RNA Folding

DOI: 10.1002/anie.200502928

Modulation of RNA Tertiary Folding by Incorporation of Caged Nucleotides**

*Claudia Höbartner and Scott K. Silverman**

Photochemical caging groups[1] are often used to study biomolecular interactions, such as those between cofactors and enzymes or between neurotransmitters and their receptors.[2] In a caging experiment, a functionally important part of a molecule is modified by the strategic introduction of a non-natural moiety that disrupts key contacts or interferes with a critical reaction step. When the non-natural moiety is removed by photolysis, the native biomolecular interactions are restored. Photochemical caging has been used to study protein folding,[3] RNA catalysis,[4] secondary structure changes in RNA,[5] and other RNA-related phenomena,[6] but not higher-order tertiary folding of RNA.[7] When a caging group is appended at any specific nucleotide position, the effect on RNA tertiary folding is difficult to predict. This is unlike the situation of RNA catalysis (for which caging of an active-site hydroxy-group nucleophile is certain to disrupt reactivity) or of RNA secondary structure changes (for which reliable Watson–Crick interactions provide a strong predictive tool). Herein we report how an RNA molecule with a well-defined tertiary structure responds to the introduction of individual caging groups. Our results, which demonstrate varying local and global structural effects, have several implications for ongoing efforts to understand RNA folding.

We appended the photochemically removable (*S*)-1-(2-nitrophenyl)ethyl (NPE) group onto the nucleobase of each of the four standard RNA nucleotides (Figure 1).[8] The four corresponding phosphoramidites (NPE-U, NPE-C, NPE-A, and NPE-G) were prepared for conventional solid-phase RNA synthesis by using the 2′-*O*-triisopropylsilyloxymethyl (TOM) methodology.[9] The preparative route for the caged guanosine phosphoramidite was based on a recent report,[5] and the routes for the other three phosphoramidites are detailed in the Supporting Information. For each of the four phosphoramidites, solid-phase synthesis was used to prepare

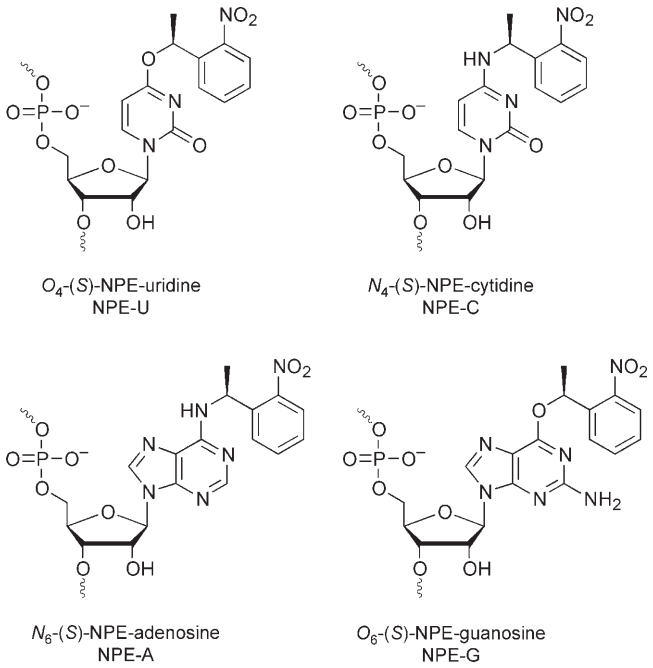

Figure 1. Structures of the (*S*)-NPE-caged RNA nucleotides.

15-mer and 24-mer RNA oligonucleotides, each of which incorporate a single caged RNA nucleotide at a specific internal position. In all cases, denaturing polyacrylamide gel electrophoresis (PAGE), anion-exchange HPLC analysis, and ESI mass spectrometry revealed that decaging proceeds to completion in less than 10 min by using a 300-W Xe-arc-lamp source (300–400 nm; Figure 2) or in approximately 30–60 min by using a handheld UV illuminator (365 nm; data not shown).[10]

