Transfer RNA as a Potential Building Block for Nanotechnology

Harold S. Bernhardt and Warren P. Tate

Biochemistry Department, University of Otago, Dunedin (e-mail: berha498@student.otago.ac.nz; warren.tate@stonebow.otago.ac.nz).

Nucleic acid nanotechnology

Introduction

It is almost 25 years ago since Nadrian Seeman first proposed the synthesis of X-shaped nucleic acid structures based on the Holliday junction, the cross-over structure of two strands of DNA that occurs as an intermediate in genetic recombination.[1] Since then, Seeman and other groups have produced these and other DNA structures, including a cube and an octahedron.[2-4] One of the main goals of nucleic acid nanotechnology has been the synthesis of components which will form self-assembling 3D arrays with such possible uses as molecular sieves, or scaffolds on which to fix macromolecules for X-ray crystallography, or for use in molecular electronics.

RNA as a potential building block

Over the last five years RNA has been used increasingly in this work. Although more labile than DNA, it has the advantage of being able to form a more diverse range of structures due to its more complex chemistry. It still retains the key attribute of nucleic acids: the specificity of Watson-Crick base-pairing interactions that allows the precise positioning of component parts to form a larger, predictable structure. Systems based upon base-pairing interactions have included the dimerization initiation site of HIV RNA and the right angle (RA) motif found in ribosomal RNA,[5,6] while those based on non-base-pairing tertiary structural associations have included the GAAA tetraloop-receptor interaction which occurs widely in biological systems.[7] A necessary requirement of such systems is inducibility - the ability to be able to control the rate of association; the addition of Mg^{2+} is commonly used to promote aggregation.

Is transfer RNA a candidate building block?

Perhaps surprisingly, as yet there have been no published reports using transfer RNA (tRNA) as a building block for the construction of nanomaterials (tRNA transfers individual amino acids to the growing polypeptide chain during ribosomal protein synthesis). This is surprising, since tRNA possesses a number of attributes that make it particularly suitable. It has a small size and high stability, a well-characterized cloverleaf (2D) and L-shape (3D) structure, and, most importantly, it has the ability to form dimers and higher aggregates through interactions between its single-stranded loops (anticodon loop, D-loop and T-loop). One of the best known examples is yeast $tRNA^{Asp(GUC)}$ (which transfers aspartic acid to the growing polypeptide chain), which forms dimers in solution and in the crystal through its quasi-self-complementary GUC anticodon (Fig. 1).[8] Another example is the A14G mutant human mitochondrial $tRNA^{Leu(UAA)}$ (which transfers leucine to the growing polypeptide chain in the synthesis of mitochondrial proteins), where the mutation creates a self-complementary six base sequence that allows formation of a dimer (Fig. 2).[9] This mutation is associated with two disease conditions, maternally-inherited diabetes and deafness (MIDD), and mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (MELAS).[9]

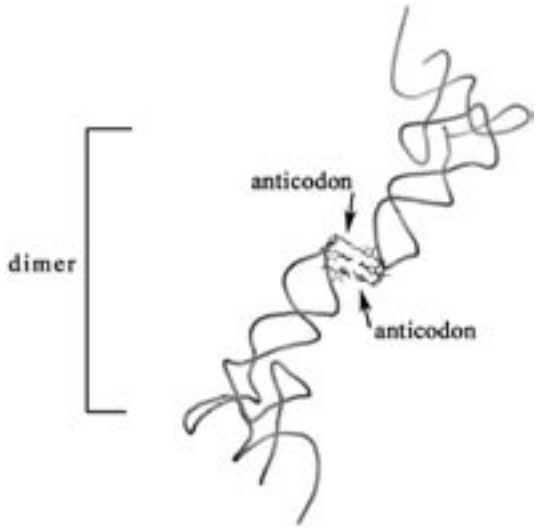

Fig. 1. A dimer from Yeast $tRNA^{Asp(GUC)}$. The graphic (taken from PDB file (see ref. 8) has only anticodons represented fully and shows 3D L-shape of tRNA.

