococcus agalactiae pre-tmRNA, v2 NEM316	41	37	44	47	43	46	347
Streptococcus mutans pre-tmRNA	60	56	63	64	62	63	346
Streptococcus gordonii pre-tmRNA	50	51	62	64	62	64	346
Streptococcus pneumoniae pre-tmRNA, v1 Tigr4, R6	68	63	27	27	8	8	345
Streptococcus pneumoniae pre-tmRNA, v2 R6, 670-6B	68	63	45	49	19	21	345
Streptococcus pneumoniae pre-tmRNA, v3 23F	31	28	50	53	50	53	345
Streptococcus mitis pre-tmRNA	68	63	42	46	43	48	345
Streptococcus sobrinus pre-tmRNA	48	44	53	55	33	32	346
Streptococcus thermophilus pre-tmRNA, v2 LMD-9	71	62	41	46	42	46	347
Streptococcus thermophilus pre-tmRNA, v3 CNRZ1066, LMG 18311	68	63	41	46	42	46	347
Lactococcus lactis pre-tmRNA, v1 I11403, CNRZ124, I11403, CNRZ124, IL986, CNRZ188	40	36	18	22	17	21	356
L. lactis pre-tmRNA, v2 MG1363, CNRZ112, CNRZ355, IL582, CNRZ119, IL960, IL2961, SK11	33	34	18	23	42	50	356
Actinomyces naeslundii pre-tmRNA	75	79	46	49	54	55	369
Brevibacterium linens pre-tmRNA	42	43	32	33	32	33	376
Arthrobacter FB24 tmRNA	49	60	40	45	39	45	369
Tropheryma whipplei pre-tmRNA	46	47	36	40	34	39	367
Clavibacter michiganensis tmRNA	52	55	31	38	47	54	376
Leifsonia xyli tmRNA	43	51	24	29	25	29	394
Corynebacterium diphtheriae tmRNA	65	56	52	55	54	57	385
Corynebacterium glutamicum pre-tmRNA	63	62	67	72	49	53	376
Corynebacterium efficiens pre-tmRNA	81	77	62	62	64	67	374
Mycobacterium avium pre-tmRNA, v1 104	32	34	31	35	39	44	366
Mycobacterium avium pre-tmRNA, v2 k10	32	34	31	35	39	44	367
Mycobacterium bovis pre-tmRNA	46	47	26	31	27	30	367
Mycobacterium microti pre-tmRNA, v1	46	47	26	31	27	30	367
Mycobacterium tuberculosis pre-tmRNA, v1 H37Rv, CDC1551	46	47	26	31	27	30	367
Mycobacterium africanum pre-tmRNA (ends unsequenced)	46	47	26	31	27	30	367
Mycobacterium tuberculosis pre-tmRNA, v2 210	41	44	28	31	38	41	367
Mycobacterium leprae pre-tmRNA	43	42	29	30	29	30	368
Mycobacterium marinum pre-tmRNA, v1	54	52	35	39	35	39	366
Mycobacterium smegmatis tmRNA	60	58	31	35	31	35	369
Nocardia farcinica pre-tmRNA	38	39	57	61	38	42	366
Propionibacterium acnes pre-tmRNA, v1	42	48	42	47	41	45	369
Propionibacterium acnes pre-tmRNA, v2	46	53	42	47	38	43	369
Propionibacterium acnes pseudo-tmRNA, v1 (defective, missing 5-prime end)	12	14	15	19	15	19	325
Propionibacterium acnes pseudo-tmRNA, v2 (defective, missing 5-prime end)	20	24	10	12	15	19	325
Streptomyces coelicolor pre-tmRNA	61	63	28	31	33	36	397
Streptomyces scabies pseudo-tmRNA (defective, missing 3-prime end)	40	46	34	41	33	42	339
Streptomyces collinus tmRNA	26	28	22	26	23	26	386
Streptomyces aureofaciens tmRNA	23	25	24	26	11	12	382
Streptomyces avermitilis pre-tmRNA	52	56	28	33	29	33	387
Thermobifida fusca tmRNA	54	55	40	42	42	46	372
Kineococcus radiotolerans pre-tmRNA	41	43	31	33	31	33	364
Frankia sp. Cc13 pre-tmRNA	52	54	53	57	33	37	372
Frankia sp. EAN1pec pre-tmRNA	42	44	25	28	30	32	375
Bifidobacterium longum tmRNA, v1 DJO10A	61	65	48	59	47	54	397
Bifidobacterium longum tmRNA, v2 NCC2705	52	58	27	32	37	45	397
Rubrobacter xylanophilus pre-tmRNA	42	49	29	33	29	33	358
Synechococcus elongatus pre-tmRNA	38	41	28	31	23	25	391
Synechococcus PCC7002 pre-tmRNA	46	46	27	31	34	39	393
Synechocystis PCC6803 pre-tmRNA	48	43	21	21	21	21	396
Crocospaera watsonii pre-tmRNA	47	45	23	26	23	26	383
Thermosynechococcus elongatus pre-tmRNA	56	57	54	55	55	55	383
Trichodesmium erythraeum pre-tmRNA	56	60	29	31	28	30	395
Nostoc PCC7120 pre-tmRNA	50	48	17	20	13	14	388
Nostoc punctiforme pre-tmRNA	45	41	24	27	30	34	388
Anabaena variabilis pre-tmRNA	17	17	18	20	23	25	388
Mesostigma viride plastid pre-tmRNA	46	57	45	61	45	61	356
Nephroselmis olivacea plastid tmRNA	33	54	48	77	48	77	327
Cyanophora paradoxa plastid pre-tmRNA	44	65	51	83	51	83	292
Thalassiosira weissflogii plastid pre-tmRNA	30	60	34	72	33	68	345

