New frontiers in studying pseudouridine formation in RNA

Ute Kothe

Pseudouridines

It is more than 50 years ago that a fifth nucleotide has been discovered to be part of RNA.[1, 2] This fifth nucleotide, as it is still sometimes called, is the most abundant, post-transcriptionally modified nucleotide in cellular RNA called pseudouridine.[3] As the name implies, pseudouridines are modified uridines which differ from their parent nucleotide by having a unique C-C glycosidic bond instead of the canonical N-C glycosidic bond (Fig. 1).[2] This modified nucleotide retains the ability to base-pair with adenine; however the additional imino group in the base allows for further hydrogen bonds to form, in particular bridging a water molecule between the base and the preceding phosphodiester bond.[4] As a result, pseudouridines are thought to rigidify RNA in particular by improving base stacking,[3] but roles for the specific function of the modification in non-coding RNA such as ribosomal and spliceosomal RNAs have also been discussed.[5, 6] As pseudouridine has been known for half a century, this raises the question: why are we still interested in this seemingly small adjustment to cellular RNA? There are two main answers to this question which will be both addressed in this review. First, only the last years have allowed us to gain a detailed understanding of pseudouridine synthases which has in turn raised several new questions. And second, we are only beginning to identify the importance of pseudouridine formation for the cell and its potential applications.

Stand-alone Pseudouridine synthases

Each organism contains several pseudouridine synthases which are responsible for the site-specific modification of many different RNA target sites.[7] To date, very comprehensive knowledge has been accumulated on the different bacterial pseudouridine synthases, their respective target sites and also their crystal structures. This information reveals that, despite significant sequence variety, all pseudouridine synthases share a common fold in their catalytic domain which is characterized by a central 8-stranded mixed β-sheet that is surrounded by several loops and α-helices (Fig. 2A and B). In one instance (TruD), this catalytic domain is interestingly formed by a circular permutation in its primary sequence.[8-10] Based on sequence and structure comparison, pseudouridine synthases are classified into six families (Table 1). Representatives of five families are found in all domains of life and are named according to a bacterial representative[7] whereas the sixth family is represented so far by the single enzyme Pus10, present in many archaea and some eukaryotes.[11, 12] The active site of pseudouridine synthases is located in a cleft in the middle of the catalytic domain where the universally conserved aspartate residue is found; this so called catalytic aspartate residue is essential for pseudouridine formation. In addition, the active site contains an aromatic residue, usually a tyrosine, that stacks with the target uridine and might play a role as a general base in a late step of the reaction chemistry,[13] as well as a conserved positive arginine residue, that might stabilize or activate the catalytic aspartate residue (Fig. 2A).[14]

Aside from the catalytic domain, pseudouridine synthases differ significantly from each other in particular with respect to additional domains which are believed to contribute to specific RNA binding.[14, 15] For example, pseudouridine synthases of the TruB family are characterized by an additional C-terminal PUA domain (found in pseudouridine synthases and archaeosine-transglycosylases) that is believed to bind to the tRNA acceptor stem and possibly also the 3'CCA end of tRNA[16] or the 3'ACA end of H/ACA guide RNA (vide infra).[17] Enzymes of the RsuA family and some members of the RluA family have an N-terminal S4-domain resembling the ribosomal S4 protein and may have C-terminal extensions.[7] The TruD protein harbors a unique domain which is inserted into the catalytic domain.[8-10] And lastly, the TruA enzyme has been found to operate as a homodimer where each subunit contains an active site and tRNA is bound across both subunits.[18]

Stand-alone pseudouridine synthases recognize their target sites through a variety of molecular mechanisms and with different degrees of specificity. *Escherichia coli* pseudouridine synthases are responsible for generating pseudouridines at seven positions in specific tRNAs as well as eleven pseudouridines in 16S and 23S ribosomal RNA (rRNA). RluA is the only pseudouridine synthase that is capable of modifying both tRNA and rRNA,[19] but other enzymes also target several different sites within tRNA (TruA) or rRNA (RluC, RluD) (Table 1). The specificity of RNA

recognition by pseudouridine synthases can either be achieved by identifying conserved sequence elements within the RNA or by detecting a particular RNA conformation. Interestingly, all crystal structures of pseudouridine synthases in complex with RNA, so far, reveal different mechanisms of recognizing RNA conformations, such as recognition of: (1) the native RNA conformation (TruB),[14] (2) a distorted RNA structure which is induced upon binding the RNA by the protein (RluA and RluF)[20, 21] or even (3) the flexibility of RNA structure allowing it to adopt different conformations (TruA).[18] In all cases, the pseudouridine synthases have to gain access to the target uracil base in order to modify it, and therefore this base is typically flipped out of the RNA structure and into the active site of the modification enzyme. Along with the target uracil, up to two additional bases might be flipped out and buried in binding pockets of the enzyme.[14] Taken together, these mechanisms of target site recognition ensure that pseudouridines are only found in a small number of defined positions in cellular RNA.

While the study of eukaryotic pseudouridine synthases is lagging behind the investigation of their bacterial counterparts, a few interesting points can be noted. So far, eleven pseudouridine synthases in eukaryotes (mostly in *Saccharomyces cerevisiae*) have been characterized, and their target sites have been identified (Table 1). Strikingly, pseudouridine synthases are not universally conserved, and not all bacterial enzymes have orthologues in eukaryotes. Most notably, no eukaryotic enzyme of the RsuA family has been identified so far. Furthermore, upon comparing bacterial and eukaryotic pseudouridine synthases, it becomes evident that the latter enzymes are mainly responsible for modifying tRNA, sometimes small nuclear RNA (snRNA) and only in one case 5S rRNA and mitochondrial 21S rRNA. This is in clear contrast to the bacterial enzymes which often target ribosomal RNA, and reveals the evolution of pseudouridine synthases to recognize very different and new target sites in the eukaryotic kingdom.

