

Chemical Synthesis of Mono- and Bis-Labeled Pre-MicroRNAs**

Ugo Pradère, Andreas Brunschweiler, Luca F. R. Gebert, Matije Lucic, Martina Roos, and Jonathan Hall*

Chemically synthesized single and multi-labeled oligonucleotides are enabling tools in molecular diagnostics,^[1–3] nanotechnologies,^[4,5] and other related fields. Most applications involve relatively short DNAs and RNAs; however the demand for labeled long RNA (>60 nt) has increased in line with our growing appreciation of a rich, only partially understood biology of structured non-coding RNAs.

A common secondary structure motif in RNA is the hairpin, which is composed of a terminal loop region and a double-stranded RNA stem that often bears mismatches and bulges. This motif is central to several biological functions including RNA maturation, RNA folding, ribozyme function, and protection of RNA from degradation.^[6] The hairpin also serves as a recognition site for RNA binding proteins as, for example, during microRNA (miRNA) biogenesis.^[7–9] Currently however, the difficult synthesis of long functionalized RNAs, for example, pre-miRNAs, hampers research on their biological properties (e.g. processing, metabolism, interaction partners, localization in cells, etc.).

Long, unmodified RNAs can be prepared by *in vitro* transcription from DNA templates. Functional groups can be incorporated by enzymatic or chemical ligation at the 5'-end,^[10–12] at the 3'-end,^[13,14] or at random internal positions through the use of labeled nucleoside triphosphates during transcription.^[4] Long RNAs labeled at defined internal positions are accessible in a multi-step semi-synthetic method by enzymatic ligation of short, labeled RNA strands. These are typically prepared by solid-phase synthesis using modified phosphoramidites.^[15–20] Considerable progress in labeling short oligonucleotides has been made by applying Cu^I-catalyzed azide/alkyne cycloaddition (“click chemistry”).^[21,22] This reaction was first used by the groups of Carell^[23] and Seela^[24] for labeling short DNA. In their “post-synthetic” approach, an alkyne-bearing nucleoside was incorporated into an oligodeoxyribonucleotide, and subsequently

reacted with an appropriate azide while still linked to the controlled pore glass (CPG) solid support. Elegant strategies were employed to achieve multiple labeling of short DNAs, for example through distinct protecting groups on the alkyne function,^[25] or through combined solid- and solution-phase click chemistry.^[26] The subsequent introduction of Cu^I-stabilizing ligands^[27,28] or of activated cycloaddition partners^[29] rendered these protocols efficient also for RNA. This procedure was first employed for solution phase labeling of RNAs,^[27,30–33] and was then adapted to solid phase conjugation of small molecules to short interfering RNAs (siRNA) with the aim to improve their targeting properties.^[34,35]

In the course of our work on pre-miRNAs we have needed to position labels at precise internal positions such that 1) the label did not disturb the cellular function of the RNA, and 2) it remained attached to the desired RNA sequence after biochemical processing into its 5p- and 3p-arms^[36] (Figure 1).

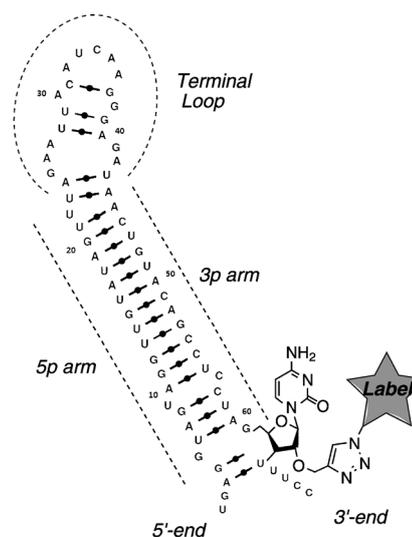


Figure 1. Representation of a pre-microRNA, internally labeled with a 2'-O-propargyl-ribonucleoside in its 3p arm.