tmRNA	McQFold		RNAFold		pknotsRG		sequence length
	PPV	Sens.	PPV	Sens.	PPV	Sens.	
<i>Thalassiosira pseudonana</i> plastid pre-tmRNA	39	64	31	66	31	66	341
<i>Odontella sinensis</i> plastid pre-tmRNA	40	79	29	65	35	75	368
<i>Emiliana huxleyi</i> plastid pre-tmRNA	23	50	15	32	15	32	288
<i>Cyanidium caldarium</i> plastid pre-tmRNA	57	64	58	87	58	87	252
<i>Cyanidioschyzon merolae</i> plastid pre-tmRNA	44	100	44	100	42	96	233
<i>Porphyra purpurea</i> plastid pre-tmRNA	41	62	39	70	38	68	320
<i>Gracilaria tenuistipitata</i> plastid pre-tmRNA	22	60	30	93	29	90	393
<i>Guillardia theta</i> plastid pre-tmRNA	16	31	6	13	6	13	327
<i>Mycobacteriophage Bxz1</i> pre-tmRNA	10	18	36	66	36	66	435
<i>Bacillusphage G</i> pre-tmRNA	34	77	32	77	32	74	309
<i>Aquifex aeolicus</i> tmRNA	65	67	69	72	69	72	347
<i>Sulforhodospira azorensis</i> pre-tmRNA	85	87	83	80	83	80	349
<i>Thermotoga maritima</i> tmRNA, v1 MSB8	50	48	73	70	81	78	356
<i>Thermotoga neapolitana</i> tmRNA	71	73	73	72	73	75	357
<i>Thermodesulfobacterium commune</i> tmRNA	72	74	75	71	74	75	355
<i>Deinococcus radiodurans</i> tmRNA	50	49	41	43	46	47	349
<i>Deinococcus geothermalis</i> tmRNA	30	30	20	19	20	19	348
<i>Thermus thermophilus</i> tmRNA, v1 HB8	40	44	41	44	41	44	349
<i>Thermus thermophilus</i> tmRNA, v2 HB8, HB27	47	51	41	44	41	44	349
<i>Thermomicrobium roseum</i> tmRNA	65	66	59	65	59	64	353
<i>Dehalococcoides ethenogenes</i> tmRNA	45	43	42	40	48	47	352
<i>Leptospirillum groupII</i> tmRNA	56	58	20	21	33	32	359
<i>Leptospirillum groupIII</i> tmRNA	37	39	20	21	20	21	359
<i>Thermodesulfobrevibrio yellowstonii</i> tmRNA	66	67	28	29	28	29	350
<i>Bacteroides fragilis</i> pre-tmRNA, v1 NCTC 9434, YCH46	37	39	14	17	14	17	396
<i>Bacteroides fragilis</i> pre-tmRNA, v2 638R	37	39	14	17	14	17	396
<i>Bacteroides thetaiotaomicron</i> pre-tmRNA	36	39	19	23	20	23	396
<i>Tannerella forsythensis</i> pre-tmRNA	15	17	12	14	12	14	398
<i>Porphyromonas gingivalis</i> tmRNA	47	46	20	24	38	42	407