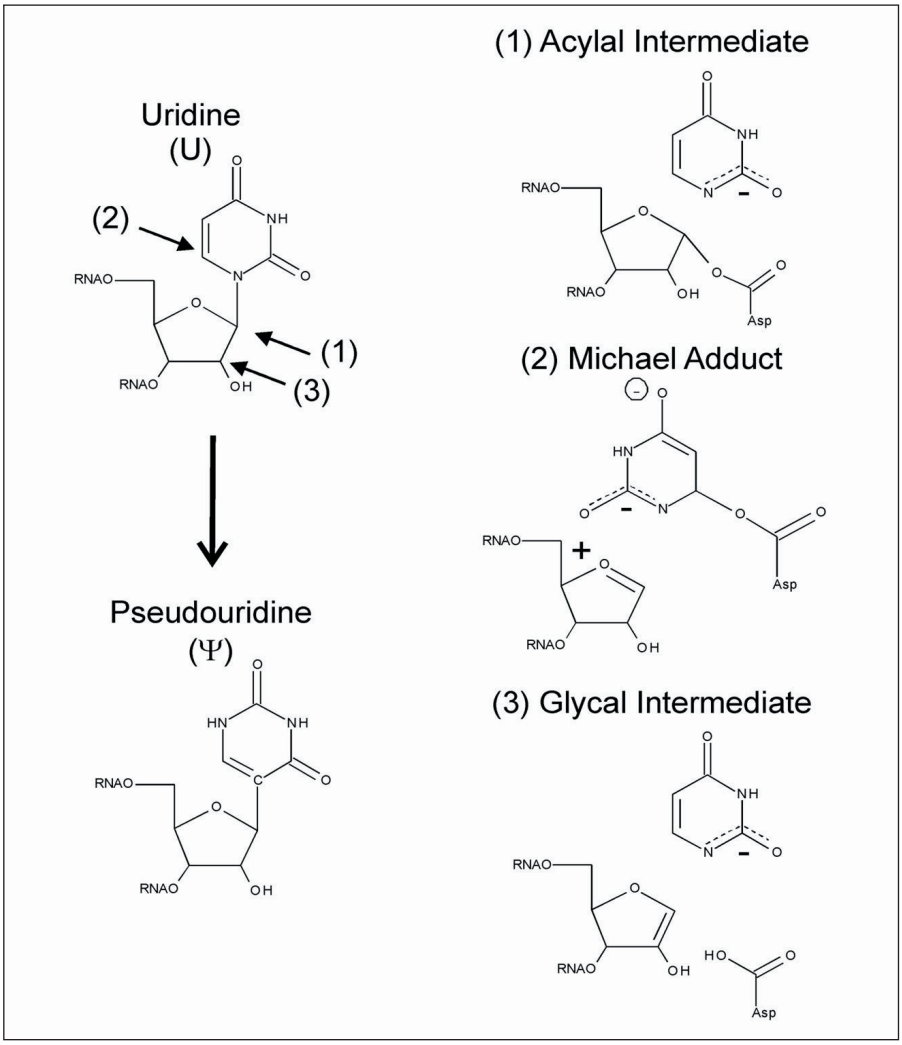

Figure 1: Structures of uridine and pseudouridine together with potential intermediates during the different suggested catalytic mechanisms. The main structural difference between uridine and pseudouridine is the conversion of an N-C to a C-C glycosidic bond resulting in the presence of an additional imino group in the aromatic base in pseudouridine. During the three proposed catalytic mechanisms, the catalytic aspartate either attacks C'1 of the ribose generating an acylal intermediate (1), the C6 of the uracil base forming a Michael adduct (2), or the C2' of the ribose resulting in a glycal intermediate (3).

H/ACA small nucleolar ribonucleoproteins as pseudouridine synthases

The most interesting and only essential eukaryotic pseudouridine synthase is the enzyme Cbf5, which belongs to the TruB family, and is also found in archaea. This is the only pseudouridine synthase that does not act alone, but functions in complex with three other proteins, Nop10, Gar1, and Nhp2, as well as with an H/ACA guide RNA, forming an H/ACA small nucleolar ribonucleoprotein (snoRNP) complex (Fig. 2B and C).[22, 23] The clear advantage of these H/ACA snoRNPs over stand-alone pseudouridine synthases is the greatly expanded diversity in recognizing target RNA sites. By applying a "division of labor" approach, the Cbf5 enzyme is

Familiy	Prokaryotic Enzyme	Target site (*E. coli*)	Eukaryotic enzyme	Target site (*S. cerevisiae*)
TruA			Pus1[60]	tRNA U27/28/34/36 U2snRNA U26/27/28/34/36/65/67
			Pus2[61]	tRNA U27/28
	TruA[62]	tRNA U38/39/40	Pus3[63]	tRNA U38/39
TruB	TruB[64]	tRNA U55	Pus4[65]	tRNA U55
			Cbf5[24]	many – with help of guide RNA
RsuA	RsuA[66]	16S rRNA U516		
	RluB[67]	23S rRNA U2605		
	RluE[67]	23S rRNA U2457		
	RluF[67]	23S rRNA U2504		
RluA	RluA[68]	23S rRNA U746	Pus5[69]	21S rRNA U2819 (mitochondria)
		tRNA U32	Pus6[70]	tRNA U31
	RluC[71]	23S rRNA U955/ U2504/2580	Pus8[72]	tRNA U32 (cytoplasmic)
	RluD[73]	23S rRNA U1911/ U1915/1917	Pus9[72]	tRNA U32 (mitochondrial)
	TruC[67]	tRNA U65		
TruD	TruD[74]	tRNA U13	Pus7[75]	U2snRNA U35, tRNA U13, pre-tRNA^{Tyr} U35, 5S rRNA U50
Pus10	none		Pus10[a 76]	tRNA U54/55 (archaea)

[a] Pus10 is not found in *S. cerevisiae*, but in humans and other eukaryotes; the target site has only been characterized in archaea.

Table 1. Pseudouridine synthases and their target sites in bacteria and eukaryotes.

merely responsible for catalyzing the modification[24] while it is the H/ACA guide RNA that specifically recruits target RNA into the complex for modification.[25] There are many different H/ACA guide RNAs (more than 100 in humans) that allow the H/ACA snoRNPs to modify a large number of uridines in a site-specific manner. Interestingly, for several H/ACA guide RNAs no target sites have been identified so far. Eukaryotic H/ACA guide RNAs are characterized by having two stem-loop structures that are separated by a characteristic ANANNA Hinge region and by containing a 3'ACA sequence giving rise to their naming as H/ACA guide RNAs (Fig. 2B). Within each stem-loop structure is an unpaired region called pseudouridylation pocket; both the 5' and the 3' side of the pseudouridylation pocket base-pair with the target site bulging out the target uridine and a flanking residue thereby allowing the target uridine to bind to the catalytic pocket of Cbf5.[26] The last decade has greatly contributed to our understanding of H/ACA snoRNP structure and function, in particular based on biochemical studies of archaeal ribonucleoprotein complexes. A number of detailed reviews on the complex biogenesis, function and structure of H/ACA snoRNPs have been published in the last years.[27, 28]