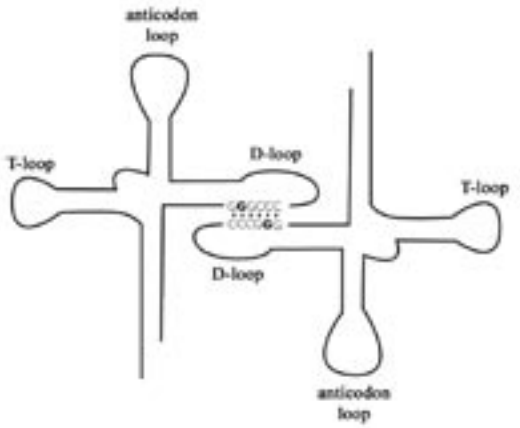

Fig. 2. Dimer from unmodified A14G mutant human mitochondrial $tRNA^{Leu(UAA)}$. The 2D schematic has only complementary sequences represented fully and shows cloverleaf secondary structure including D-loop, anticodon loop and T-loop positions (see ref. 9).

Experimental strategy

$tRNA^{Gly(GCC)}$ from E. coli

One tRNA that has been shown to aggregate is $tRNA^{Gly(GCC)}$ (which transfers glycine to the ribosome) from *E. coli*, yeast and *Bombyx mori* (silkworm).[10-12] Romby *et al.*[13] have provided experimental evidence that *E. coli* $tRNA^{Gly(GCC)}$

forms dimers through its GCC anticodon in a similar manner to yeast $tRNA^{Asp(GUC)}$, but unlike $tRNA^{Asp(GUC)}$, only at pH 4-5. They have suggested this is due to the formation of a stabilizing hemi-protonated C-C(+) base-pair between the two middle bases of the anticodons, whereas, at neutral pH, the C-C interaction is destabilizing.[13] Despite this evidence that it only forms dimers, we chose *E. coli* $tRNA^{Gly(GCC)}$ for our initial studies, partly because it appeared to have the most stable structure of the three $tRNA^{Gly(GCC)}$s, and for the reasons discussed below. As our laboratory already held samples of *E. coli* genomic DNA, our initial approach was to amplify the $tRNA^{Gly(GCC)}$ gene using PCR, and follow this with *in vitro* transcription to produce unmodified $tRNA^{Gly(GCC)}$. Although tRNAs *in vivo* are always modified at a number of positions, unmodified tRNA transcripts produced *in vitro* have been shown to have a similar, if somewhat looser structure. This difference was thought not to be a disadvantage for our studies - and perhaps even preferred, as suggested by the study of the A14G mutant human mitochondrial $tRNA^{Leu(UAA)}$ dimer that used unmodified *in vitro* transcripts.[9]

Multiple gene transcripts

E. coli possesses four identical copies of the $tRNA^{Gly(GCC)}$ gene that occur at two separate loci on opposite strands as well as opposite sides of the *E. coli* genome (Fig. 3).[14] *glyW* occurs as a single gene, while *glyV-glyX-glyY* occur as a triple gene cluster, with the three genes separated by short sequences of 36 and 35 base pairs. Since we wanted principally to amplify the single gene, we used internal primers complementary to the 5'- and 3'- ends of the tRNA, with the forward primer also containing the 17 base T7 RNA polymerase promoter sequence to enable subsequent *in vitro* transcription. We realized however, that the primers had the potential to amplify the double genes (*glyV-glyX* and *glyX-glyY*) and even the triple gene (*glyV-glyX-glyY*) (Fig. 4). The RNA transcripts of these products might possess interesting features of their own for nanotechnological building blocks. They should possess correct tRNA structures as the primary triple gene transcript is cleaved into individual tRNAs *in vivo* by RNase P, an enzyme that recognizes the tertiary structure of the tRNA.[14] The double and triple gene transcripts containing two or three GCC anticodons should also be capable of aggregating (at least at low pH) and the triple gene transcript that contains three GCC anticodons, might possibly be capable of forming 3D structures.

