Protein Production

DOI: 10.1002/anie.200605237

Cell-Free Transcription/Translation from PCR-Amplified DNA for High-Throughput NMR Studies**

*Peter S. C. Wu, Kiyoshi Ozawa, Siew P. Lim, Subhash G. Vasudevan, Nicholas E. Dixon, and Gottfried Otting**

Cell-free coupled transcription/translation systems based on *Escherichia coli* extracts have become an increasingly attractive alternative to conventional expression systems in vivo.[1–3] In particular, they allow the fast synthesis of isotope-labeled proteins for direct analysis by NMR spectroscopy without chromatographic purification of the protein,[4] they suppress metabolic conversion between different amino acids,[5] and they use expensive isotope-labeled amino acids sparingly.[2] The best protein yields are obtained in a dialysis (continuous-exchange cell-free; CECF) system in which the outside buffer solution continuously replenishes the inside reaction mixture with amino acids and adenosine 5′-triphosphate (ATP).[6] In practice, the most time-consuming step is the large-scale preparation of plasmid DNA, as CECF reactions perform much better with plasmid DNA than with linear DNA templates. This is presumably due to degradation by exonucleases present in the *E. coli* extracts.[7] For example, reported yields of chloramphenicol acetyltransferase are up to 1 mg of protein per mL of reaction mixture when polymerase chain reaction (PCR)-amplified linear DNA is used,[8] whereas the use of plasmid DNA gives a yield of 6 mg mL^{-1}.[6] Herein we present a technique that provides the same protein yields from PCR products as from plasmid DNA. This makes it possible to obtain protein NMR spectra within 24 h of gene amplification by PCR, including 7 h for cell-free protein synthesis and overnight dialysis against NMR buffer solution.

Our method is designed to generate complementary 8-base-pair, single-stranded overhangs at both ends of the PCR templates (Figure 1). These overhangs are suitable for ligation into exonuclease-insensitive cyclized templates by the endogenous ligase activity of *E. coli* S30 extracts. For all proteins tested in our laboratory, PCR products with 5′-phosphorylated overhangs produced the same protein yields in the CECF system as with plasmid DNA (Figure 2).

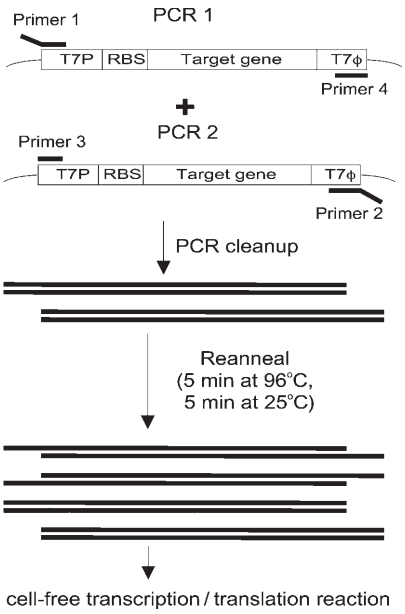

Figure 1. Generation of stable PCR templates from a construct with a T7 ϕ10 promoter, a ribosome binding site, and a T7 terminator. The primer sequences were (single-stranded overhang regions underlined): primer 1: 5′-PO_4-TTAGCTGGTCGATCCCGCGAAATTAATACG; primer 2: 5′-PO_4-CCAGCTAACAAAAAACCCCTCAAGACCCG; primer 3: 5′-PO_4-TCGATCCCGCGAAATTAATACG; primer 4: 5′-PO_4-CAAAAAACCCCTCAAGACCCG. When starting from constructs without T7 elements, the T7 sequences can be introduced in an additional PCR step (see the Supporting Information).

Evidence for cyclization was obtained indirectly. Figure 3 shows that the best yields were obtained with 1) longer overhangs and 2) 5′-phosphorylated overhangs; 3) single-stranded overhangs were essential; and 4) the precise nucleotide sequences of the overhangs were unimportant provided they were complementary.