Tackling the catalytic mechanism of pseudouridine synthases

Despite the wealth of biochemical information on pseudouridine synthases, the chemical mechanism of pseudouridylation still remains unsolved. As of now, three different chemical mechanisms have been suggested, but for a long time it was even unclear whether all pseudouridine synthases share the same mechanism or whether different enzyme families might utilize different chemistry. Three arguments now support the general opinion that all pseudouridine synthases apply the same chemical strategy. First, the structural studies have revealed a common catalytic domain including a conserved catalytic pocket with an essential aspartate residue even in only distantly related families such as TruD.[7] This finding also indicates that all pseudouridine synthases were probably derived from a common ancestral enzyme. Second, the interaction of pseudouridine synthases with RNA containing the inhibitor 5-fluorouracil has now been clarified. Originally, it seemed as if only some enzymes form covalent adducts with 5-fluorouracil (e.g. TruA)[29] while others are capable of converting this substrate to a 5-fluoro-6-hydroxypseudouridine species (e.g. TruB).[14] However, a careful study by the Mueller group has now revealed that all enzymes seem to equally react with 5-fluorouracil thus eliminating the suggested difference between the enzymes.[30] Lastly, our group has conducted a detailed kinetic study of bacterial pseudouridine synthases representing three different families which demonstrated that pseudouridine synthases are characterized by a uniformly slow catalytic step.[31] The surprisingly identical rate constants for pseudouridine formation can be explained by a common, limiting chemical mechanism.

Pseudouridine formation is a rather complex process as it minimally comprises three chemical steps including the cleavage of the N-C glycosidic bond, the rotation of the base within the enzyme's active site and the

re-attachment of the base to the ribose by forming the C-C glycosidic bond found only in pseudouridines. As mentioned above, three different catalytic mechanisms have been suggested for pseudouridine synthases which can be distinguished by the role of the catalytic aspartate; in particular, which site of the target uridine is attacked by this aspartate, and by the nature of the intermediates. Interestingly, the first step during pseudouridine formation, glycosidic bond cleavage, is the same as the reaction catalyzed by uracil-DNA-glycosylases which also employ a catalytic aspartate residue.[32] This similarity was the basis for the first proposed catalytic mechanism suggesting a nucleophilic attack by the aspartate onto the C1' of the ribose resulting in an acylal intermediate (Fig. 1).[33] Alternatively, it was instead hypothesized that the catalytic aspartate residue could form a covalent bond to C6 of the uracil base creating a Michael adduct.[29, 33] A covalent bond between the aspartate and the base could then constitute the axis of rotation for the base to position the C5 next to the C1' of the ribose. Both mechanisms were debated for more than a decade. Most recently however, the Mueller group extended the studies on the exact nature of the 5-fluoro-6-hydroxypseudouridine product and identified two isomeric hydrated products.[34] Importantly, these findings strongly disfavor the mechanism forming a Michael adduct, but are in agreement with a modified mechanism including the suggested acylal intermediate. Alternatively, the authors propose a third catalytic mechanism where the catalytic aspartate abstracts a proton from the C2' position of the ribose creating a glycal intermediate and a free uracil base. Thus, the field is completely open again and further studies are required to verify the catalytic mechanism. Clearly, new experimental approaches are needed to resolve this long-standing question. For example, kinetic isotope studies can reveal whether C2' is closely involved in catalysis; however, such studies are extremely challenging based on the large size of the substrate RNA which has to be specifically labeled with a heavy isotope. Alternatively, it might be possible to trap and isolate reaction intermediates and to characterize their chemical structure, but again this is no easy task. Hence, a half-century after the discovery of pseudouridines, we are still not able to explain how these modifications are formed, but the progress in the past decade gives rise to the hope that this important question might finally be answered in the near future.

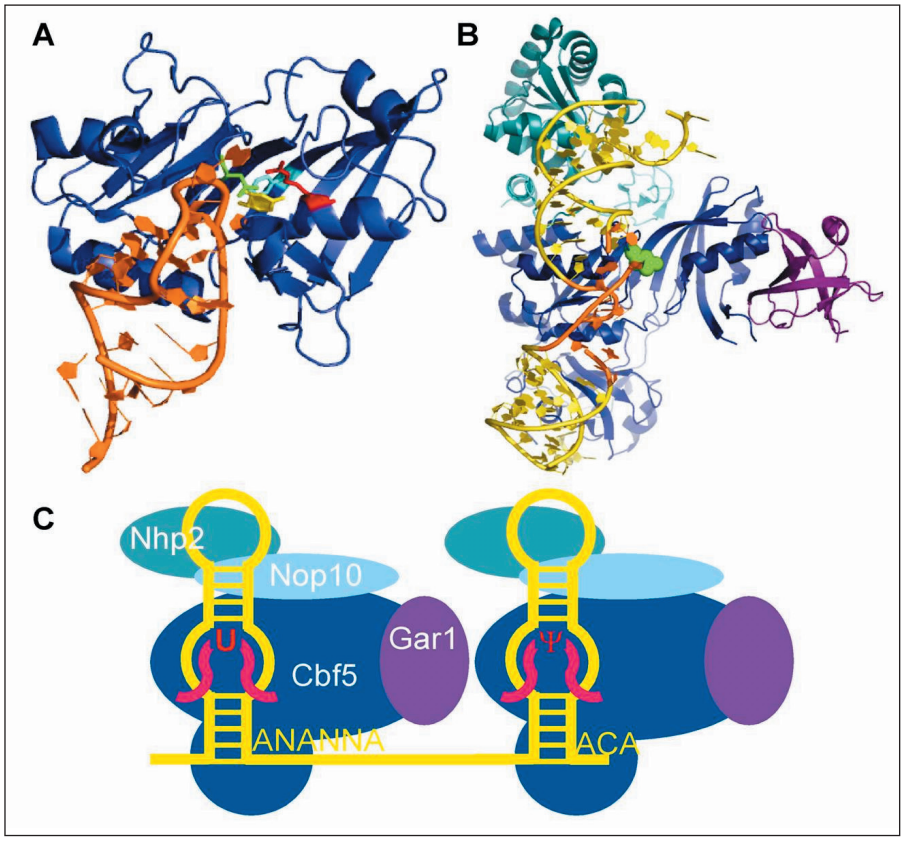

Figure 2: Pseudouridine synthases are either stand-alone enzymes or H/ACA small nucleolar ribonucleoproteins. **A.** *X-ray structure of the E. coli stand-alone pseudouridine synthase RluA[20] (blue) which modifies position 32 (yellow) in the anticodon stem loop (orange) of several tRNAs (PDB ID 2I82). This enzyme consists of only the catalytic domain; the catalytic residues of RluA are shown in stick presentation: catalytic aspartate 64 (green), tyrosine 96 (cyan), and arginine 165 (red).* **B.** *X-ray structure of the Pyrococcus furious H/ACA small ribonucleoprotein[17] (PDB ID 2HVY) containing the proteins Cbf5 (blue) including the catalytic aspartate (green), Gar1 (purple), Nop10 (cyan) and the archaeal homologue of Nhp2 (teal). The H/ACA guide RNA is shown in yellow with the unpaired pseudouridylation pocket in orange where the substrate RNA can bind through base-pairing.* **C.** *Schematic composition of a eukaryotic H/ACA small nucleolar ribonucleoprotein. As eukaryotic H/ACA guide RNA consists of two hairpins connected by a hinge region (ANANNA), these ribonucleoproteins are believed to contain two copies of each protein, Cbf5, Gar1, Nop10 and Nhp2.*