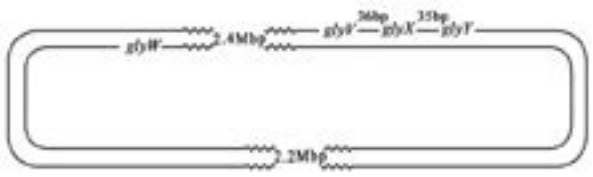

Fig. 3. *In vivo* transcribed $tRNAGly^{(GCC)}$ is a single gene product from *glyW* and a triple gene product from *glyV/glyX/glyY* (see ref. 14); (bp = base pair, Mbp = 10^6 base pairs).

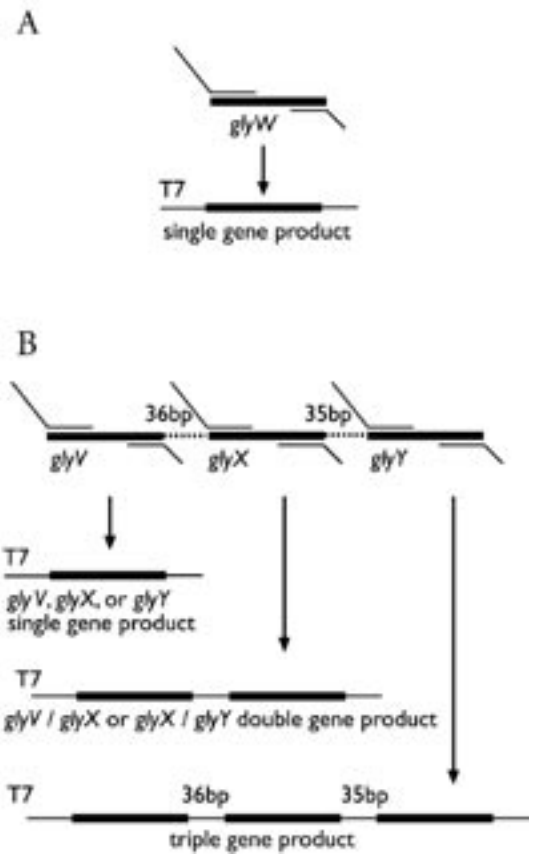

Fig. 4. (a) The predicted amplification of the single $tRNAGly^{(GCC)}$ gene product from the *glyW* locus, and (b) that of the single, double and triple $tRNAGly^{(GCC)}$ gene products from the *glyV/glyX/glyY* locus.

Experimental results

PCR of single, double, and triple gene products

Our first PCR experiment gave bands corresponding to the single, double and triple gene products, but the two double genes were not separated, as they differ in length by only a single base pair (Fig. 5). A higher molecular weight band running at >500 base pairs was unexpected and its origin has yet to be defined. Subsequent attempts to tweak the parameters and increase the yield of the double and (especially) triple gene product have so far been unsuccessful; no triple gene product was detected but consistent quantities of single and double gene products were obtained. This difficulty might be resolved by increasing the concentration of *E. coli* genomic DNA used, and/or by using a different DNA polymerase; strategies for the future.

In vitro *transcription*

In vitro transcription of particular tRNA genes can be notoriously difficult. Sometimes it has necessitated the use of tricks such as incorporating a copy of the hammerhead ribozyme gene in the tRNA gene-containing construct in order to increase the yield of tRNA product.[15] However, in our case, the DNA products generated by PCR (containing a mixture of single, double and triple $tRNA^{Gly(GCC)}$ genes) were converted successfully into RNA, utilizing direct *in vitro* transcription by T7 RNA polymerase; three bands corresponding to the single, double and triple $tRNA^{Gly(GCC)}$ gene transcripts were observed on a denaturing RNA gel (Fig. 6).