The high protein yields obtained with the PCR protocol of Figure 1 open many important opportunities for high-throughput protein synthesis. For example, mutant proteins can readily be prepared and analyzed by NMR spectroscopy without cloning or chromatographic protein purification. Here we used this protocol to create site-directed mutants by PCR for site-specific resonance assignments of a dengue virus protease construct.

[*] P. S. C. Wu, Dr. K. Ozawa, Prof. G. Otting
Research School of Chemistry
The Australian National University
Canberra ACT 0200 (Australia)
Fax: (+61) 2-6125-0750
E-mail: gottfried.otting@anu.edu.au
Homepage: http://rsc.anu.edu.au/~go/
Prof. N. E. Dixon
Department of Chemistry
University of Wollongong, NSW 2522 (Australia)
Dr. S. P. Lim, Dr. S. G. Vasudevan
Novartis Institute for Tropical Diseases Pte Ltd
10 Biopolis Road, Singapore 138670 (Singapore)

[**] This work was supported by the Australian Research Council with a Federation Fellowship to G.O. and grant funding for the 800-MHz NMR spectrometer and this project. We thank Dr. Daying Wen and Dr. Xun-Cheng Su for a ^{15}N-HSQC spectrum of the uniformly labeled protease. PCR = polymerase chain reaction.

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

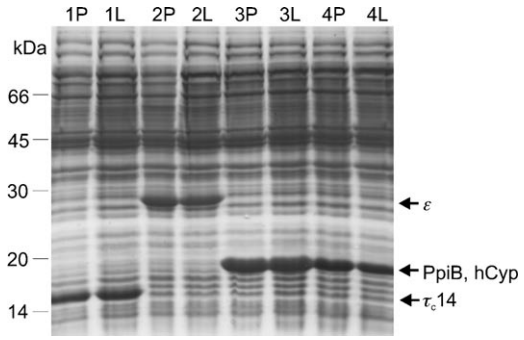

Figure 2. Comparison of protein yields obtained with plasmid DNA (P) or PCR-amplified linear DNA (L), as outlined in Figure 1, by using SDS-PAGE with Coomassie blue staining. 1) C-terminal fragment τ_C14 of the τ subunit of *E. coli* DNA polymerase III (Pol III). 2) ε Subunit of Pol III. 3) *E. coli* peptidylprolyl isomerase B (PpiB). 4) Human cyclophilin A (hCyp).

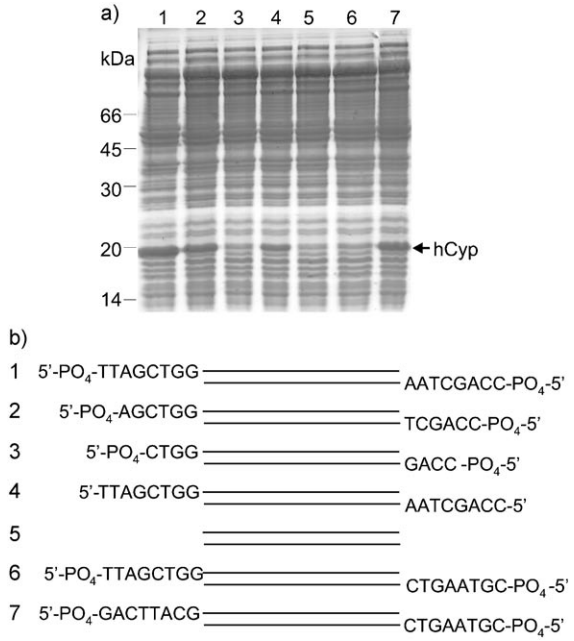

Figure 3. Expression yields achieved with different PCR templates. a) Coomassie blue-stained SDS-PAGE of hCyp synthesized by CECF reactions by using PCR products (10 $\mu g\,mL^{-1}$) with the overhangs shown in (b).