We then placed individual caged nucleotides within a large RNA molecule, P4–P6. The 160-nucleotide P4–P6 domain of the *Tetrahymena* group I intron RNA adopts a characteristic tertiary structure in the presence of millimolar concentrations of Mg^{2+}.[11,12] As revealed by X-ray crystallography, the folded domain adopts a three-dimensional structure that resembles a candy cane (Figure 3).[13,14] The two largely helical domains of the RNA are held together, in part, by a key tertiary interaction between the GAAA tetraloop and its 11-nucleotide receptor.[15] Nucleotide and functional-group alterations within this tetraloop–receptor interaction have deleterious effects on the P4–P6 thermodynamic stability.[11,12,15–17]

We sought to probe the effects on P4-P6 folding of caging specific nucleotides with the NPE group. For this purpose, we synthesized modified P4–P6 variants by using a ligation strategy that joins a caged 15-mer or 24-mer oligonucleotide with an unmodified 145-mer or 136-mer T7 RNA polymerase transcript (see Supporting Information). We prepared 14 caged P4–P6 derivatives in this fashion. The caging sites were distributed throughout the RNA and were chosen on the basis of the X-ray crystal structure (Figure 3a). Caging sites were placed in positions where the NPE group was anticipated to have a combination of three effects: 1) alteration of the hydrogen-bonding interactions within either secondary- or

[*] Dr. C. Höbartner, Prof. S. K. Silverman
Department of Chemistry
University of Illinois at Urbana-Champaign
600 South Mathews Avenue, Urbana, IL 61801 (USA)
Fax: (+1) 217-244-8024
E-mail: scott@scs.uiuc.edu

[**] This research was supported by an Erwin Schrödinger fellowship from the Austrian Science Fund to C.H. and by the UIUC Department of Chemistry. S.K.S. is a Fellow of The David and Lucile Packard Foundation. We thank Prof. Ronald Micura (Innsbruck) for use of the HPLC system, Mary Smalley and Chandra Miduturu for assistance with RNA transcriptions, and Kimberly Peterson for preliminary synthetic experiments.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

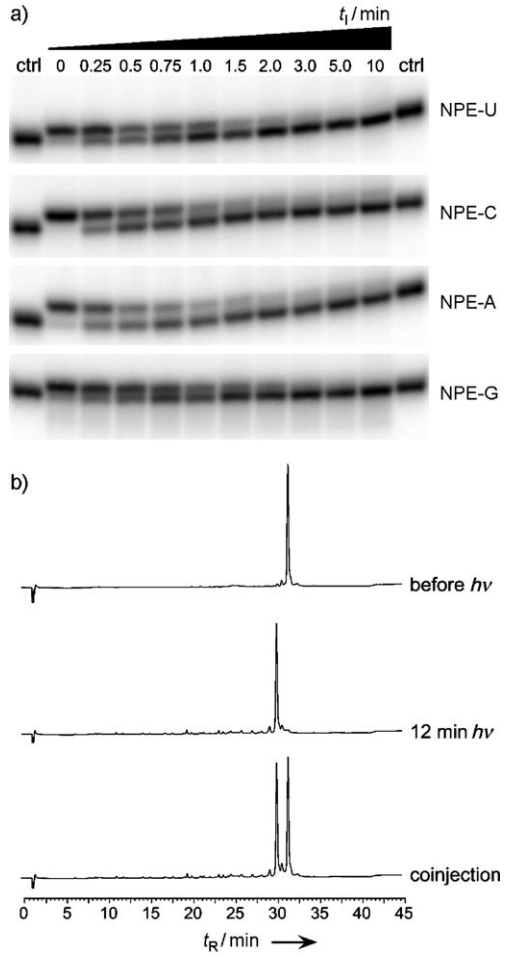

Figure 2. Photodeprotection of caged RNA oligoribonucleotides. a) Denaturing PAGE (20%) shows the photolysis time course for each of four caged 15-mer oligonucleotides (5′-GGAAUUGCGGGAAAG-3′, the underlined positions were individually caged with the NPE group). Control (ctrl) = oligonucleotide of the same sequence without the NPE group. b) Representative anion-exchange HPLC traces for the oligonucleotide with NPE-caged adenosine. Photolysis was carried out with a 300-W Xe arc lamp. (Traces for the other three caged oligonucleotides are shown in the Supporting Information, along with ESIMS data before and after photolysis for all four caged oligonucleotides.)