Identifying the cellular functions of pseudouridine synthases

Although pseudouridines are the most common post-transcriptional modification found in RNA across all domains of life, our understanding of the importance of pseudouridines and pseudouridine synthases for the cell still remains fragmented. As most pseudouridines are found in ribosomal RNA, in particular at the functional centers of the ribosome such as the peptidyltransferase center and the interface of the large and small subunit, a role of pseudouridines in ribosome bioge-

nesis and function would be expected. In the last decade, the Fournier lab in particular has analyzed the impact of deleting specific pseudouridines from ribosomal RNA in yeast. For this purpose, single H/ACA guide RNAs were knocked out or mutated to remove a single or a small number of site-specific pseudouridines that are introduced into rRNA by H/ACA small nucleolar ribonucleoproteins. Most interestingly, the lack of a single pseudouridine usually minimally affects the ribosome or the yeast cell. However, the removal of more than three pseudouridines from a single area of the ribosome such as the peptidyltransferase center,[35] helix 69 in 25S rRNA at the interface with the small ribosomal subunit,[5] the A-site finger of the large ribosomal subunit[36], or the decoding center[37], all lead to synergistic effects of varying degrees on cell growth, translational activity and ribosome formation. Interestingly, removal of just a single pseudouridine in many of the mentioned functional regions of the ribosome can also affect the accuracy of translation.[38] Taken together, these findings suggest that many pseudouridines contribute together to the function and formation of the ribosome thereby enhancing cellular fitness, even though the role of individual pseudouridines might be minimal. In accordance with these studies in yeast, three pseudouridine modifications in helix 69 of the bacterial ribosome, that are introduced by the pseudouridine synthase RluD, have also been shown to be important for ribosome function as they influence the interaction of the ribosome with release factor 2 (which is not a homologue of eukaryotic release factors).[39, 40]

Several pseudouridines are also found in spliceosomal RNA, in particular U2 snRNA, where they seem to contribute to spliceosomal function. Almost all uridines in the branch site recognition sequence of U2 snRNA are modified to pseudouridines. Conversion of these sites to 5-fluorouridines, that cannot be changed to pseudouridines, inhibited splicing in a *Xenopus oocyte* reconstitution system.[41] Similarly, the lack of several pseudouridines in the 5' region of U2 snRNA results in cumulative splicing defects, in particular the formation of the spliceosomal E complex.[42] The most important spliceosomal pseudouridine is found at position 35 in the branch site recognition sequence of yeast U2 snRNA. Deletion of the pseudouridine synthase Pus7p that generates this modification leads to growth defects under stress conditions including pre-mRNA accumulation.[6] U2 pseudouridine 35 has a structural effect on branch site recognition sequence and helps to position the branch site adenosine for nucleophilic attack of the 5' splice site.[43] In summary, at least a subset of pseudouridines seem to be important for the function of the ribosome and the spliceosome while other pseudouridines might be less critical for the function of these ribonucleoprotein machines, and may act in a synergistic structural context.

In accordance with the findings that single pseudouridines are dispensable in the cell, most pseudouridine synthases are not essential enzymes with the notable exception of Cbf5[44],that is responsible for the modification of large numbers of uridines by H/ACA small nucleolar ribonucleoproteins. Surprisingly however, yeast cells are viable, but grow poorly, when expressing only an inactive Cbf5 variant, where the catalytic aspartate has been mutated such that *in vivo* pseudouridylation is abolished.[24] This raises the question whether a function of Cbf5 other than its pseudouridylation activity is essential for the cell. Presumably, this essential function of Cbf5 resides in its contribution to pre-rRNA processing. In complex with the snR30 RNA, the only essential H/ACA box RNA in yeast, Cbf5 somehow is required for cleavage of the pre-rRNA transcript while the snR30 H/ACA box RNA seems not to direct pseudouridylation events.[45]

All *E.coli* stand-alone pseudouridine synthases can be knocked out without detectable effects on cell viability. Interestingly, a knock-out strain of *E. coli* lacking TruB (a homologue of Cbf5) grows well on its own but is out-competed by wild-type strains, indicating a contribution of TruB to the fitness of *E. coli* cells.[46] The same finding holds true for an *E.coli* RluA knock-out strain.[19] Strikingly, this phenotype of the TruB knock-out strain can be reversed by expression of a catalytically inactive TruB mutant, further suggesting that an uncharacterized function of this pseudouridine synthase (other than RNA modification) is critical for cellular fitness.[46] This "other" function of TruB might be related to thermal stress tolerance since both a TruB knock-out strain as well as a strain expressing inactive TruBD48C display sensitivity to a temperature shift from 37 to 50°C.[47]

What function of pseudouridine synthases in addition to pseudouridine formation can be important to the cell? An unambiguous answer is still lacking, but it has been hypothesized that these RNA modification enzymes might also act as RNA chaperones that contribute to the folding of their target RNA independent of the chemical modification.[46] For example, many pseudouridine synthases need to gain access to the modification site by changing the structure of their substrate RNA and by flipping nucleobases into their active site (vide supra).[7] This interaction with pseudouridine synthases might help to unfold non-native structures of the RNA and to provide the RNA with a second chance of folding to reach its native state upon dissociation from the pseudouridine synthases. Such an RNA chaperone activity of pseudouridine synthases might not be required under normal growth conditions, but only under certain stress conditions again highlighting that these enzymes contribute to the general fitness of cells without being essential. Similar to this suggestion, it has been demonstrated that knock-out of the *E.coli* pseudouridine synthase RluC, that targets the ribosomal peptidyltransferase

center, renders the cells more susceptible to certain antibiotics, thus showing again a contribution of a pseudouridine synthase to cellular fitness under certain (antibiotic) stress conditions.[48]