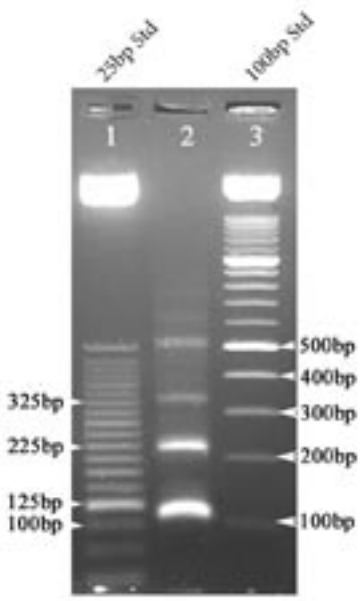

Fig. 5. PCR amplification of single, double and triple tRNAGly$^{(GCC)}$ genes from *E. coli* genomic DNA gave bands equivalent to those expected: 111, 222/223 and 334bp (with the two double gene products running as a single band) on a 2% agarose gel. The origin of an additional >500bp band shown here is unresolved. Lane 1, 25bp DNA standard (25bp band not visible); lane 2, PCR products (as described above); lane 3, 100bp DNA standard (bp = base pairs).

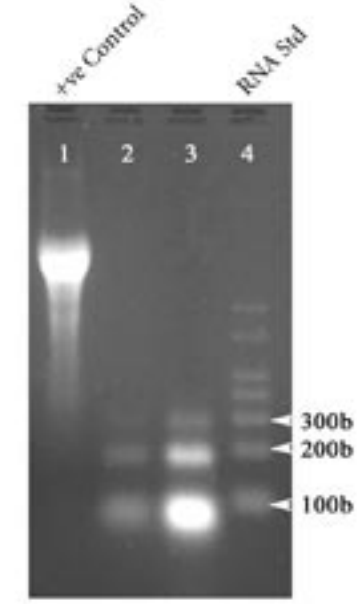

Fig. 6. Products ca. 85, 196/197 and 308b in size from *in vitro* transcription of single, double and triple tRNA$^{Gly(GCC)}$ gene using T7 RNA polymerase; lane 1 is a luciferase mRNA control and b = bases (see text).

In order to achieve these levels of transcription it was necessary to develop the *in vitro* transcription protocol by (*a*) quadrupling the initial T7 RNA polymerase concentration and reducing the concentration of rNTPs, template DNA, and RNase-inhibitor accordingly, and increasing reaction time to 16-24 h (Fig. 6, lane 2); and (*b*), by also adding an extra aliquot of T7 RNA polymerase 3 h into the reaction as described by Roy *et al*.[9] (Fig. 6, lane 3). As Fig. 6 shows, the addition of an extra aliquot of enzyme dramatically increases the level of transcription.

In vitro *transcripts of wild-type and A14G mutant human mitochondrial tRNA$^{Leu(UAA)}$*

Early in this work gifts of two pUC18 plasmids with constructs containing the wild-type and A14G mutant human mitochondrial tRNA$^{Leu(UAA)}$ genes, respectively, downstream of a T7 RNA polymerase promoter were received from Kelley's group (Department of Chemistry, Boston College). We were keen to use these as positive and negative controls for dimerization as Kelley's group had demonstrated that the A14G mutant (but not the wild-type) forms dimers under normal conditions.[9]

The plasmids were linearized using *Mva1* restriction enzyme. The plasmid fragments were separated on a 2% agarose gel, and the 155 base pair fragment that contained either the wild-type or A14G mutant tRNA$^{Leu(UAA)}$ gene with the T7 promoter was then purified by use of a gel-extraction kit. Finally, the wild-type and A14G mutant plasmid DNA fragments were incubated in an *in vitro* transcription reaction modified as previously described. Even then, and using similar amounts of template DNA for each of the three tRNAs, the two tRNA$^{Leu(UAA)}$ plasmid fragments produced a much lower level of transcription than the single tRNA$^{Gly(GCC)}$ gene PCR product as shown in Fig. 7. The single tRNA$^{Gly(GCC)}$ gene was obtained by gel extraction of the single gene band on a preparative 2% agarose gel of a PCR reaction similar to that shown in Fig. 5.