<i>Prevotella intermedia</i> pre-tmRNA	37	43	20	24	51	62	398
<i>Prevotella ruminicola</i> pre-tmRNA	45	49	37	42	50	60	399
<i>Cytophaga hutchinsonii</i> tmRNA	28	33	19	25	19	25	404
<i>Salinibacter ruber</i> pre-tmRNA	17	17	35	36	37	37	367
<i>Chlorobium tepidum</i> tmRNA	21	21	32	37	31	36	404
<i>Chlorochromium aggregatum</i> tmRNA	58	59	56	62	56	62	401
<i>Gemmata obscuriglobus</i> pre-tmRNA	25	28	44	47	20	22	410
<i>Rhodopirellula baltica</i> pre-tmRNA	20	23	44	49	41	48	378
<i>Chlamydia trachomatis</i> pre-tmRNA, v1 D/UW-3/CX	29	29	32	35	32	36	417
<i>Chlamydia trachomatis</i> pre-tmRNA, v2 LGV	35	35	33	37	32	36	417
<i>Chlamydia trachomatis</i> pre-tmRNA, v3 Jali	29	29	32	35	32	36	417
<i>Chlamydia muridarum</i> pre-tmRNA	30	30	26	29	35	38	418
<i>Chlamydia pneumoniae</i> pre-tmRNA	30	35	32	38	29	35	423
<i>Chlamydia caviae</i> pre-tmRNA	40	41	48	55	49	56	422
<i>Chlamydia abortus</i> pre-tmRNA	48	47	50	56	44	51	422
<i>Parachlamydia UWE25</i> pre-tmRNA	49	64	24	29	28	35	402
<i>Verrucomicrobium spinosum</i> tmRNA	24	23	37	38	35	40	356
<i>Borrelia burgdorferi</i> pre-tmRNA	53	51	37	36	37	36	359
<i>Borrelia garinii</i> pre-tmRNA	47	43	39	41	41	43	360
<i>Treponema pallidum</i> pre-tmRNA	72	74	46	51	49	54	352
<i>Treponema denticola</i> pre-tmRNA	54	50	47	46	51	53	348
<i>Leptospira interrogans</i> pre-tmRNA	47	46	60	60	59	59	347
<i>Fibrobacter succinogenes</i> pre-tmRNA	35	36	39	42	33	34	358
<i>Acidobacterium capsulatum</i> tmRNA	41	41	22	23	23	23	352
<i>Acidobacterium Ellin6076</i> tmRNA	37	38	37	41	35	39	352
Uncultured farm2 tmRNA	53	55	41	44	42	44	356
<i>Fusobacterium nucleatum</i> tmRNA, v1 ATCC 25586	32	26	22	24	21	23	344
<i>Fusobacterium nucleatum</i> tmRNA, v2 ATCC 10953	30	26	64	61	60	62	344
<i>Fusobacterium nucleatum</i> tmRNA, v3 ATCC 49256	43	37	53	53	60	64	344
<i>Dictyoglomus thermophilum</i> tmRNA	87	85	75	72	85	85	350
Uncultured farm3 tmRNA	10	13	11	16	18	27	365
Uncultured farm8 tmRNA	21	23	18	21	18	21	356