Using pseudouridine formation to regulate gene expression

A very new and exciting question in the field of RNA modification is whether pseudouridines are used by the cell or could be used as artificial tools in order to regulate gene expression. The last few years have in fact revealed astonishing findings supporting both suggestions. First, the potential of pseudouridines to contribute to gene regulation in the cell will be reviewed. The best example of such regulation is the inducible pseudouridylation of U2 snRNA which influences pre-mRNA splicing.[49] As described above, several pseudouridines, in particular pseudouridine 35 in U2 snRNA are important for splicing. In addition to these constitutive pseudouridines, the Yu group has identified two inducible pseudouridines in yeast U2 snRNA that are introduced by the stand-alone pseudouridine synthase Pus7p and by H/ACA small nucleolar ribonucleoproteins containing the guide RNA snR81, respectively. These pseudouridines are only found under nutrient deprivation or upon heat shock. At least one of these pseudouridines (Ψ_{93} introduced by snR81 snoRNPs) influences pre-mRNA splicing.[49]

The second example of cellular regulation of gene expression is the pseudouridylation of steroid receptor RNA activator. Here, the stand-alone pseudouridine synthase Pus1 modifies steroid receptor RNA activator and acts thereby as co-activator for nuclear receptors such as the retinoic acid receptor by forming a steroid-independent trimeric complex of Pus1, steroid receptor RNA activator and retinoic acid receptor.[50] The physiological significance of this co-activation is currently unclear.

Finally, recent studies have revealed that the pseudouridylation machinery is linked to microRNAs that are major regulators of gene expression. Strikingly, H/ACA small nucleolar RNAs, which normally act as part of H/ACA small nucleolar ribonucleoproteins, are sometimes also precursors of microRNAs.[51, 52] Hence, some of these RNAs have a dual functionality and can act both as guide RNAs directing pseudouridylation as well as precursors for well-established microRNAs which regulate gene expression. The evolutionary link between snoRNAs and microRNAs is an interesting question that warrants further investigation.[53]

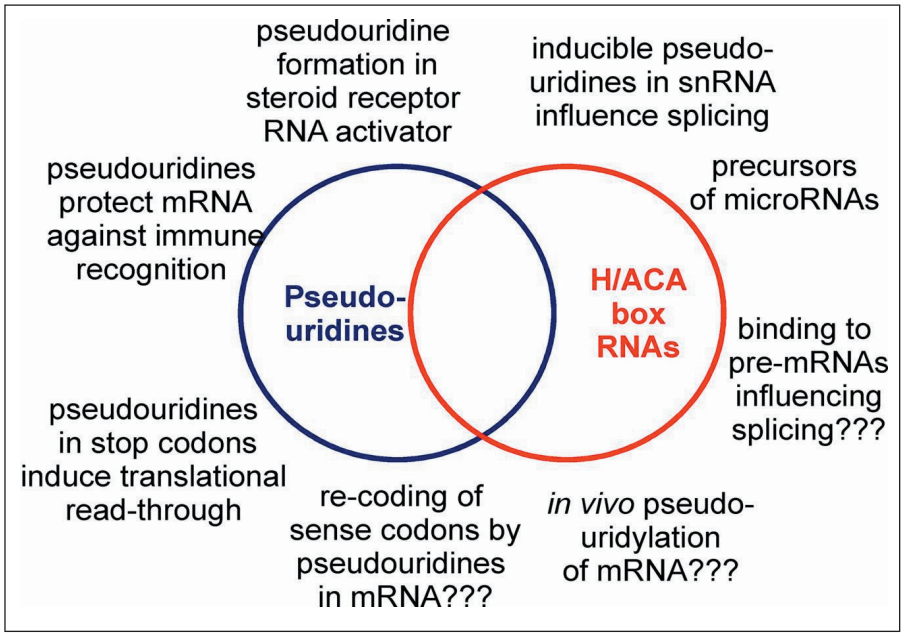

Figure 3: Regulation of gene expression by pseudouridines and H/ACA box RNAs. *Several reports have now revealed how pseudouridine formation in snRNA, mRNA and steroid receptor RNA activator can influence gene expression (see main text for details). Furthermore, H/ACA box RNAs can be precursors of microRNAs independent of their function in directing pseudouridylation. Other mechanisms of regulating gene expression by pseudouridines or the pseudouridylation machinery can be envisioned and might be discovered in the future (indicated by question marks).*

Beyond the described cases of regulation of gene expression by the pseudouridylation machinery, there may be many more mechanisms. For example, C/D box small nucleolar guide RNAs target post-transcriptional 2'O-ribose methylation of RNA and therefore act similarly to H/ACA snoRNAs as guide RNAs. In addition, they have also been shown to have other functions, such as interacting with pre-mRNAs to influence alternative splicing patterns.[54, 55] As there are also numerous H/ACA snoRNAs for which no target pseudouridine site has been identified yet; it might be possible that similar regulatory mechanisms exist for H/ACA guide RNAs. So far, no pseudouridines have been detected in cellular mRNA, but this can be attributed to the low abundance of some of these RNAs and the technical difficulty of detecting pseudouridines. As these modified nucleotides base-pair to adenine like uridines and also have the same molecular mass, high-throughput detection of pseudouridines by reverse-transcriptase based sequencing strategies or mass spectrometry of cellular RNA is currently not possible.

With our increasing understanding of pseudouridine formation, it now becomes possible to use this knowledge as a tool to manipulate cells. First, it has been recognized that incorporation of pseudouridines into mRNA by *in vitro* transcription renders such mRNAs more stable and thus translationally

active upon delivery to cells.[56] This effect stems from the inability of pseudouridine-containing mRNAs to activate 2'-5'-oligoadenylate synthetase and subsequently RNase L. Furthermore, pseudouridine-containing mRNAs are resistant to degradation by RNase L.[57] Thus, incorporating pseudouridines into mRNAs constitutes a mechanism to overcome immune recognition and in the future may allow using such mRNAs in gene replacement and vaccination.