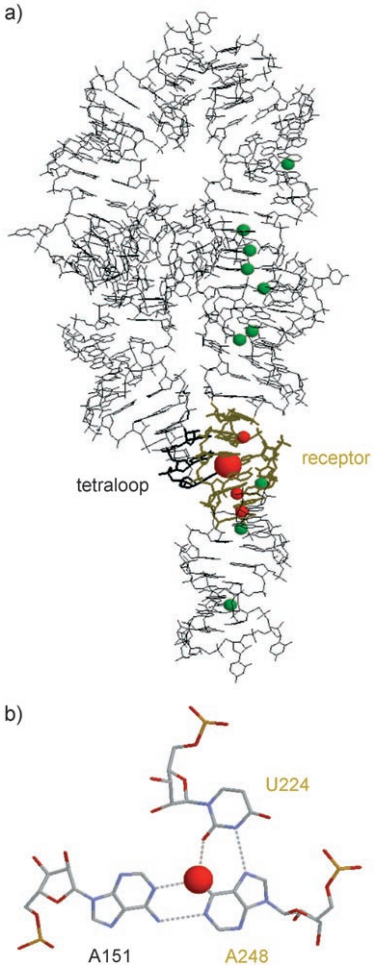

Figure 3. The 160-nucleotide P4–P6 RNA and sites of caging. a) X-ray crystal structure[13] with caging sites marked. Spheres are centered on the nucleobase atoms that bear the NPE caging groups and are colored red or green as in the data plots. The sphere marking the N6 atom of A248 is larger than the other spheres. Adenosine nucleotides A151, A152, and A153 of the GAAA tetraloop are thickened for emphasis. See Supporting Information for a picture of secondary structure. b) The base triple involving A248.

tertiary-structure elements; 2) disruption of base-stacking interactions; and 3) introduction of steric clashes (see Supporting Information for detailed images). In the NPE-A248 P4–P6 derivative, for example, the NPE caging group is placed on an adenine N6 group that is located in a densely packed region of RNA structure and at the interface between the GAAA tetraloop and the tetraloop receptor (Figure 3a). Because A248 is the central nucleotide of a base triple (Figure 3b) and because there appears to be no room to accommodate the bulky NPE group, NPE caging of A248 is expected to disrupt several hydrogen bonds and introduce a steric clash directly within the key tetraloop–receptor tertiary interaction. The determination of the consequences of such disruptions on RNA tertiary folding was a main experimental goal of this study.

The effects of individual caged nucleotides on P4–P6 folding were analyzed by nondenaturing (native) PAGE, which reports on global folding (Figure 4a).[17] The P4–P6 domain has a Mg^{2+}-dependent mobility on native PAGE; that is, when the RNA folds at higher Mg^{2+} concentrations, it adopts a more compact structure and thus migrates faster through the gel. Thermodynamic disruptions to the P4–P6 structure result in an increased Mg^{2+} requirement for folding, as quantified by the Mg^{2+} midpoint ($[Mg^{2+}]_{1/2}$ value) of a mobility versus $[Mg^{2+}]$ titration curve.[15,17,18] Our data allowed all 14 of the caged P4–P6 derivatives to be classified readily into just two types (Figure 4b): 1) those in which the caging group has a substantially destabilizing thermodynamic impact on tertiary folding (red), and 2) those in which the thermodynamic impact is small (green). Regardless of the effect of the caging group, its photolytic removal before analysis by native PAGE restored the RNA folding to that of the unmodified wild-type P4–P6 (Figure 4).

When the differential effects of caging the various nucleotides are considered in the context of their locations