However, we found that none of the literature protocols reported for RNA functionalization on solid support were suitable for the labeling of pre-miRNAs. Rajeev's conditions,^[34] which involve microwave irradiation on a large scale (0.6 μmol) were not compatible with the 50 nmol synthesis scale of pre-miRNAs. In our hands, Carell's conditions,^[25] although effective on short RNAs afforded labeled pre-miRNA in very low yields (10% or less) (e.g. Table 1, entry 1). The negative influence of oligoribonucleotide length

[*] Dr. U. Pradère, Dr. L. F. R. Gebert, M. Lucic, M. Roos, Prof. Dr. J. Hall
Institute of Pharmaceutical Sciences, Department of Chemistry and Applied Biosciences, ETH Zürich
8093 Zürich (Switzerland)
E-mail: jonathan.hall@pharma.ethz.ch

Dr. A. Brunschweiler
Faculty of Chemistry and Chemical Biology
TU Dortmund University
Otto-Hahn-Strasse 6, 44227 Dortmund (Germany)

[**] This work was supported by grants from the Swiss National Science Foundation (CRSII3_127454) and the ETH (ETH-01 11-2) to J.H. We thank H. Towbin and H. Jahns for help with the manuscript.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201304986>.

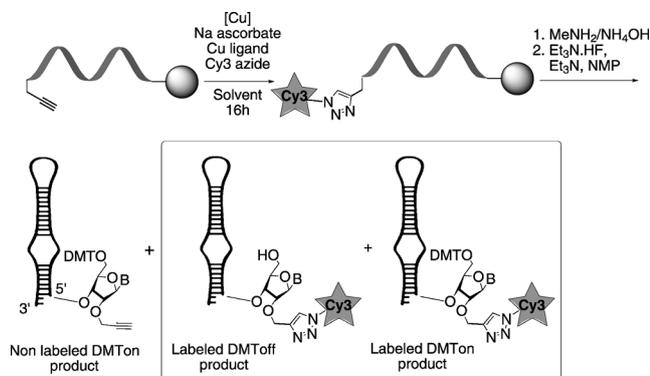
Table 1: Results from post-synthetic labeling of 5'-end modified pre-miR-21 with Cy3 azide **1** under various reaction conditions overnight.

Entry	Cat. ^[a]	Co-solvents with H ₂ O (vol.)	T [°C]	Degradation ^[b] [%]	SM ^[c] [%]	Yields ^[d] [%] (% as DMToff)
1	CuBr	<i>t</i> BuOH/DMSO (1.2 mL)	45	11	89	–
2	CuSO ₄	MeOH/THF (1.2 mL)	45	35	37	28 (5)
3	CuSO ₄	MeOH/DMF (1.2 mL)	45	14	20	67 (11)
4	CuSO ₄	<i>t</i> BuOH/DMSO (1.2 mL)	45	12	86	3 (3)
5	CuSO ₄	MeOH/DMF (1.2 mL)	25	–	94	6
6	CuSO ₄	MeOH/DMF (1.2 mL)	65	68	4	28 (25)
7	CuBr	MeOH/DMF (1.2 mL)	45	18	37	46 (8)
8	CuSO ₄	MeOH/DMF (400 μL)	45	18	–	83 (14)
9	CuSO ₄	MeOH/DMF (133 μL)	45	22	–	78 (9)

[a] CuSO₄ refers to CuSO₄·5 H₂O. [b] Estimated RNA degradation during the click reaction: calculated by subtracting the integral of peak D in the reference spectrum (failure sequences alone) from the integral of peak D under a given set of click conditions (failure sequences + degradation). [c] SM is unreacted starting material. [d] Yield in the click conjugation step: calculated from the peak integrals (A + B)/(A + B + C + D), where integral of D is corrected for failure sequences using the reference spectrum (see [b] and text).

on reactivity was surprising and obliged us to investigate new protocols to label pre-miRNAs.

Although dedicated phosphoramidites have been developed for the synthesis of long RNAs^[37,38] we found that regular 2'-*O*-*tert*-butyldimethylsilyl (TBDMS) phosphoramidites afford in good yields pre-miRNAs in lengths up to 70 nt, even with the incorporation of one or two alkyne-modified nucleosides for click conjugations. 2'-*O*-Propargyl cytidine was selected as coupling partner because of a straightforward synthesis (cf. Supporting Information) and the positioning of the label in the minor groove of the duplex, thereby minimizing interference with base-pairing. In addition, an enhanced reactivity of the alkyne group, presumably due to



Scheme 1. Approach employed to screen reaction conditions (Table 1): click labeling of **1** to 5'-end modified pre-miR-21.

the α -oxygen substituent, was observed in a comparison with the cytidine bearing a longer 2'-*O*-pentynyl linker during click reactions (data not shown).