Second and very impressively, the Yu group recently demonstrated that directed pseudouridylation of stop codons in mRNA results in translational read-through, thus opening the exciting possibility of using site-directed pseudouridylation for targeting aberrant mRNAs produced in genetic diseases that are based on the occurrence of premature stop codons.[58] The authors made use of the conserved secondary structure of H/ACA guide RNAs to design novel guide RNAs that base-pair to mRNA nucleotides flanking the target uridine residue at the first position of the stop codon, thereby specifying this residue for modification by H/ACA small nucleolar ribonucleoproteins. For all three stop codons, this strategy resulted in significant read-through of pre-mature stop codons both during *in vitro* translation as well as in *S. cerevisiae*; this effect was enhanced if the nonsense-mediated decay pathway was inactivated. Hence, disease-specific guide RNAs might be designed in the future and delivered to affected cells in order to use intrinsic H/ACA small nucleolar ribonucleoproteins to modify mutated mRNA and to thereby overcome premature termination in these patients. Surprisingly, the pseudouridine-containing codons are read by very specific tRNAs resulting in the incorporation of not more than two different amino acids at the corresponding position in the encoded polypeptides.[58] This altered decoding potential by the ribosome was entirely unexpected and may allow expansion of the genetic code. Future studies will reveal whether the modification of sense codons also likely leads to the incorporation of specific, different amino acids and would therefore open even more possibilities for developing alternatives and additions to the genetic code.[59]

Conclusions

Although pseudouridines have been discovered more than half a century ago, the last decade has revealed remarkable features of pseudouridines and pseudouridine synthases thereby revolutionizing our understanding of this post-transcriptional RNA modification. We are only beginning to comprehend the molecular mechanism of target recognition and chemical conversion of uridines by pseudouridine synthases which are catalyzing a remarkably complex chemical reaction. It seems that the foundation has now been laid to address these questions in the future. Even more complex is the biological function of pseudouridines which have synergistic effects on ribosome and spliceosome function and contribute the cellular fitness under certain stress conditions. Very surprisingly, the actual chemical conversion of uridines to pseudouridines might have secondary importance to the presence of the pseudouridine synthases themselves. These enzymes may have additional cellular functions such as acting as RNA chaperones to enhance the folding of cellular RNAs. This intriguing hypothesis certainly needs further investigation to clearly identify the cellular role of pseudouridine synthases. Lastly, a number of very recent studies suggest that pseudouridine formation may be regulated in certain cases in the cell and may thus contribute to the regulation of gene expression under stress conditions or in specific tissues. Also, H/ACA guide RNAs can have dual functionality as they can be precursors of microRNAs, but it is not clear under which conditions a guide RNA is further processed to a microRNA. A better understanding of such regulatory functions may significantly enhance our knowledge of pseudouridine function in health and disease in the future. As a first glimpse of how to use pseudouridine formation to influence cells, an amazing new possibility of changing the genetic code by introducing pseudouridines into mRNAs has recently emerged, which has large potential in the treatment of inherited diseases as well as in engineering novel polypeptides using an expanded genetic code. After more than 50 years of research, the field of pseudouridine research seems more exciting and promising than ever before.

Acknowledgements

Research in the Kothe lab is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Foundation for Innovation (CFI), and Alberta Innovates Health Solutions (AI-HS).

References

1. Davis FF, Allen FW. Ribonucleic acids from yeast which contain a fifth nucleotide. *J Biol Chem* 1957; **227**:907-15.
2. Cohn WE. Pseudouridine, a carbon-carbon linked ribonucleoside in ribonucleic acids: isolation, structure, and chemical characteristics. *J Biol Chem* 1960; **235**:1488-98.
3. Charette M, Gray MW. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* 2000; **49**:341-51.
4. Arnez JG, Steitz TA. Crystal structure of unmodified tRNA(Gln) complexed with glutaminyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry* 1994; **33**:7560-7.
5. Liang XH, Liu Q, Fournier MJ. rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Molecular Cell* 2007; **28**:965-77.
6. Yang C, McPheeters DS, Yu YT. Psi35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in Saccharomyces cerevisiae. *J Biol Chem* 2005; **280**:6655-62.
7. Hamma T, Ferre-D'Amare AR. Pseudouridine synthases. *Chem Biol* 2006; **13**:1125-35.
8. Hoang C, Ferre-D'Amare AR. Crystal structure of the highly divergent pseudouridine synthase TruD reveals a circular permutation of a conserved fold. *RNA* 2004; **10**:1026-33.
9. Ericsson UB, Nordlund P, Hallberg BM. X-ray structure of tRNA pseudouridine synthase TruD reveals an inserted domain with a novel fold. *FEBS Lett* 2004; **565**:59-64.
10. Kaya Y, Del Campo M, Ofengand J, Malhotra A. Crystal structure of TruD, a novel pseudouridine synthase with a new protein fold. *J Biol Chem* 2004; **279**:18107-10.
11. McCleverty CJ, Hornsby M, Spraggon G, Kreusch A. Crystal structure of human Pus10, a novel pseudouridine synthase. *Journal of Molecular Biology* 2007; **373**:1243-54.
12. Watanabe Y, Gray MW. Evolutionary appearance of genes encoding proteins associated with box H/ACA snoRNAs: cbf5p in Euglena gracilis, an early diverging eukaryote, and candidate Gar1p and Nop10p homologs in archaebacteria. *Nucleic Acids Research* 2000; **28**:2342-52.

