Decarboxylative Claisen condensation catalyzed by *in vitro* selected ribozymes†

Youngha Ryu,‡ Kil-Joong Kim, Charles A. Roessner and A. Ian Scott*

Received (in Cambridge, MA, USA) 6th December 2005, Accepted 9th February 2006
First published as an Advance Article on the web 22nd February 2006
DOI: 10.1039/b517160a

The *in vitro* selection of RNAs catalyzing the decarboxylative Claisen condensation provides evidence for the synthesis of fatty acids, the building blocks of lipids and membranes, in the "RNA world".

Life on earth as we know it today is confined within membranes consisting mainly of lipids and membrane proteins. Lipids are composed of fatty acid building blocks which, in turn, are synthesized by fatty acid synthases *via* the decarboxylative Claisen condensation (Fig. 1).

Polyketides are also synthesized *via* the Claisen condensation by ketosynthases. Although evolution has resulted in the enzymatic synthesis of fatty acids, they may have already existed in the prebiotic world,[1] and mechanisms for their synthesis without enzymes have been proposed using RNA templates[2] or self-replication.[3] According to the "RNA world" hypothesis, however, an RNA catalyst, or ribozyme, for the Claisen condensation would be anticipated.[4] Naturally occurring ribozymes for RNA splicing,[5,6] RNA editing,[7] peptide bond formation[8] and control of gene expression[9] have been discovered, and an *in vitro* selection technique (SELEX)[10] has provided ribozymes for many biologically important chemical reactions including RNA polymerization,[11] alkylation,[12] transesterification,[13] aminoacylation,[14] peptide bond formation,[15] nucleotide formation,[16] and carbon–carbon bond formation *via* the Diels–Alder[17,18] and aldol reactions.[19]

Since acetyl-CoA and malonyl-CoA are the normal substrates in fatty acid and polyketide biosynthesis *via* the Claisen condensation, of particular significance for the synthesis of lipids is the reported selection of ribozymes that synthesize CoA and acyl-CoAs.[20–22] However, there have been no reports of the isolation of ribozymes capable of utilizing acyl-CoAs as the building blocks for the synthesis of fatty acids (or polyketides) *via* the Claisen condensation. We describe here the isolation of ribozymes with decarboxylative Claisen condensation activity using an *in vitro* selection scheme based on the SELEX procedure. Briefly, the strategy (Fig. 2) involved incubating a pool of malonyl-PEG-RNA conjugates with biotinylated acyl-CoA (Biotin-LC-CoA). Catalytic RNA molecules that became covalently attached to biotin due to condensation *via* the decarboxylative Claisen reaction were separated from the non-active members of the pool using streptavidin agarose affinity chromatography and further enriched by iterative selection cycles. The malonyl-PEG-RNA pool was generated by: (1) transcription of a random DNA library (142 random nt in a 222 nt transcript) in the presence of an excess of the modified base HS-TEG-GMP to provide part of the linker and a free thiol at the 5′ end of the RNA molecules followed by: (2) malonylation with malonyl-TEG-thiosulfate. The resultant malonyl-PEG-RNA pool was then incubated with the substrate Biotin-LC-CoA. RNA able to form a C–C bond *via* the Claisen condensation, now covalently linked to the biotin tag, was selected by streptavidin affinity chromatography. The active RNA molecules were eluted from the streptavidin column by reduction of the disulfide linkage with dithiothreitol (DTT). The RNAs thus selected were reverse-transcribed to complementary DNA (cDNA), amplified by PCR, transcribed to 5′-HS-PEG-RNA,

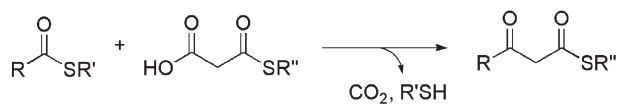

Fig. 1 The decarboxylative Claisen condensation. In fatty acid and polyketide synthesis, R′ = R″ = Coenzyme A (CoA).

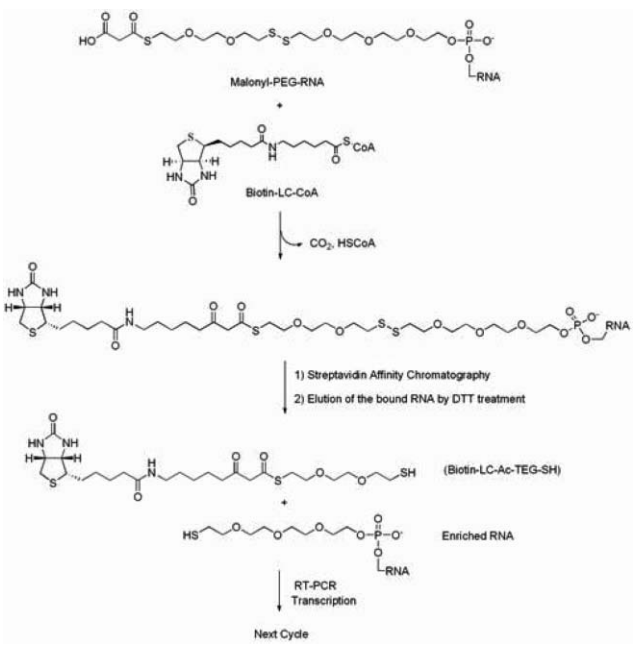

Fig. 2 Scheme for the selection of a ketosynthase ribozyme.