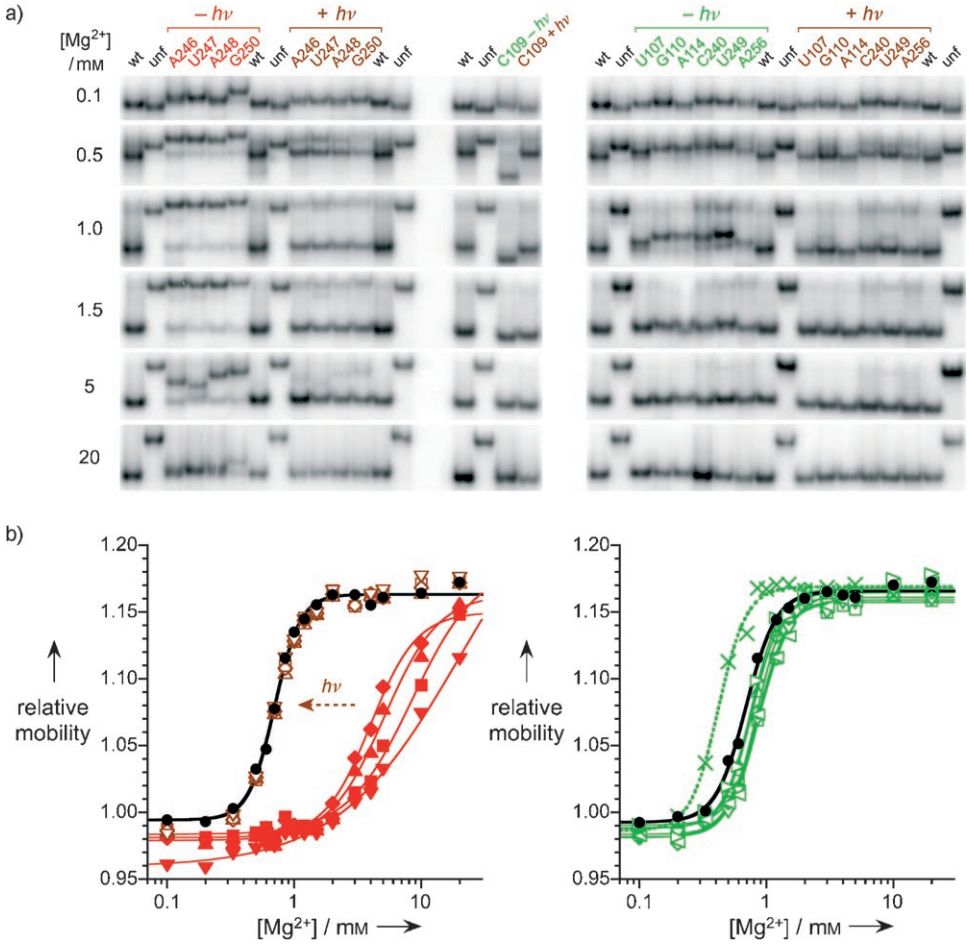

Figure 4. Native PAGE analysis of caged P4–P6 RNAs. a) Representative gel images. wt = unmodified wild-type P4–P6; unf = unfolded control that cannot fold even at high $[Mg^{2+}]$.[11,17,23] b) Mg^{2+} dependence of native gel mobility, computed for each caged derivative relative to the unfolded control. The wild-type P4–P6 is in black (●). Data in red represent positions where an unambiguous thermodynamic disruption is observed (▲ A246; ◆ U247; ■ A248; ▼ G250). Data in green represent positions where little effect is observed (× C109; the remaining symbols represent other positions from panel (a)). Photolysis of samples before native PAGE restored folding to that equivalent to the wild-type P4–P6 (brown). See the Supporting Information for experimental details.

within P4–P6, a clear pattern emerges. Caging of a nucleotide far from the tetraloop–receptor interaction has little effect on global folding. In sharp contrast, caging of a nucleotide that is located within (or very close to) the tetraloop–receptor interaction, such as A248, introduces a distinct thermodynamic destabilization. Nevertheless, native PAGE indicates that each of the latter derivatives eventually adopts the native global tertiary structure when the Mg^{2+} concentration is increased sufficiently (≥ 20 mM). For the NPE-caged A248 derivative, this finding is surprising because the tetraloop–receptor interaction must distort considerably to accommodate the NPE group, although the precise nature of the distortion cannot be discerned from these data. The only exception to the clear pattern is that the NPE-caged U249 derivative is not destabilized even though it is located within the receptor. Significantly, U249 is a bulged nucleotide in which the nucleobase does not form any tertiary hydrogen bonds, and inspection of the X-ray crystal structure suggests that the NPE caging group should be accommodated without causing a steric clash.