In order to determine an optimal set of reaction conditions, we investigated systematically variations of the copper salt, solvents, reaction volumes, and temperatures in a series of side-by-side comparisons (Scheme 1). The oncomir pre-miR-21 (60 nt) was selected as a test hairpin. Initially, 2'-*O*-propargyl cytidine was incorporated at its 5' terminus to avoid the labeling of non-full length failure sequences ($n-1$, $n-2$, etc.) and thereby simplifying interpretation of HPLC spectra. Ten CPG supports bearing the pre-miR-21 sequence were synthesized, mixed, and redistributed equally into 10 wells to ensure an identical starting point for each reaction condition. In one well, the CPG was not subjected to click conjugation: After a standard work-up, its HPLC chromatogram was used as a reference spectrum (see below). Cy3 azide (**1**) was prepared (see Supporting Information) and used for the labeling on the remaining nine CPGs. This hydrophobic molecule increases the HPLC retention time of a conjugated RNA providing sufficient separation of the products (Figure 2) and consequently their unambiguous identification by mass spectrometry (MS). All reactions were run over 16 h. Partial loss of the terminal DMT group during the click reaction was always observed, leading to two product peaks in the chromatogram (Figure 2: peaks A and B). However, these products were easily distinguished from the unreacted starting material (SM) (peak C) and the fast migrating RNA (peak D) comprising the failure sequences together with RNA degradation products from the click reaction. The results of the nine experiments are summarized in Table 1.

Analysis of data in Table 1 led us to an optimized set of conditions for the conjugation reaction. The best reaction solvent mixture was H₂O/MeOH/DMF: H₂O/*t*BuOH/DMSO was unsatisfactory in the presence of either CuBr (entry 1) or CuSO₄·5 H₂O (entry 4) catalysts with no or trace quantities of products, respectively. Surprisingly, DMF was a superior co-solvent to THF, increasing the product yield from 28%

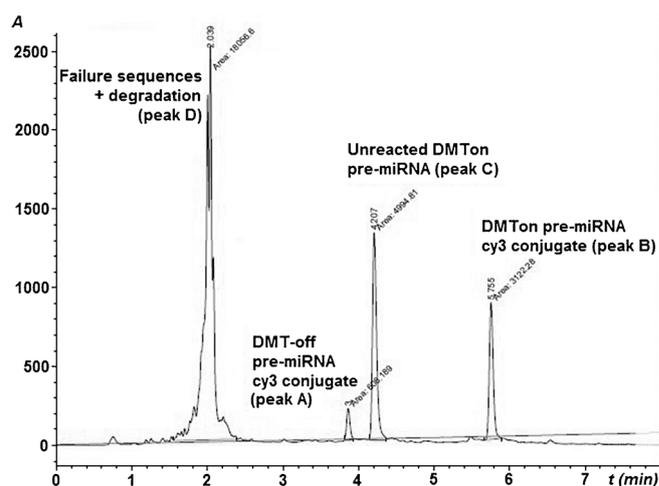


Figure 2. HPLC chromatogram of the products from Table 1, entry 2. DMTon: dimethoxytrityl-protected oligoribonucleotide; DMToff: fully deprotected oligoribonucleotide; Cy3: Cyanine3.

(entry 2) to 67% (entry 3). $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (entry 3, 67%) was superior to CuBr in $\text{H}_2\text{O}/\text{MeOH}/\text{DMF}$ (entry 7, 46%). Overnight reaction at 45 °C (entry 3, 67%) was better than at 25 °C (entry 5, 6%) and 65 °C (entry 6, 28%), with minimum RNA degradation. Finally, total conversion was reached with reaction volumes of 400 μL (entry 8, 83%) and 133 μL (entry 9, 78%), though the former led to less degradation. Consequently, the reaction conditions of entry 8 were adopted for the labeling of pre-miRNAs.

We then assessed the efficiency of our protocol with Cy3 azide (**1**) at positions within pre-miR-21 selected to show the general applicability of the method: at the 3' terminus, in the hairpin's 5p and 3p arms, and in its terminal loop. Reaction conversions were estimated by comparison of the DMTon and DMToff conjugates with the unreacted pre-miRNA. The 5' terminus, the 5p arm, and the loop were efficiently labeled in conversions of >95%, 82%, and >95%, respectively (Table 2). Surprisingly, labeling close to or at the 3' terminus was less efficient with respectively 68% and 38% conversions. Double labeling was successfully carried out within the 3p and 5p arms as well as the 3'-terminus and the loop in conversions of 64% and 33%, respectively.