The dengue virus NS2B/NS3 protease complex is essential for dengue virus replication.[9] ^{15}N-HSQC spectra of $^{15}N/^{13}C$-labeled protease yielded very broad signals (full-width at half maximum > 50 Hz) owing to nonspecific aggregation or chemical exchange processes (see the Supporting Information).[10] In addition, many residues were highly mobile, resulting in poor spectral resolution, and the protein was prone to degradation. These properties prevent conventional resonance assignment strategies, yet, well-resolved cross-peaks were observed for samples that were selectively labeled with $^{15}N/^{13}C$ isoleucine. We assigned these cross-peaks by systematic site-directed mutagenesis of each of the sixteen isoleucine residues in the amino acid sequence to valine. As only isoleucine was isotope labeled, each peak missing from the HSQC spectrum of a mutant protein identified the assignment of the corresponding isoleucine residue in the wild-type protein.

The mutant constructs with the requisite terminal single-stranded overhangs were generated by using a modified overlap extension protocol (see the Supporting Information). The PCR products were used directly as templates for expression. Mutagenesis thus added little to the time required for sample preparation, and many mutations were easily made in parallel. The ^{15}N-HSQC and ^{13}C-HSQC spectra were recorded by using an experiment with simultaneous evolution of ^{15}N and ^{13}C resonances in the indirect frequency dimension.[11] In this way, 15 of the 16 isoleucine amide cross-peaks and 11 of the 22 isoleucine methyl cross-peaks were assigned (most of the isoleucine methyl peaks are unresolved, see Figure 4b). The same amide cross-peak was assigned to Ile 123 and Ile 165 of NS3 because mutation of either residue led to its disappearance (see Figure 4a and the Supporting Information). As both residues are close in the three-dimensional structure,[9] mutations at either site could affect the cross-peak, resulting in an ambiguous assignment.

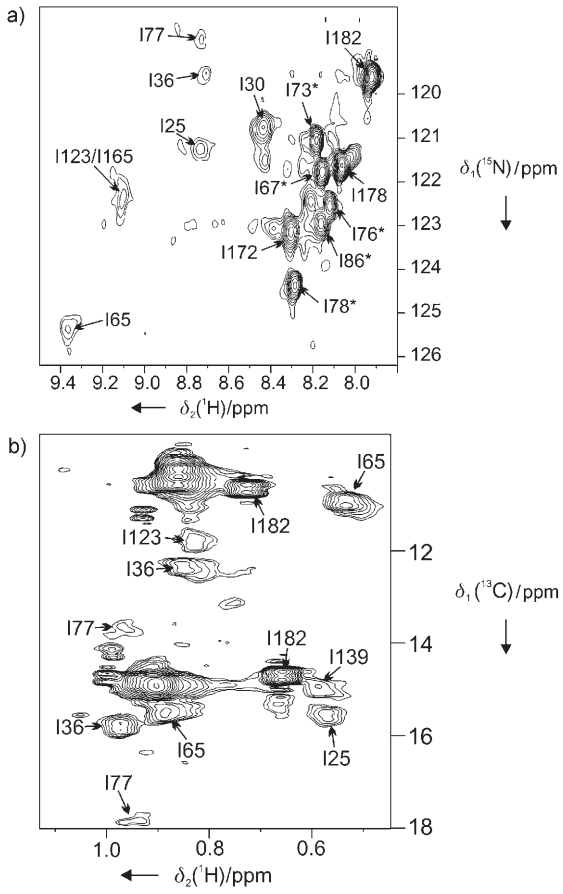

Figure 4. ^{15}N- and ^{13}C-HSQC spectra of a 200 μM solution of $^{15}N/^{13}C$-Ile-labeled dengue virus NS2B/NS3 protease at 25 °C, pH 6.9. NS2B resonances are marked with a star. a) Amide region of the I139V mutant. b) Methyl region of the I78V mutant. (The HSQC spectra of these two mutants were indistinguishable from those of the wild-type protein in the amide and methyl region, respectively.)

The resonance assignments obtained for the NS2B/NS3 protease attribute the intense and narrow cross-peaks to those parts of the protein for which no electron density was observed in the crystal structure,[9] indicating that these segments are highly mobile in solution. Furthermore, these data show that all structured parts of the protein are affected by line broadening. Nonetheless, the assigned cross-peaks are now available for NMR-spectroscopy-based ligand screening.