A.2 tRNA

The structures of the 21 tRNA sequences given in the BRaliBase, cf. [GG04, GJBM⁺03, SMS⁺76], <http://www.binf.ku.dk/~pgardner/bralibase>, in March 2007 do not contain pseudo-knots. We skipped one of the sequences because it contains an undetermined nucleotide, which cannot be handled by the current version of McQFold. The 20 sequences we used are listed in the following table. The first 11 of them are highly similar to each other and therefore should not be considered as a representative sample of tRNA molecules.

tRNA	McQFold		RNAFold		pknotsRG		sequence length
	PPV	Sens.	PPV	Sens.	PPV	Sens.	
Yeast Phe-tRNA	95	100	30	35	33	35	73
Yeast major Phe-tRNA	95	100	24	25	24	25	73
E. coli tRNA-Phe	95	100	48	55	24	25	73
Yeast minor Phe-tRNA	95	100	27	30	27	30	73
N. crassa tRNA-Phe	94	85	89	85	94	85	73
Lupin minor Phe-tRNA	94	85	33	40	35	40	73
Barley tRNA-Phe	35	40	35	40	36	40	73
Lupin major Phe-tRNA	94	85	17	20	17	20	73
E. gracilis Phe-tRNA	76	80	28	35	28	35	73
X. leavis Phe-tRNA	76	80	28	35	28	35	73
Drosophila-Phe	76	80	52	55	52	55	73
H. volcanii Lys-tRNA-1	67	70	44	55	44	55	74
Bacteriophage T4 Pro-tRNA	87	100	87	100	87	100	74
E. coli tRNA-Val-1	30	35	29	30	29	30	73
E. coli glyW gene	65	75	62	75	91	100	73
Yeast Met-tRNA	100	100	75	75	75	75	73
Bacteriophage T4 Thr-tRNA	95	100	95	95	95	95	73
Yeast mitochondrial Phe-tRNA	61	55	50	55	50	55	72
T.thermophila cytoplasmic tRNA-Tyr	86	95	75	75	75	75	73
H.volcanii Phe-tRNA	95	100	42	55	44	55	74

A.3 RNase P

The RNase P sequences and structures given in the “RNase P Database” (cf. [Bro99], <http://www.mbio.ncsu.edu/RNaseP>) are subdivided into 20 groups. For each of these groups we selected the sequence top-most in the corresponding table among all sequences completely available in March 2007. We skipped the group “unknown bacterium” because it only contained incomplete sequences.

RNase P	McQFold		RNAFold		pknotsRG		sequence length
	PPV	Sens.	PPV	Sens.	PPV	Sens.	
A. tumefaciens	65	60	82	76	82	76	402
A. eutrophus	57	46	65	58	64	58	341
A. ferrooxidans	33	29	44	39	45	41	344
D. desulfuricans	60	56	52	50	52	50	360
C. diphtheriae	59	53	55	50	57	50	425
C. hydrogenoformans	76	76	67	68	67	68	331
Anabaena-PCC7120	65	54	63	54	59	54	457
B. thetaiotaomicron	42	33	37	32	56	48	361
B. burgdorferi	37	31	50	51	50	51	346
C. trachomatis	60	54	66	64	67	65	408
C. limicola	20	18	44	39	34	31	386
C. aurantiacus	42	39	54	53	55	54	328
D. radiodurans	55	48	44	41	43	41	486
T. maritima	45	43	54	54	49	49	338
Halobacterium-NRC1	57	53	29	32	42	42	375
A. pernix	48	46	24	25	24	25	330
A. telluris	36	67	26	56	30	62	364
R. americana-mito	46	46	39	36	39	36	312
C. paradoxa-cyanelle	49	46	52	46	52	46	351