Table 2: Post-synthetic single and double homo-labeling of pre-miR-21 at different positions with azide **1**.

Product	Position of the label (nt) ^[a]	Conv. ^[b]
ORN-1	5'-end (1)	> 95%
ORN-2	5p (13)	82%
ORN-3	loop (32)	> 95%
ORN-4	3p (56)	68%
ORN-5	3'-end (60)	38%
ORN-6	5p/3p (32/56)	64% ^[c]
ORN-7	3'-end/loop (32/60)	33% ^[d]

[a] Nucleotide number of the modification from the 5'-terminus; 3p: 3p arm; 5p: 5p arm. [b] Estimated conversion in the click step (calculated from the peak integrals $(A+B)/(A+B+C)$). [c] Single-labeled products as the major side products (36%). [d] Single-labeled products as the major side products (57%) while 10% of the hairpin remained unlabeled.

The protocol was generally applicable for pre-miRNAs and with a variety of labels (Table 3). We conjugated pre-let-7a-2 (67 nt), pre-miR-106a (59 nt), pre-miR-122 (58 nt), and pre-miR-32 (63 nt) to a selection of functional groups commonly used in chemical biology. These included Cy3 (**1**), which we used for cellular localization studies, biotin (**2**) for immobilization, trioxalen (**3**) for cross-linking studies, and black hole quencher 1 (BHQ-1) (**4**) for FRET assays (Figure 3, synthesis in the Supporting Information). Following our procedure, 19 labeled pre-miRNAs were quickly prepared. Conversions of the conjugation reaction are reported in Table 3 which unexpectedly indicates that the efficiency of conjugation is more dependent on the labeled position than on the nature of the label. The 5' terminus, the 5p arm, and the terminal loop of the pre-miRNAs were all functionalized in good to excellent conversions (55–>95%, 81–94% and 80–>95%, respectively). Conversely, the 3p arm and the 3' terminus of the pre-miRNAs were less efficiently labeled with conversions of 46–92% (average of 60%) and 33–54%,

Table 3: Post-synthetic labeling of pre-miRNAs at different positions.

Product	Pre-miRNA	Position of the label ^[a]	Label	Conv. ^[b]
ORN-8	pre-miR-21	3p (56)	biotin	46%
ORN-9	pre-miR-21	loop (44)	BHQ	90%
ORN-10	pre-miR-21	3p (56)	BHQ	53%
ORN-11	pre-miR-21	3p (34)	trioxalen	92%
ORN-12	pre-let-7a-2	3p (62)	biotin	64%
ORN-13	pre-let-7a-2	loop (31)	Cy3	90%
ORN-14	pre-let-7a-2	3'-end (67)	BHQ	33%
ORN-15	pre-let-7a-2	5p (13)	trioxalen	91%
ORN-16	pre-miR-106a	3p (55)	biotin	60%
ORN-17	pre-miR-106a	5p (8)	BHQ	92%
ORN-18	pre-miR-106a	loop (35)	trioxalen	89%
ORN-19	pre-miR-122	5'-end (1)	BHQ	55%
ORN-20	pre-miR-122	3p (52)	biotin	53%
ORN-21	pre-miR-122	5p (11)	Cy3	94%
ORN-22	pre-miR-122	loop (26)	trioxalen	93%
ORN-23	pre-miR-32	loop (42)	biotin	80%
ORN-24	pre-miR-32	5p (6)	Cy3	81%
ORN-25	pre-miR-32	3'-end (62)	Cy3	54%
ORN-26	pre-miR-32	5'-end (1)	trioxalen	93%

[a] Nucleotide number of the modification from the 5'-terminus; 3p: 3p arm; 5p: 5p arm. [b] Estimated conversion in the click step (calculated from the peak integrals $(A+B)/(A+B+C)$).

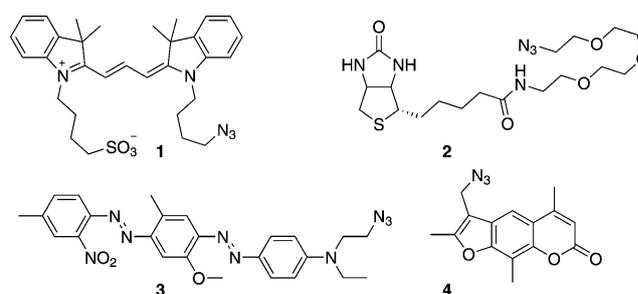


Figure 3. Labels used to functionalize pre-miRNAs in this study.

respectively. This confirmed that labeling becomes sluggish at positions closer to the solid support, possibly because of steric hindrance.