To provide an assessment of RNA folding that is independent of the native gel experiments, chemical probing with dimethyl sulfate was also performed. This reagent methylates either N1 of adenine or N3 of cytosine but only if the nucleobase is locally accessible to the reagent.[11,12,19] Tertiary folding of the RNA molecule buries some of the nucleobase functional groups within its interior;[20] this renders these functional groups inaccessible to dimethyl sulfate and thus protects them from methylation. Because the functional groups methylated by dimethyl sulfate are located on the Watson–Crick faces of the nucleobases, methylation inhibits primer extension by reverse transcriptase and leads to an abort band when the RTase reaction assay is examined by denaturing PAGE. The adenosine nucleotides of the GAAA tetraloop in unfolded P4–P6 (for example, P4–P6 in the absence of Mg^{2+}) are methylated readily by dimethyl sulfate. These nucleotides, however, are substantially protected from methylation when P4–P6 is folded properly owing to their sequestration in the tetraloop–receptor interaction.[11,12]

For several of the caged P4–P6 RNAs, dimethyl sulfate probing was performed at varying Mg^{2+} concentrations, and the accessibility of the GAAA tetraloop was monitored (Figure 5a). These experiments provide information on local structure that cannot be determined with native PAGE. Depending on the caging site, the dimethyl sulfate probing results revealed a difference in the accessibilities of the GAAA tetraloop at high Mg^{2+} concentrations (Figure 5b). The tested caging sites that are remote from the tetraloop–receptor interaction (U107, C109) had essentially no effect on Mg^{2+}-dependent accessibility of the tetraloop. In contrast, caging of nucleotides within or near the receptor (A246, U247, A248, and G250) significantly suppressed protection of the tetraloop from dimethyl sulfate modification upon an increase in the Mg^{2+} concentration. As with the native gel assays, U249 located within the receptor provides a predict-

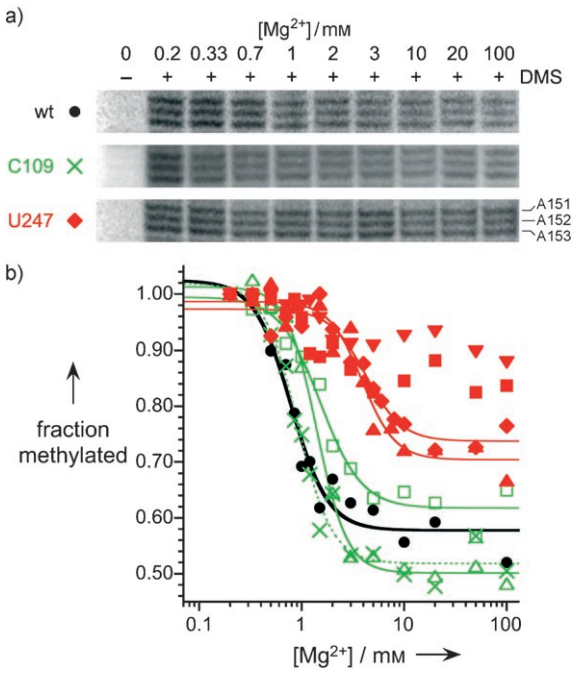

Figure 5. Chemical probing experiments with dimethyl sulfate of caged P4–P6 RNAs. a) Representative gel images for reverse transcriptase (RTase) assays of caged P4–P6 treated with dimethyl sulfate. The presence of an RTase abort band indicates methylation of the corresponding nucleobase with dimethyl sulfate, which must have been accessible to the reagent. The GAAA tetraloop includes nucleotides A151–A153. b) Mg^{2+} dependence of fraction methylated for the tetraloop adenosine nucleotides when various nucleotides are caged (● wild-type RNA; △ U107, × C109; □ U249; ▲ A246; ◆ U247; ■ A248; ▼ G250). A high fraction methylated implies high accessibility of the tetraloop to dimethyl sulfate. Photolysis of the samples before dimethyl sulfate probing restored protection from methylation (see Supporting Information).

able exception in this regard. These data show that the tetraloop–receptor interaction can be substantially perturbed on the local structural level, even when the global structure is essentially native. This conspicuous difference between local and global structural effects due to caging of specific nucleotides increases our understanding of how 'key interactions' contribute to RNA tertiary structure. Our results may also relate to the concept of compact, misfolded RNA states.[21] Some of the caged P4–P6 derivatives such as NPE-caged A248 adopt globally compact structures (as revealed by native PAGE) that nevertheless lack at least one native tertiary interaction (as indicated by dimethyl sulfate probing).