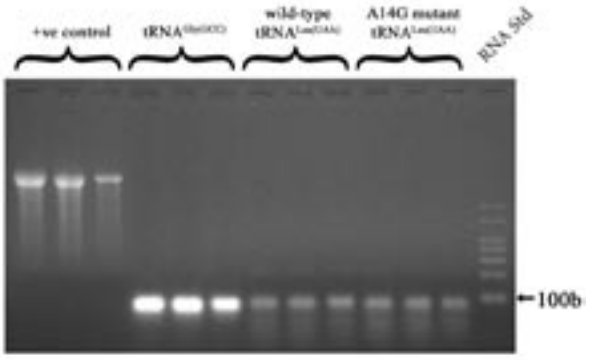

Fig. 7. *In vitro* transcription of single tRNA$^{Gly(GCC)}$ gene and wild-type and A14G mutant human mitochondrial tRNA$^{Leu(UAA)}$ genes; luciferase mRNA positive control; 1st lane of each set of 3: *in vitro* transcription; 2nd lane: DNase cleavage of the template DNA; 3rd lane: RNA extraction.

Reportedly, *in vitro* transcription from a plasmid fragment (as is the case with wild-type/A14G mutant tRNA$^{Leu(UAA)}$) is three to four times more efficient than that directly from a PCR product (as is the case with tRNA$^{Gly(GCC)}$). This makes the difference in level of transcription even more striking and probably results from the different internal promoter sequences of *E. coli* tRNA$^{Gly(GCC)}$ and human mitochondrial tRNA$^{Leu(UAA)}$. As already noted, each of the constructs has a 17 base T7 RNA polymerase promoter sequence upstream of the tRNA. However, the first six bases of the tRNA itself also play a part in the binding of T7 RNA polymerase by functioning as an internal promoter. The internal promoter sequence of *E. coli* tRNA$^{Gly(GCC)}$ is much closer to the ideal T7 sequence than that of human mitochondrial tRNA$^{Leu(UAA)}$ thus allowing more efficient transcription.

Aggregation of building block tRNAs analyzed by native gel electrophoresis

Using single $tRNA^{Gly(GCC)}$ transcript and wild-type and A14G mutant human mitochondrial $tRNA^{Leu(UAA)}$, preliminary experiments to effect aggregation consisted of heating the tRNAs to 70°C for 5 min, rapid cooling on wet ice, and addition of Mg^{2+} to give a 10 mM final concentration as described by Roy *et al*.[9] Aliquots of these reactions have been electrophoresed using native PAGE at either 4°C or RT for 1-4 h. The results provide preliminary evidence of multiple bands with the A14G mutant human mitochondrial $tRNA^{Leu(UAA)}$, *i.e.* aggregation. In order to observe aggregation of the $tRNA^{Gly(GCC)}$ single transcript the gel should be at pH 4-5; the use of a size separation column to isolate the dimer is yet to be investigated.

In search of the first tRNA

An RNA world

In addition to possible applications in nanotechnology, we are also interested in the possible relationship between tRNA aggregation and the origin and evolution of tRNA function, the genetic code and protein synthesis. It appears likely that life was based on RNA prior to the advent of protein synthesis as it can function both as an enzyme and a carrier of genetic information; this is known as the *RNA world*. In light of this hypothesis, and the central role of tRNA in protein synthesis, it seems almost certain that tRNA is an extremely ancient molecule, with its primary, secondary and tertiary structure highly conserved over all three kingdoms of life.