Finally, we examined the double hetero-labeling of pre-miRNAs which is more challenging because it requires incorporation and selective functionalization of two modified ribonucleotides. We elected to use a procedure similar to the method of Wengel^[26] by interrupting solid phase synthesis for the first functionalization. The second functionalization—still on solid support—was performed after the final coupling step, as for the mono-functionalized pre-miRNAs. We found it advantageous to perform the first click reaction as late as possible in the synthesis, that is, just prior to incorporation of the second modified cytidine. This procedure permitted us to prepare four hetero double-labeled pre-miRNAs bearing combinations of Cy3/biotin and Cy3/BHQ-1 in low to moderate conversions (20–61%) (Table 4). This procedure, although quite efficient for some sequences, still suffers from some limitations, particularly when the first click reaction does not proceed to completion. This leads to functionalization at both positions with the second label and product mixtures that could not be easily separated.

- [3] A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, S. Tyagi, *Nat. Methods* **2008**, *5*, 877–879.
- [4] G. A. Burley, J. Gierlich, M. R. Mofid, H. Nir, S. Tal, Y. Eichen, T. Carell, *J. Am. Chem. Soc.* **2006**, *128*, 1398–1399.
- [5] J. J. Storhoff, C. A. Mirkin, *Chem. Rev.* **1999**, *99*, 1849–1862.
- [6] P. Svoboda, A. Di Cara, *Cell. Mol. Life Sci.* **2006**, *63*, 901–908.
- [7] H. Towbin, P. Wenter, B. Guennewig, J. Imig, J. A. Zagalak, A. P. Gerber, J. Hall, *Nucleic Acids Res.* **2013**, *41*, e47.
- [8] G. Michlewski, S. Guil, C. A. Semple, J. F. Caceres, *Mol. Cell* **2008**, *32*, 383–393.
- [9] H. Siomi, M. C. Siomi, *Mol. Cell* **2010**, *38*, 323–332.
- [10] J. C. Schlatterer, A. Jäschke, *Biochem. Biophys. Res. Commun.* **2006**, *344*, 887–892.
- [11] H. Vogel, C. Richert, *ChemBioChem* **2012**, *13*, 1474–1482.
- [12] S. Chirayil, R. Chirayil, K. J. Luebke, *Nucleic Acids Res.* **2009**, *37*, 5486–5497.
- [13] T.-p. Wu, K.-c. Ruan, W.-y. Liu, *Nucleic Acids Res.* **1996**, *24*, 3472–3273.
- [14] D. Proudnikov, A. Mirzabekov, *Nucleic Acids Res.* **1996**, *24*, 4535–4542.
- [15] B. P. Davies, C. Arenz, *Angew. Chem.* **2006**, *118*, 5676–5679; *Angew. Chem. Int. Ed.* **2006**, *45*, 5550–5552.
- [16] M. J. Moore, C. C. Query, *Methods Enzymol.* **2000**, *317*, 109–123.
- [17] K. H. Yeom, I. Heo, J. Lee, S. Hohng, V. N. Kim, C. Joo, *EMBO Rep.* **2011**, *12*, 690–696.
- [18] K. Fauster, C. Kreutz, R. Micura, *Angew. Chem.* **2012**, *124*, 13257–13261; *Angew. Chem. Int. Ed.* **2012**, *51*, 13080–13084.
- [19] K. Fauster, M. Hartl, T. Santner, M. Aigner, C. Kreutz, K. Bister, E. Ennifar, R. Micura, *ACS Chem. Biol.* **2012**, *7*, 581–589.
- [20] C. M. Dojahn, M. Hesse, C. Arenz, *Chem. Commun.* **2013**, *49*, 3128–3130.
- [21] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064.
- [22] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2708–2711; *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.