Department of Chemistry, Texas A&M University, College Station, TX, USA 77843-3255. E-mail: Scott@mail.chem.tamu.edu; Fax: 01 979 845 5299; Tel: 01 979 845 3243

† Electronic supplementary information (ESI) available: Experimental details, RNA sequence and structure analysis, and ESI mass spectrograms. See DOI: 10.1039/b517160a

‡ *Present address*: Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92307, USA.

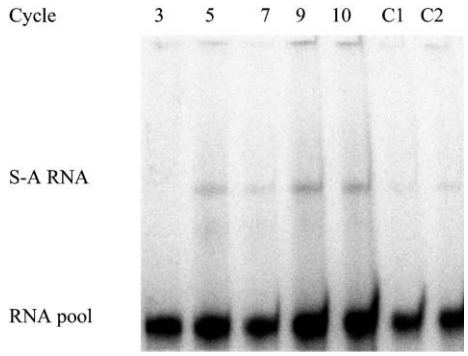

Fig. 3 Gel mobility shift assay of certain cycles of the ketosynthase selection process. S-A RNA is the RNA bound to streptavidin. The presence of streptavidin in the band was confirmed by Coomassie blue staining. Control reactions for pool-10 RNA: C-1, no streptavidin; C-2, RNA treated with DTT before streptavidin incubation. The relative amount of S-A RNA: 3, 0.1%; 5, 2.9%; 7, 1.2%; 9, 5.1%; 10, 4.2%; C1, 0.2%; C2, 0.6%.

and again converted to malonyl-PEG-RNA to enter the next cycle of selection. The cycle was repeated 10 times, and reactions performed with the RNA from several cycles were chosen for analysis by a gel mobility shift assay, which demonstrated the enrichment of streptavidin-bound RNA molecules (Fig. 3).

As Biotin-LC-CoA was used as an electrophile for the Claisen condensation, it could react with bases of RNA through simple acylation to make an amide bond as well as with malonic acid through Claisen condensation. As a result, it could produce RNA connected to the Biotin tag through amide without disulfide bond, which would not be cleaved during the main reaction and make a complex with Streptavidin to show a band shift at the same position as S-A RNA. To verify the possibility of a non-specific acylation reaction, we performed the reaction of Biotin-LC-CoA with an RNA obtained without HS-TEG-GMP and malonylating agent. The result showed the band at the same position (data not shown). Supposing that the amount of non-specific reaction and S–S cleavage by HSCoA (discussed later) are constant in all reactions, the activity of RNA increased from almost zero of round 3 to about 5% after round 9.

To isolate single ketosynthase ribozymes, cDNA, generated from the cycle 10 RNA by reverse transcription and PCR amplification, was inserted between the *Eco*RI and *Bam*HI sites of plasmid pHN1+[23] and transformed into *E. coli* strain XA90. Plasmid DNA was isolated from 29 individual clones. The insert was amplified by PCR and transcribed to provide RNA for analysis by the gel mobility shift assay. Of the 29 individual clones, clones 10 and 29 demonstrated enhanced ketosynthase activity (8.8% and 13%, respectively. Gel mobility shift assay is shown in supplemental material). Primary sequence alignment showed that the 181-nucleotide RNA clone 10 and the 172-nucleotide RNA clone 29 have 39% identity in the random regions (see supplementary material). No common motif could be found from the minimum energy secondary structures predicted for two individual clones. However, the consensus secondary structure predicted from the aligned primary sequence by the RNAalignfold program[24] showed a high degree of similarity between the two RNAs (see supplemental material).

Mass analysis was used to confirm that the decarboxylative Claisen condensation was being catalyzed by the RNA. Malonyl-PEG-RNA derived from clone 29 was reacted with Biotin-LC-CoA, and a chloroform extract of the reaction mixture and an aqueous layer followed by intentional reduction with DTT (see supplemental material) were analyzed by MALDI-TOF mass spectrometry. A mass peak at 564 [M + H^+] (see mass spectra in the supplemental material) was observed only in the chloroform extract, which is identical to that expected for biotin-LC-Ac-TEG-SH, the compound resulting from cleavage of the disulfide bond in the RNA-bound product (Fig. 2). Authentic biotin-LC-Ac-TEG-SH was synthesized and provided a nearly identical spectrum on mass analysis. Also observed in both spectra are peaks at 586 and 602 corresponding to [M + Na^+] and [M + K^+].