In summary, we have shown herein that nucleotides with caged nucleobases are readily incorporated into large RNA molecules that have complex tertiary structures, and the effects of the caged nucleotides are assessable with appropriate biochemical assays. From our results, we drew two conclusions that relate to the control of RNA folding by using caged nucleotides. First, incorporation of a bulky caging group can thermodynamically disrupt global RNA tertiary structure, but in all tested cases the disruption is relatively modest because it can be overcome by a sufficient increase in the Mg^{2+} concentration. This indicates that even 'key' tertiary contacts, such as the tetraloop–receptor interaction, are not strictly required for RNA tertiary folding. It also suggests that a chemically more sophisticated strategy (e.g., one that involves the attachment of a covalent tether at more than one position) is required to induce qualitative changes in global RNA structure that cannot be recovered simply by an increase in the Mg^{2+} concentration.[22] Second, comparisons of the native PAGE and dimethyl sulfate probing data indicate a clear differentiation between local and global structural effects upon the introduction of a chemical perturbation. This relates to the ability of individual caging groups to induce misfolded RNA states. A future direction for these studies will be to combine individual caging groups or photocleavable covalent tethers with time-resolved decaging to monitor RNA tertiary-folding processes in real time.

Received: August 18, 2005
Published online: October 17, 2005

Keywords: caging · nucleobases · oligonucleotides · photolysis · RNA structures

[1] J. H. Kaplan, B. Forbush III, J. F. Hoffman, *Biochemistry* **1978**, *17*, 1929–1935.
[2] a) H. A. Lester, J. M. Nerbonne, *Annu. Rev. Biophys. Bioeng.* **1982**, *11*, 151–175; b) A. M. Gurney, H. A. Lester, *Physiol. Rev.* **1987**, *67*, 583–617; c) J. A. McCray, D. R. Trentham, *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 239–270; d) S. R. Adams, R. Y. Tsien, *Annu. Rev. Physiol.* **1993**, *55*, 755–784; e) B. E. Cohen, B. L. Stoddard, D. E. Koshland, Jr., *Biochemistry* **1997**, *36*, 9035–9044; f) J. C. Miller, S. K. Silverman, P. M. England, D. A. Dougherty, H. A. Lester, *Neuron* **1998**, *20*, 619–624; g) K. Curley, D. S. Lawrence, *Curr. Opin. Chem. Biol.* **1999**, *3*, 84–88; h) A. P. Pelliccioli, J. Wirz, *Photochem. Photobiol. Sci.* **2002**, *1*, 441–458; i) J. E. T. Corrie in *Dynamic Studies in Biology* (Eds.: M. Goeldner, R. Givens), Wiley-VCH, Weinheim, **2005**, pp. 1–28.
[3] a) H. S. M. Lu, M. Volk, Y. Kholodenko, E. Gooding, R. M. Hochstrasser, W. F. DeGrado, *J. Am. Chem. Soc.* **1997**, *119*, 7173–7180; b) K. C. Hansen, R. S. Rock, R. W. Larsen, S. I. Chan, *J. Am. Chem. Soc.* **2000**, *122*, 11567–11568.
[4] a) S. G. Chaulk, A. M. MacMillan, *Nucleic Acids Res.* **1998**, *26*, 3173–3178; b) S. G. Chaulk, A. M. MacMillan, *Angew. Chem.* **2001**, *113*, 2207–2210; *Angew. Chem. Int. Ed.* **2001**, *40*, 2149–2152.
[5] P. Wenter, B. Fürtig, A. Hainard, H. Schwalbe, S. Pitsch, *Angew. Chem.* **2005**, *117*, 2656–2659; *Angew. Chem. Int. Ed.* **2005**, *44*, 2600–2603.
[6] a) L. Kröck, A. Heckel, *Angew. Chem.* **2005**, *117*, 475–477; *Angew. Chem. Int. Ed.* **2005**, *44*, 471–473; b) S. Shah, S. Rangarajan, S. H. Friedman, *Angew. Chem.* **2005**, *117*, 1352–1356; *Angew. Chem. Int. Ed.* **2005**, *44*, 1328–1332.
[7] a) P. Brion, E. Westhof, *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113–137; b) I. Tinoco, Jr., C. Bustamante, *J. Mol. Biol.* **1999**, *293*, 271–281.
[8] The *S* isomer of the NPE group was chosen to be consistent with the report of Wenter et al.[5]
[9] a) S. Pitsch, P. A. Weiss, L. Jenny, A. Stutz, X. Wu, *Helv. Chim. Acta* **2001**, *84*, 3773–3795; b) R. Micura, *Angew. Chem.* **2002**, *114*, 2369–2373; *Angew. Chem. Int. Ed.* **2002**, *41*, 2265–2269.
[10] The caged uridine derivative with the NPE group at N3 instead of at O4 was also prepared. However, its photocleavage is inefficient (only ≈50% deprotection in 10 min with the arc

lamp) and results in a mixture of two cleavage products, only one of which has an unmodified uracil nucleobase. Therefore, we did not pursue the N_3-NPE-caged uridine derivative any further. See Supporting Information for details.