- [23] J. Gierlich, G. A. Burley, P. M. Gramlich, D. M. Hammond, T. Carell, *Org. Lett.* **2006**, *8*, 3639–3642.
- [24] F. Seela, V. R. Sirivolu, *Chem. Biodiversity* **2006**, *3*, 509–514.
- [25] P. M. E. Gramlich, S. Warncke, J. Gierlich, T. Carell, *Angew. Chem.* **2008**, *120*, 3491–3493; *Angew. Chem. Int. Ed.* **2008**, *47*, 3442–3444.
- [26] I. K. Astakhova, J. Wengel, *Chem. Eur. J.* **2013**, *19*, 1112–1122.
- [27] V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem.* **2009**, *121*, 10063–10067; *Angew. Chem. Int. Ed.* **2009**, *48*, 9879–9883.
- [28] T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org. Lett.* **2004**, *6*, 2853–2855.
- [29] I. Singh, C. Freeman, A. Madder, J. S. Vyle, F. Heaney, *Org. Biomol. Chem.* **2012**, *10*, 6633–6639.
- [30] E. Paredes, S. R. Das, *ChemBioChem* **2011**, *12*, 125–131.
- [31] S. Kellner, S. Seidu-Larry, J. Burhenne, Y. Motorin, M. Helm, *Nucleic Acids Res.* **2011**, *39*, 7348–7360.
- [32] L. Eltepu, M. Jayaraman, K. G. Rajeev, M. Manoharan, *Chem. Commun.* **2013**, *49*, 184–186.
- [33] W. Wang, K. Chen, D. Qu, W. Chi, W. Xiong, Y. Huang, J. Wen, S. Feng, B. Zhang, *Tetrahedron Lett.* **2012**, *53*, 6747–6750.
- [34] T. Yamada, C. G. Peng, S. Matsuda, H. Addepalli, K. N. Jayaprakash, M. R. Alam, K. Mills, M. A. Maier, K. Charisse, M. Sekine, M. Manoharan, K. G. Rajeev, *J. Org. Chem.* **2011**, *76*, 1198–1211.
- [35] J. Willibald, J. Harder, K. Sparrer, K. K. Conzelmann, T. Carell, *J. Am. Chem. Soc.* **2012**, *134*, 12330–12333.
- [36] G. Hutvagner, J. McLachlan, A. E. Pasquinelli, E. Balint, T. Tuschl, P. D. Zamore, *Science* **2001**, *293*, 834–838.
- [37] S. Pitsch, P. A. Weiss, L. Jenny, A. Stutz, X. Wu, *Helv. Chim. Acta* **2001**, *84*, 3773–3795.
- [38] Y. Shiba, H. Masuda, N. Watanabe, T. Ego, K. Takagaki, K. Ishiyama, T. Ohgi, J. Yano, *Nucleic Acids Res.* **2007**, *35*, 3287–3296.
- [39] J. S. Yang, M. D. Phillips, D. Betel, P. Mu, A. Ventura, A. C. Siepel, K. C. Chen, E. C. Lai, *RNA* **2011**, *17*, 312–326.
- [40] A. Birmingham, E. Anderson, K. Sullivan, A. Reynolds, Q. Boese, D. Leake, J. Karpilov, A. Khvorova, *Nat. Protoc.* **2007**, *2*, 2068–2078.
- [41] S. Sakamoto, K. Aoki, T. Higuchi, H. Todaka, K. Morisawa, N. Tamaki, E. Hatano, A. Fukushima, T. Taniguchi, Y. Agata, *Mol. Cell. Biol.* **2009**, *29*, 3754–3769.
- [42] C. Okada, E. Yamashita, S. J. Lee, S. Shibata, J. Katahira, A. Nakagawa, Y. Yoneda, T. Tsukihara, *Science* **2009**, *326*, 1275–1279.

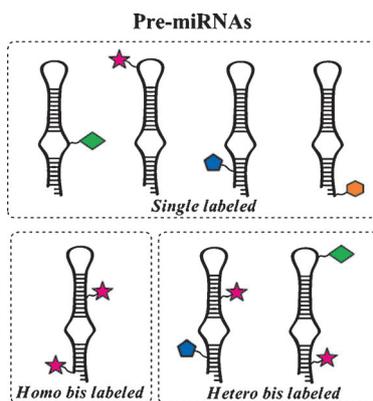
Communications



Pre-microRNA Labeling

U. Pradère, A. Brunschweiler,
L. F. R. Gebert, M. Lucic, M. Roos,
J. Hall*     

Chemical Synthesis of Mono- and Bis-
Labeled Pre-MicroRNAs



A chemical method for the post-synthetic labeling of pre-miRNAs on solid support using easily accessible reagents was developed. The procedure was employed to generate a library of 31 pre-microRNAs carrying labels commonly used in chemical biology, including Cy3, trioxalen, biotin, and BHQ-1.