13. Phannachet K, Elias Y, Huang RH. Dissecting the roles of a strictly conserved tyrosine in substrate recognition and catalysis by pseudouridine 55 synthase. *Biochemistry* 2005; **44**:15488-94.
14. Hoang C, Ferre-D'Amare AR. Cocrystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell* 2001; **107**:929-39.
15. Sivaraman J, Sauve V, Larocque R, Stura EA, Schrag JD, Cygler M, Matte A. Structure of the 16S rRNA pseudouridine synthase RsuA bound to uracil and UMP. *Nat Struct Biol* 2002; **9**:353-8.
16. Pan H, Agarwalla S, Moustakas DT, Finer-Moore J, Stroud RM. Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit. *Proc Natl Acad Sci U S A* 2003; **100**:12648-53.
17. Li L, Ye K. Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature* 2006; **443**:302-7.
18. Hur S, Stroud RM. How U38, 39, and 40 of many tRNAs become the targets for pseudouridylation by TruA. *Molecular Cell* 2007; **26**:189-203.
19. Raychaudhuri S, Niu L, Conrad J, Lane BG, Ofengand J. Functional effect of deletion and mutation of the Escherichia coli ribosomal RNA and tRNA pseudouridine synthase RluA. *J Biol Chem* 1999; **274**:18880-6.
20. Hoang C, Chen J, Vizthum CA, Kandel JM, Hamilton CS, Mueller EG, Ferre-D'Amare AR. Crystal structure of pseudouridine synthase RluA: indirect sequence readout through protein-induced RNA structure. *Molecular Cell* 2006; **24**:535-45.
21. Alian A, DeGiovanni A, Griner SL, Finer-Moore JS, Stroud RM. Crystal structure of an RluF-RNA complex: a base-pair rearrangement is the key to selectivity of RluF for U2604 of the ribosome. *Journal of Molecular Biology* 2009; **388**:785-800.
22. Ganot P, Bortolin ML, Kiss T. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* 1997; **89**:799-809.
23. Ni J, Tien AL, Fournier MJ. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell* 1997; **89**:565-73.
24. Zebarjadian Y, King T, Fournier MJ, Clarke L, Carbon J. Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA. *Mol Cell Biol* 1999; **19**:7461-72.
25. Ganot P, Caizergues-Ferrer M, Kiss T. The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev* 1997; **11**:941-56.
26. Duan J, Li L, Lu J, Wang W, Ye K. Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Molecular Cell* 2009; **34**:427-39.
27. Hamma T, Ferre-D'Amare AR. The box H/ACA ribonucleoprotein complex: interplay of RNA and protein structures in post-transcriptional RNA modification. *J Biol Chem* 2010; **285**:805-9.
28. Kiss T, Fayet-Lebaron E, Jady BE. Box H/ACA small ribonucleoproteins. *Molecular Cell* 2010; **37**:597-606.
29. Gu X, Liu Y, Santi DV. The mechanism of pseudouridine synthase I as deduced from its interaction with 5-fluorouracil-tRNA. *Proc Natl Acad Sci U S A* 1999; **96**:14270-5.
30. McDonald MK, Miracco EJ, Chen J, Xie Y, Mueller EG. The Handling of the Mechanistic Probe 5-Fluorouridine by the Pseudouridine Synthase TruA and Its Consistency with the Handling of the Same Probe by the Pseudouridine Synthases TruB and RluA. *Biochemistry* 2010; **50**:426-36.
31. Wright JR, Keffer-Wilkes LC, Dobing SR, Kothe U. Pre-steady-state kinetic analysis of the three Escherichia coli pseudouridine synthases TruB, TruA, and RluA reveals uniformly slow catalysis. *RNA* 2011; **17**:2074-84.
32. Stivers JT, Jiang YL. A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chemical Reviews* 2003; **103**:2729-59.
33. Huang L, Pookanjanatavip M, Gu X, Santi DV. A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. *Biochemistry* 1998; **37**:344-51.
34. Miracco EJ, Mueller EG. The products of 5-fluorouridine by the action of the pseudouridine synthase TruB disfavor one mechanism and suggest another. *J Am Chem Soc* 2011; **133**:11826-9.
35. King TH, Liu B, McCully RR, Fournier MJ. Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Molecular Cell* 2003; **11**:425-35.
36. Liang XH, Liu Q, Fournier MJ. Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA* 2009; **15**:1716-28.
37. Piekna-Przybylska D, Przybylski P, Baudin-Baillieu A, Rousset JP, Fournier MJ. Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit. *J Biol Chem* 2008; **283**:26026-36.
38. Baudin-Baillieu A, Fabret C, Liang XH, Piekna-Przybylska D, Fournier MJ, Rousset JP. Nucleotide modifications in three functionally important regions of the Saccharomyces cerevisiae ribosome affect translation accuracy. *Nucleic Acids Research* 2009; **37**:7665-77.
39. Ejby M, Sorensen MA, Pedersen S. Pseudouridylation of helix 69 of 23S rRNA is necessary for an effective translation termination. *Proc Natl Acad Sci U S A* 2007; **104**:19410-5.
40. Kipper K, Sild S, Hetenyi C, Remme J, Liiv A. Pseudouridylation of 23S rRNA helix 69 promotes peptide release by release factor RF2 but not by release factor RF1. *Biochimie* 2011; **93**:834-44.
41. Zhao X, Yu YT. Pseudouridines in and near the branch site recognition region of U2 snRNA are required for snRNP biogenesis and pre-mRNA splicing in Xenopus oocytes. *RNA* 2004; **10**:681-90.
42. Donmez G, Hartmuth K, Luhrmann R. Modified nucleotides at the 5' end of human U2 snRNA are required for spliceosomal E-complex formation. *RNA* 2004; **10**:1925-33.
43. Newby MI, Greenbaum NL. Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat Struct Biol* 2002; **9**:958-65.
44. Jiang W, Middleton K, Yoon HJ, Fouquet C, Carbon J. An essential yeast protein, CBF5p, binds in vitro to centromeres and microtubules. *Mol Cell Biol* 1993; **13**:4884-93.
45. Eliceiri GL. The vertebrate E1/U17 small nucleolar ribonucleoprotein particle. *Journal of cellular biochemistry* 2006; **98**:486-95.
46. Gutgsell N, Englund N, Niu L, Kaya Y, Lane BG, Ofengand J. Deletion of the Escherichia coli pseudouridine synthase gene truB blocks formation of pseudouridine 55 in tRNA in vivo, does not affect exponential growth, but confers a strong selective disadvantage in competition with wild-type cells. *RNA* 2000; **6**:1870-81.
47. Kinghorn SM, O'Byrne CP, Booth IR, Stansfield I. Physiological analysis of the role of truB in Escherichia coli: a role for tRNA modification in extreme temperature resistance. *Microbiology* 2002; **148**:3511-20.
48. Toh SM, Mankin AS. An indigenous posttranscriptional modification in the ribosomal peptidyl transferase center confers resistance to an array of protein synthesis inhibitors. *Journal of Molecular Biology* 2008; **380**:593-7.
49. Wu G, Xiao M, Yang C, Yu YT. U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *The EMBO Journal* 2011; **30**:79-89.
50. Zhao X, Patton JR, Davis SL, Florence B, Ames SJ, Spanjaard RA. Regulation of nuclear receptor activity by a pseudouridine synthase through posttranscriptional modification of steroid receptor RNA activator. *Molecular Cell* 2004; **15**:549-58.
51. Scott MS, Avolio F, Ono M, Lamond AI, Barton GJ. Human miRNA precursors with box H/ACA snoRNA features. *PLoS Computational Biology* 2009; **5**:e1000507.
52. Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N, Meister G. A human snoRNA with microRNA-like functions. *Molecular Cell* 2008; **32**:519-28.
53. Scott MS, Ono M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie* 2011; **93**:1987-92.
54. Kishore S, Stamm S. The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science* 2006; **311**:230-2.
55. Bazeley PS, Shepelev V, Talebizadeh Z, Butler MG, Fedorova L, Filatov V, Fedorov A. snoTARGET shows that human orphan snoRNA targets locate close to alternative splice junctions. *Gene* 2008; **408**:172-9.
56. Kariko K, Muramatsu H, Welsh FA, Ludwig J, Kato H, Akira S, Weissman D. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Molecular Therapy: the journal of the American Society of Gene Therapy* 2008; **16**:1833-40.
57. Anderson BR, Muramatsu H, Jha BK, Silverman RH, Weissman D, Kariko K. Nucleoside modifications in RNA limit activation of 2'-5'-oligoadenylate synthetase and increase resistance to cleavage by RNase L. *Nucleic Acids Research* 2011; **39**:9329-38.
58. Karijolich J, Yu YT. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* 2011; **474**:395-8.
59. Parisien M, Yi C, Pan T. Rationalization and prediction of selective decoding of pseudouridine-modified nonsense and sense codons. *RNA* 2012; **18**:355-67.
60. Simos G, Tekotte H, Grosjean H, Segref A, Sharma K, Tollervey D, Hurt EC. Nuclear pore proteins are involved in the biogenesis of functional tRNA. *The EMBO Journal* 1996; **15**:2270-84.
61. Behm-Ansmant I, Branlant C, Motorin Y. The Saccharomyces cerevisiae Pus2 protein encoded by YGL063w ORF is a mitochondrial tRNA:Psi27/28-synthase. *RNA* 2007; **13**:1641-7.
62. Turnbough CL, Jr., Neill RJ, Landsberg R, Ames BN. Pseudouridylation of tRNAs and its role in regulation in Salmonella typhimurium. *J Biol Chem* 1979; **254**:5111-9.
63. Lecointe F, Simos G, Sauer A, Hurt EC, Motorin Y, Grosjean H. Characterization of yeast protein Deg1 as pseudouridine synthase (Pus3) catalyzing the formation of psi 38 and psi 39 in tRNA anticodon loop. *J Biol Chem* 1998; **273**:1316-23.
64. Nurse K, Wrzesinski J, Bakin A, Lane BG, Ofengand J. Purification, cloning, and properties of the tRNA psi 55 synthase from Escherichia coli. *RNA* 1995; **1**:102-12.
65. Becker HF, Motorin Y, Planta RJ, Grosjean H. The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of psi55 in both mitochondrial and cytoplasmic tRNAs. *Nucleic Acids Research* 1997; **25**:4493-9.
66. Wrzesinski J, Bakin A, Nurse K, Lane BG, Ofengand J. Purification, cloning, and properties of the 16S RNA pseudouridine 516 synthase from Escherichia coli. *Biochemistry* 1995; **34**:8904-13.