Hairpin loops

Di Giulio has proposed that tRNA arose by the duplication of a shorter hairpin loop.[16] Paul Schimmel's group has demonstrated that hairpin loops combining a single-stranded 3'-terminal CCA sequence with stem sequences related to those of a number of tRNAs are specifically aminoacylated by contemporary aminoacyl-tRNA synthetases, and so could have participated in an early version of non-templated protein synthesis in the absence of the equivalent of an mRNA message.[17] The same group has also demonstrated lateral (side-by-side) H-bonding interactions between these loops that could have brought the two 3'-CCA ends into proximity, enabling the polymerization of amino acids into proteins prior to the advent of the ribosome.[18] Due to the symmetry of base-pairing interactions, it seems possible that these single hairpin loops may have been in equilibrium with hairpin loop-duplexes having a 2D structure similar to tRNA (Fig. 8).

Split tRNAs

Recently, it has been discovered that the archaeal species *Nanoarchaeum equitans* has four split tRNA genes with the split occurring between bases 37 and 38 in the anticodon loop.[19] This is also the most common insertion position for tRNA introns. The anticodon loop sequence of the overwhelming majority of glycine tRNAs that participate in protein synthesis is 5'- YUNCCA$^{\downarrow}$N - 3', where 'Y' denotes U or C, 'N' any base, and the NCC anticodon is underlined.[20] If one were to cleave glycine tRNA between bases 37 and 38, the position of the arrow in the above, one would be left with two *half* tRNAs of almost identical size, each with a 3'-terminal -CCA, and so conceivably both descended from a single hairpin loop that was aminoacylated by glycine and participated in non-templated protein synthesis.

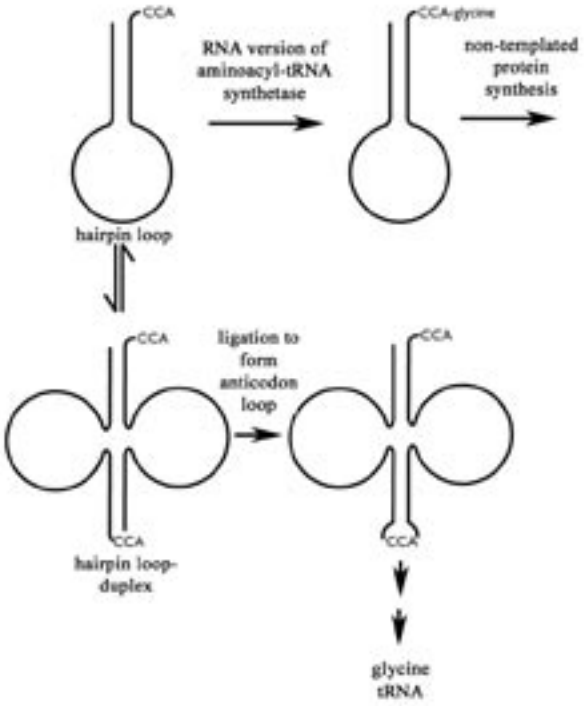

Fig. 8. Proposed evolution of glycine tRNA from hairpin loop.

Origin of the genetic code

Thus it seems possible that tRNA arose by the ligation of two identical hairpin loops, with CCA becoming part of the anticodon loop, (N)CC the first anticodon/GG(N) the first codon (coding for glycine), and glycine tRNA the first tRNA (Fig. 8). That glycine tRNA might have been the first tRNA is not inconsistent with the argument that glycine was possibly the first amino acid. After all it is the smallest, simplest, and only achiral amino acid, and it is almost always produced in experiments that seek to simulate earth's proposed early atmosphere. One of the most highly conserved sequences among cytoplasmic tRNAs is the D-loop -GG- sequence, which in $tRNA^{Gly(GCC)}$ forms part of a seven base sequence complementary to the anticodon loop. It is possible that this highly conserved -GG- sequence is a descendent of the ancestor of the first codon, and that the first anticodon-codon interaction originally had a role in tRNA aggregation.