In conclusion, PCR amplification of DNA templates for high-yield protein synthesis in cell-free systems coupled with NMR spectroscopic analysis opens many attractive applications in structural biology, including protein production from complementary DNA (cDNA), optimization of domain boundaries and protein solubility, production of mutants, and introduction of stop codons for site-specific incorporation of non-natural amino acids.[12] Without the need for large-scale plasmid preparation, cloning, or protein purification, these applications have become highly practical.

Experimental Section

PCR and cell-free protein expression: PCR was performed by using Vent polymerase (New England Biolabs), which has 3′→5′ proof-reading activity and produces mostly blunt-ended PCR products as required for proper overhang generation. Further details are provided in the Supporting Information. PCR products were purified by using the QiaQuick PCR purification kit (Qiagen). *E. coli* cell-free coupled transcription/translation reactions were carried out as described,[5] with the exception that the plasmid templates were replaced by PCR-generated DNA templates by using 10 μg DNA per mL of cell-free reaction mixture. $^{15}N/^{13}C$-labeled isoleucine was from Cambridge Isotope Laboratories.

NMR spectroscopy: Dengue virus NS2B/NS3 protease was prepared from the CF40.gly.NS3pro construct used for the crystal structure determination,[9] except that it was preceded by the sequence MASMTG. Following cell-free synthesis in a reaction mixture (1 mL; typical protein yield 1.4 mg), the samples were dialyzed against NMR buffer solution (20 mM Tris-HCl, 50 mM NaCl; pH 6.9; Tris = tris(hydroxymethyl)aminomethane) and concentrated to 200 μL by using a Centricon-10 ultrafilter (Amicon). 10 % D_2O was added and NMR spectra recorded in 3-mm NMR tubes by using a Bruker Avance 800 MHz NMR spectrometer with a cryoprobe. Simultaneous $^{15}N/^{13}C$ HSQC spectra[11] were recorded in 2.5 h per spectrum by using $t_{1max} = 32$ ms and $t_{2max} = 93$ ms.

Received: December 27, 2006
Published online: March 22, 2007

Keywords: cell-free protein synthesis · high-throughput methods · NMR spectroscopy · polymerase chain reaction

[1] S. Yokoyama, *Curr. Opin. Chem. Biol.* **2003**, *7*, 39 – 43.
[2] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. M. Ono, P. Güntert, *Nature* **2006**, *440*, 52 – 57.
[3] K. Ozawa, P. S. C. Wu, N. E. Dixon, G. Otting, *FEBS J.* **2006**, *273*, 4154 – 4159.
[4] L. Guignard, K. Ozawa, S. E. Pursglove, G. Otting, N. E. Dixon, *FEBS Lett.* **2002**, *524*, 159 – 162.
[5] K. Ozawa, M. J. Headlam, P. M. Schaeffer, B. R. Henderson, N. E. Dixon, G. Otting, *Eur. J. Biochem.* **2004**, *271*, 4084 – 4093.
[6] T. Kigawa, T. Yabuki, Y. Yoshida, M. Tsutsui, Y. Ito, T. Shibata, S. Yokoyama, *FEBS Lett.* **1999**, *442*, 15 – 19.
[7] a) H. L. Yang, L. Ivashkiv, H. Z. Chen, G. Zubay, M. Cashel, *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 7029 – 7033; b) S. A. Lesley, M. A. D. Brow, R. R. Burgess, *J. Biol. Chem.* **1991**, *266*, 2632 – 2638.
[8] G. H. Hahn, D. M. Kim, *Anal. Biochem.* **2006**, *355*, 151 – 153.
[9] P. Erbel, N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan, U. Hommel, *Nat. Struct. Mol. Biol.* **2006**, *13*, 372 – 373.
[10] S. Melino, S. Fucito, A. Campagna, F. Wrubl, A. Gamarnik, D. O. Cicero, M. Paci, *FEBS J.* **2006**, *273*, 3650 – 3662.
[11] R. Boelens, M. Burgering, R. H. Fogh, R. Kaptein, *J. Biomol. NMR* **1994**, *4*, 201 – 213.
[12] J. Xie, P. G. Schultz, *Curr. Opin. Chem. Biol.* **2005**, *9*, 548 – 554.