67. Del Campo M, Kaya Y, Ofengand J. Identification and site of action of the remaining four putative pseudouridine synthases in Escherichia coli. *RNA* 2001; **7**:1603-15.
68. Wrzesinski J, Nurse K, Bakin A, Lane BG, Ofengand J. A dual-specificity pseudouridine synthase: an Escherichia coli synthase purified and cloned on the basis of its specificity for psi 746 in 23S RNA is also specific for psi 32 in tRNA(phe). *RNA* 1995; **1**:437-48.
69. Ansmant I, Massenet S, Grosjean H, Motorin Y, Branlant C. Identification of the Saccharomyces cerevisiae RNA:pseudouridine synthase responsible for formation of psi(2819) in 21S mitochondrial ribosomal RNA. *Nucleic Acids Research* 2000; **28**:1941-6.
70. Ansmant I, Motorin Y, Massenet S, Grosjean H, Branlant C. Identification and characterization of the tRNA:Psi 31-synthase (Pus6p) of Saccharomyces cerevisiae. *J Biol Chem* 2001; **276**:34934-40.
71. Conrad J, Sun D, Englund N, Ofengand J. The rluC gene of Escherichia coli codes for a pseudouridine synthase that is solely responsible for synthesis of pseudouridine at positions 955, 2504, and 2580 in 23 S ribosomal RNA. *J Biol Chem* 1998; **273**:18562-6.
72. Behm-Ansmant I, Grosjean H, Massenet S, Motorin Y, Branlant C. Pseudouridylation at position 32 of mitochondrial and cytoplasmic tRNAs requires two distinct enzymes in Saccharomyces cerevisiae. *J Biol Chem* 2004; **279**:52998-3006.
73. Raychaudhuri S, Conrad J, Hall BG, Ofengand J. A pseudouridine synthase required for the formation of two universally conserved pseudouridines in ribosomal RNA is essential for normal growth of Escherichia coli. *RNA* 1998; **4**:1407-17.
74. Kaya Y, Ofengand J. A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. *RNA* 2003; **9**:711-21.
75. Behm-Ansmant I, Urban A, Ma X, Yu YT, Motorin Y, Branlant C. The Saccharomyces cerevisiae U2 snRNA:pseudouridine-synthase Pus7p is a novel multisite-multisubstrate RNA:Psi-synthase also acting on tRNAs. *RNA* 2003; **9**:1371-82.
76. Roovers M, Hale C, Tricot C, Terns MP, Terns RM, Grosjean H, Droogmans L. Formation of the conserved pseudouridine at position 55 in archaeal tRNA. *Nucleic Acids Research* 2006; **34**:4293-301.

Dr. Ute Kothe

Assistant Professor
Department of Chemistry & Biochemistry
Alberta RNA Research and Training Institute
University of Lethbridge
4401 University Drive
Lethbridge, Alberta
T1K 3M4
Canada
Email: ute.kothe@uleth.ca
Web: http://people.uleth.ca/~ute.kothe/

Ute Kothe studied Biochemistry at the universities in Regensburg, Bochum and Witten. From 2002 to 2006 she completed her Ph.D. studies in physical biochemistry under the supervision of Dr. Marina Rodnina at the University of Witten/Herdecke where she investigated the kinetics and accuracy of bacterial protein synthesis. Subsequently, Ute Kothe accepted a position as Assistant Professor at the University of Lethbridge in Alberta, Canada in 2006 where she is now tenured. Her research group investigates the early steps of ribosome biogenesis with a special focus on RNA modification and folding. By using model systems ranging from bacteria, to archaea and yeast as well as a combination of molecular biology, genetics, biochemistry and biophysics, in particular kinetics and fluorescence, Ute Kothe's research currently addresses the molecular mechanisms underlying RNA-protein interactions during modification and RNA folding. Ultimately, these investigations will not only increase our fundamental knowledge on how the cells builds complex ribonucleoproteins, but will also aim at identifying novel targets in the treatment of cancer, other proliferative diseases and inherited diseases which all affect ribosome biogenesis.