[11] F. L. Murphy, T. R. Cech, *Biochemistry* **1993**, *32*, 5291 – 5300.
[12] F. L. Murphy, T. R. Cech, *J. Mol. Biol.* **1994**, *236*, 49 – 63.
[13] J. H. Cate, A. R. Gooding, E. Podell, K. Zhou, B. L. Golden, C. E. Kundrot, T. R. Cech, J. A. Doudna, *Science* **1996**, *273*, 1678 – 1685.
[14] K. Juneau, E. Podell, D. J. Harrington, T. R. Cech, *Structure* **2001**, *9*, 221 – 231.
[15] B. T. Young, S. K. Silverman, *Biochemistry* **2002**, *41*, 12271 – 12276.
[16] D. L. Abramovitz, A. M. Pyle, *J. Mol. Biol.* **1997**, *266*, 493 – 506.
[17] S. K. Silverman, T. R. Cech, *Biochemistry* **1999**, *38*, 8691 – 8702.
[18] a) S. K. Silverman, T. R. Cech, *Biochemistry* **1999**, *38*, 14224 – 14237; b) S. K. Silverman, M. Zheng, M. Wu, I. Tinoco, Jr., T. R. Cech, *RNA* **1999**, *5*, 1665 – 1674; c) S. K. Silverman, M. L. Deras, S. A. Woodson, S. A. Scaringe, T. R. Cech, *Biochemistry* **2000**, *39*, 12465 – 12475; d) S. K. Silverman, T. R. Cech, *RNA* **2001**, *7*, 161 – 166; e) E. A. Doherty, R. T. Batey, B. Masquida, J. A. Doudna, *Nat. Struct. Biol.* **2001**, *8*, 339 – 343; f) S. Matsumura, Y. Ikawa, T. Inoue, *Nucleic Acids Res.* **2003**, *31*, 5544 – 5551.
[19] D. A. Peattie, *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 1760 – 1764.
[20] D. W. Celander, T. R. Cech, *Science* **1991**, *251*, 401 – 407.
[21] a) S. A. Woodson, *Nat. Struct. Biol.* **2000**, *7*, 349 – 352; b) K. L. Buchmueller, A. E. Webb, D. A. Richardson, K. M. Weeks, *Nat. Struct. Biol.* **2000**, *7*, 362 – 366; c) R. Russell, I. S. Millett, S. Doniach, D. Herschlag, *Nat. Struct. Biol.* **2000**, *7*, 367 – 370; d) R. Russell, I. S. Millett, M. W. Tate, L. W. Kwok, B. Nakatani, S. M. Gruner, S. G. Mochrie, V. Pande, S. Doniach, D. Herschlag, L. Pollack, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4266 – 4271; e) X. W. Fang, P. Thiyagarajan, T. R. Sosnick, T. Pan, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8518 – 8523; f) K. Takamoto, Q. He, S. Morris, M. R. Chance, M. Brenowitz, *Nat. Struct. Biol.* **2002**, *9*, 928 – 933; g) K. L. Buchmueller, K. M. Weeks, *Biochemistry* **2003**, *42*, 13869 – 13878; h) K. Takamoto, R. Das, Q. He, S. Doniach, M. Brenowitz, D. Herschlag, M. R. Chance, *J. Mol. Biol.* **2004**, *343*, 1195 – 1206.
[22] a) C. R. Allerson, G. L. Verdine, *Chem. Biol.* **1995**, *2*, 667 – 675; b) G. D. Glick, *Biopolymers* **1998**, *48*, 83 – 96; c) S. B. Cohen, T. R. Cech, *J. Am. Chem. Soc.* **1997**, *119*, 6259 – 6268.
[23] A. A. Szewczak, T. R. Cech, *RNA* **1997**, *3*, 838 – 849.