Summary

We hope that this article gives a taste of the potential for this field of study. Not only may it lead to the creation of 3D arrays for use in nanotechnological applications, but also it may be of critical importance to our understanding of the origin and evolution of tRNA, the genetic code and protein synthesis.

Acknowledgements

We thank Bronwyn Carlisle for the graphics, Tony Cardno for help with computing, Pascale Romby for advice,

Philippe Dumas for the PDB file of the $tRNA^{Asp}$ dimer (these last two of CNRS-IBMC, Strasbourg), and Lisa Wittenhagen from the Kelley group (Boston College, MA) for the plasmids gifts and for many helpful discussions. We acknowledge the support of the Marsden Fund of NZ for this project (to W.P.T).

References

1. Seeman, N. C. *J. Theor. Biol.* **1982**, *99*, 237-247.

2 Seeman, N. C.; Maestre, M. F.; Ma, R. I.; Kallenbach, N.R. *Prog. Clin. Biol. Res.* **1985**, *172A*, 99-108.

3. Chen, J.; Seeman, N. C. *Nature* **1991**, *350*, 631-633.

4. Shih, W. M.; Quispe, J. D.; Joyce, G. F. *Nature* **2004**, *427*, 618-621.

5 Horiya, S.; Li, X.; Kawai, G.; Saito, R.; Katoh, A.; Kobayashi, K.; Harada, K. *Chem. Biol.* **2003**, *10*, 645-654.

6. Chworos, A.; Severcan, I.; Koyfman, A. Y.; Weinkam, P.; Oroudjev, E.; Hansma, H. G.; Jaeger, L. *Science* **2004**, *306*, 2068-2072.

7. Ikawa, Y.; Fukada, K.; Watanabe, S.; Shiraishi, H.; Inoue, T. *Structure* **2002**, *10*, 527-534.

8. Moras, D.; Dock, A.-C.; Dumas, P.; Westhof, E.; Romby, P.; Ebel, J.-P.; Giegé, R. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 932-936.

9. Roy, M. D.; Wittenhagen, L. M.; Kelley, S. O. *RNA* **2005**, *11*, 254-260.

10. Söll, D.; Cherayil, J. D.; Bock, R. M. *J. Mol. Biol.* **1967**, *29*, 97-112.

11. Hampel, A.; Cherayil, J.; Bock, R. M. *Biochim. Biophys. Acta* **1971**, *228*, 482-491.

12. Amano, M.; Kyogoku, Y. *Eur. J. Biochem.* **1993**, *217*, 131-136.

13. Romby, P.; Westhof, E.; Moras, D.; Giegé, R.; Houssier, C.; Grosjean, H. *J. Biomol. Struct. Dyn.* **1986**, *4*, 193-203.

14. Tucker, S. D.; Gopalakrishnan, A. S.; Bollinger, R.; Dowhan, W.; Murgola, E. J. *J. Bacteriol.* **1982**, *152*, 773-779.

15. Kelley, S. O.; Steinberg, S. V.; Schimmel, P. J. *Biol. Chem.* **2001**, *276*, 10607-10611.

16. Di Giulio, M. *J. Theor. Biol.* **1999**, *197*, 403-414.

17. Musier-Forsyth, K.; Schimmel, P. *Acc. Chem. Res.* **1999**, *32*, 368-375.

18. Henderson, B. S.; Schimmel, P. *Bioorg. Med. Chem.* **1997**, *5*, 1071-1079.

19. Randau, L.; Munch, R.; Hohn, M. J.; Jahn, D.; Söll, D. *Nature* **2005**, *433*, 537-541.

20. Sprinzl, M.; Vassilenko, K. S. *Nucleic Acids Res.* **2005**, *33*, D139-D140; http://www.tRNA.uni-bayreuth.de