A.4 Pseudoknot Fragments

We used the contiguous pseudoknot fragments from the PseudoBase, cf. [vBGP⁺00], <http://biology.leidenuniv.nl/~batenburg/PKBGet.html>, for which all nucleotides were fully determined. We grouped the sequences in the same way as in the PseudoBase:

Group of Pseudoknot Fragments	McQFold		RNAFold		pknotsRG		number of sequences	mean length of sequences
	PPV	Sens.	PPV	Sens.	PPV	Sens.		
Aptamers	96	90	79	65	83	83	6	45.83
Artificial	100	62	100	62	100	100	1	26
mrna	60	65	45	46	45	45	7	63.71
Others	59	55	65	61	64	64	4	68.25
OtherViral3UTR	89	65	84	58	90	90	90	30.83
OtherViral5UTR	94	71	91	48	94	94	2	34.5
Ribozymes	78	77	51	47	70	70	4	90
rrna	56	54	53	50	51	51	3	48.33
tmRNA	62	49	68	50	70	70	10	57.5
ViralFrameshift	61	63	39	40	67	67	17	57.35
ViralOthers	69	56	71	59	77	77	20	49.35
ViralReadthrough	62	78	6	7	48	48	7	61.71
ViralTRNALike	67	70	59	59	67	67	57	53.25

References

- [BBC⁺00] P. Baldi, S. Brunak, Y. Chauvin, C. A. Andersen, and H. Nielsen. Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*, 16(5):412–24, 2000.
- [Bro99] J. W. Brown. The Ribonuclease P Database. *Nucleic Acids Res*, 27(1):314, 1999.
- [DE04] R. D. Dowell and S. R. Eddy. Evaluation of several lightweight stochastic context-free grammars for RNA secondary structure prediction. *BMC Bioinformatics*, 5:71, 2004.
- [GG04] P. P. Gardner and R. Giegerich. A comprehensive comparison of comparative RNA structure prediction approaches. *BMC Bioinformatics*, 5:140, 2004.
- [GJBM⁺03] S. Griffiths-Jones, A. Bateman, M. Marshall, A. Khanna, and S. R. Eddy. Rfam: an RNA family database. *Nucleic Acids Res*, 31(1):439–41, 2003.
- [Hof03] I. L. Hofacker. Vienna RNA secondary structure server. *Nucleic Acids Res*, 31(13):3429–31, 2003.
- [Mat04] D. H. Mathews. Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA*, 10(8):1178–90, 2004.
- [RG04] J. Reeder and R. Giegerich. Design, implementation and evaluation of a practical pseudoknot folding algorithm based on thermodynamics. *BMC Bioinformatics*, 5:104, 2004.
- [SMS⁺76] M. Sundaralingam, H. Mizuno, C. D. Stout, S. T. Rao, M. Liedman, and N. Yathindra. Mechanisms of chain folding in nucleic acids. The (omega, omega) plot and its correlation to the nucleotide geometry in yeast tRNAPhe1. *Nucleic Acids Res*, 3(10):2471–84, 1976.
- [vBGP⁺00] F. H. van Batenburg, A. P. Gulyaev, C. W. Pleij, J. Ng, and J. Oliehoek. PseudoBase: a database with RNA pseudoknots. *Nucleic Acids Res*, 28(1):201–4, 2000.
- [WB98] K. P. Williams and D. P. Bartel. The tmRNA Website. *Nucleic Acids Res*, 26(1):163–5, 1998.
- [Wil00] K. P. Williams. The tmRNA website. *Nucleic Acids Res*, 28(1):168, 2000.
- [ZS81] M. Zuker and P. Stiegler. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res*, 9(1):133–48, 1981.