In a control experiment to determine whether the product resulted from the catalytic activity of RNA or from chemical decarboxylative condensation, biotin-LC-CoA was incubated with malonyl-PEG under the same conditions as the selection reaction, but without RNA. No trace of the product expected from the decarboxylative Claisen condensation of these two substrates was seen by MALDI MS analysis (not shown). This result, along with the observation that two of the selected clones demonstrated significantly higher activity than the other 27, supports the conclusion that the isolated product can be formed only through catalytic action followed by reduction of the disulfide bond of the product shown in Fig. 2.

The isolated compound is observed, however, even without the addition of reducing agents prior to chloroform extraction of the reaction mixture, but we believe that CoA released during the condensation reaction and by hydrolysis of the biotin-LC-CoA substrate could be responsible for reduction of the disulfide linkage. This was confirmed by the reaction of HSCoA with 2-Pyr-S-S-PEG-AMP by monitoring the release of 2-thiopyridine that shows a strong absorbance at 343 nm as a result of cleavage of disulfide. When only one equivalent of HSCoA was added to an aqueous solution of 2-Pyr-S-S-PEG-AMP, the reduction took place immediately on mixing. Therefore, the amount of HSCoA released from the decarboxylative Claisen condensation seems to be sufficient to cleave the disulfide bond of the product. Furthermore, extra HSCoA could be supplied from the hydrolysis of Biotin-LC-CoA and non-specific acylation to the reaction mixture. This unexpected cleavage not only limited the amount of RNA-linked product available for enrichment by streptavidin-agarose column chromatography, but has so far complicated kinetic studies. Even without these studies, however, we believe that we now have sufficient evidence that RNA can catalyze the Claisen reaction as do the fatty acid synthase and ketosynthase enzymes. The success of this procedure constitutes the first experimental proof that the "RNA world" may, indeed, have been capable of synthesizing fatty acids, the precursors of the lipids essential for the structure of membranes, and lends further credence to the hypothesis that RNA could carry out most of the chemical reactions required in the absence of proteins.

We thank the Robert A.Welch Foundation and the Texas Advanced Technology and Research Program for financial support, Larry Harris-Haller (Gene Technologies Lab, Biology Department, TAMU) for DNA sequence analysis, and Shane Tichy and Vanessa Santiago (Laboratory for Biological Mass Spectrometry, Chemistry Department, TAMU) for mass analysis.

Notes and references

1 D. Deamer, *Astrobiology*, 2002, **2**, 371.
2 A. I. Scott, *Tetrahedron Lett.*, 1997, **38**, 4961.
3 J. W. Szostak, D. P. Bartel and P. L. Luisi, *Nature*, 2001, **409**, 387.
4 W. Gilbert, *Nature*, 1986, **319**, 618.
5 T. R. Cech, A. J. Saug and P. J. Grabowski, *Cell*, 1981, **27**, 487.
6 S. Valadkhan and J. L. Manley, *Nature*, 2001, **413**, 701.
7 C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace and S. Altman, *Cell*, 1983, **35**, 849.
8 N. Ban, P. Nissen, J. Hansen, P. B. Moore and T. A. Steitz, *Science*, 2000, **289**, 905.
9 W. C. Winkler, A. Nahvi, A. Roth, J. A. Collins and R. R. Breaker, *Nature*, 2004, **428**, 281.
10 D. S. Wilson and J. W. Szostak, *Annu. Rev. Biochem.*, 1999, **68**, 611.
11 W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner and D. P. Bartel, *Science*, 2001, **292**, 1319.
12 C. Wilson and J. W. Szostak, *Nature*, 1995, **374**, 777.
13 A. Jenne and M. Famulok, *Chem. Biol.*, 1998, **5**, 23.
14 M. Illangasekare, G. Sanchez, T. Nickles and M. Yarus, *Science*, 1995, **267**, 643.
15 B. Zhang and T. R. Cech, *Nature*, 1997, **390**, 96.
16 P. J. Unrau and D. P. Bartel, *Nature*, 1998, **395**, 260.
17 T. M. Tarasow, S. L. Tarasow and B. E. Eaton, *Nature*, 1997, **389**, 54.
18 B. Seelig and A. Jäschke, *Chem. Biol.*, 1999, **6**, 167.
19 S. Fusz, A. Eisenfuhr, S. G. Srivatasan, A. Heckel and M. Famulok, *Chem. Biol.*, 2005, **12**, 941.
20 F. Huang, C. W. Bugg and M. Yarus, *Biochemistry*, 2000, **39**, 15548.
21 V. R. Jadhav and M. Yarus, *Biochemistry*, 2002, **41**, 723.
22 T. M. Coleman and F. Huang, *Chem. Biol.*, 2002, **9**, 1227.
23 S. L. Schreiber and G. L. Verdine, *Tetrahedron*, 1991, **47**, 2543.
24 I. L. Hofacker, M. Fekete and P. F. Stadler, *J. Mol. Biol.*, 2002, **319**, 1059.