LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-TR-223848

Biosynthesis of the Cyclotide MCoTI-II using an Engineered Intein

J. Cantor, J. A. Camarero

August 21, 2006

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

Jason Cantor

DHS Internship @ Lawrence Livermore National Laboratories; Summer 2006

Project: *Biosynthesis of the Cyclotide MCoTI-II using an Engineered Intein*

Adviser: Julio Camarero

• The Project (Background, Motivation, Rationale)

Cyclotides are an emerging family of naturally occurring circular mini-proteins (~30-40 amino acids) characterized by six conserved Cys residues (forming 3 disulfide bridges) that create a topologically unique structure designated as a *cyclic cysteine knot* (CCK). The cysteine knot motif, which is embedded within the macrocylic backbone, is described as two disulfide bridges that form a ring that is penetrated by the third disulfide bridge. The cyclic backbone and CCK motif together confer cyclotides with a remarkable stability and resistance to proteolytic, chemical, and thermal degradation. Further, cyclotides are functionally diverse and display a wide range of functions including uterotonic activity, trypsin inhibition, cytotoxicity, neurotensin binding, anti-HIV, antimicrobial, and insecticidal activity. Together, these characteristics make cyclotides attractive candidates for both drug design and agricultural applications, both in their native forms and as molecular scaffolds for the incorporation of novel bioactivities. [1]

The ability to manipulate production of cyclotides within biological systems is critical for mutagenesis studies, production of grafted products, and the mass production of cyclotides with novel activities. My adviser's hope is to achieve this capability by employing recombinant DNA expression techniques to generate large combinatorial libraries of cyclotides. The advantage in creating a biosynthetic library (containing ~10^6-10^{10} members/library vs. chemically based libraries with typical values ranging from

~10^3-10^5 members/library) is that it can be lead to the *in vivo* application of biological screening and selection methodologies based on a specific clone's ability to affect certain cellular processes.

**

MCoTI-II is a natural cyclotide isolated from the seeds of the spinal gourd *Momordica cochinchinensis*. It belongs to the **trypsin inhibitory** subfamily of cyclotides and is very different in sequence to the other cyclotide subfamilies. In fact, it is more homologous to a family of acyclic trypsin inhibitors known as squash protease inhibitors [2]. One factor for choosing MCoTI-II was that my adviser believed that we could take advantage of the protein's ability to bind trypsin and therefore make isolation/purification of correctly folded MCoTI-II relatively straightforward compared to other cyclotides. With successful biosynthesis of MCoTI-II, the aspiration would be to then utilize the protein's unique structural characteristics in order to evolve cyclotides capable of inhibiting bacterial toxins (e.g. Lethal Factor from *Bacillus Anthracis*) or other harmful bio-agents.

**

Protein splicing is a naturally occurring, post-translational, self-processing that releases an **in**ternal pro**tein** sequence (Intein) from a protein precursor resulting in ligation of extein polypeptide sequences. Engineering of a naturally occurring intein and its subsequent inclusion into an expression vector allows for the biosynthesis of circular polypeptides through the intramolecular version of native chemical ligation (NCL) [3]. NCL is an extremely specific ligation reaction in which two fully unprotected polypeptides, one containing a C-terminal α-thioester group and the other a N-terminal

Cys residue, react chemoselectively under neutral aqueous conditions with the formation of a native peptide bond. It's well established that when these two reactive groups are located in the same precursor, the chemical ligation proceeds in an intramolecular manner thus resulting in the efficient formation of a circular polypeptide. Recent advances in protein engineering have also made possible the introduction of the C-terminal α-thioester group and N-terminal Cys residue into recombinant proteins. These developments make the use of NCL between recombinant fragments possible and the general term for this technology is Expressed Protein Ligation (EPL). [4]

**

My project focused on the biosynthesis of the cyclotide MCoTI-II in *E. coli* based on the use of an engineered intein for *in vivo* circularization of a linear precursor via EPL. This is a necessary technology to apply in order to allow the backbone cyclization of the MCoTI-II linear precursors inside *E. coli*. Since MCoTI-II contains six conserved Cys residues, there are six potential linear precursors of the protein – each unique in which Cys residue acts as the N-terminus. As a result, it is possible that protein biosynthesis in *E. coli* could be achieved with varying degrees of success depending on which precursor is being used. Thus, my specific role in this project involved exploring the relative degrees of expression, folding, and cleavage by the engineered intein, for three of the MCoTI-II linear precursors.

- The Project (My Specific Work)

As mentioned above, I worked with three different linear precursors of the cyclotide MCoTI-II, which I will designate MC-1, MC-4, and MC-6 (where the number

indicates which Cys residue acts as the N-terminus and thus the point of cyclization). Most of the first half of my summer involved the molecular cloning of the three precursors above into each of two different expression vectors (pTXB1, pTYB1). Each expression vector contained an engineered intein and chitin-binding domain (CBD) upstream of the cloned gene, but differed in the identity of the engineered intein – thus, we could determine whether a specific intein could also influence biosynthesis success. Cloning each precursor involved typical molecular biology techniques such as DNA oligo hybridization, digestion, ligation, and colony PCR screening. Ultimately, I obtained 2-3 positive clones for each precursor in each expression vector.

The second half of my summer focused on the actual expression and subsequent analysis of MCoTI-II precursors. The goals during this time were to first determine whether biosynthesis of the protein was even plausible in *E. coli*, and then to begin determining how to optimally obtain the correctly folded (and thus active) protein. Expression and purification of each precursor (in either expression vector) involved techniques such as SDS-PAGE, sonication for cell lysis, and preparation of chitin bead columns used to trap proteins that were fused to the chitin binding domain mentioned earlier (these would be cyclotide-intein-CBD fusion products). Once precursors were isolated via chitin bead columns, I subjected small samples to mix with various thiols, causing cleavage of any remaining uncleaved cyclotide-intein fusions. Following that step, my adviser and I used RP-HPLC and ES-MS to characterize our products and to determine the presence of correctly folded MCoTI-II. In addition, as mentioned earlier, we could further confirm the presence of correctly folded MCoTI-II by taking advantage of its native activity of trypsin inhibition. Therefore, I also prepared a trypsin-sepharose

column through which we could run a sample and again characterize the binding products via RP-HPLC and ES-MS.

- The Project (My Achievements, Contributions, etc.)

My first achievement was the successful molecular cloning of MC-1, MC-4, and MC-6 into two different expression vectors in a short period of time. Without this step, no further analysis could've been conducted. Next, I was able to confirm successful expression of MCoTI-II from the linear precursors MC-1 and MC-6 (both in pTXB1). As mentioned earlier, confirmation was achieved first characterizing thiol-cleaved chitin bound samples via SDS-PAGE, RP-HPLC, and ES-MS. Then I obtained even further evidence that we in fact had active protein using a trypsin-sepharose column. Products that bound to this particular column were analyzed using RP-HPLC and ES-MS, and an identical retention peak and mass associated with the correctly folded protein were again obtained. This set of experiments proved in fact that MCoTI-II could be successfully biosynthesized in the *E. coli* cytoplasm. When analyzing the same two precursors (MC-1, MC-6) in pTYB1, it appeared that *in vivo* cleavage was much better (one dominant SDS-PAGE band rather than two) but overall expression was lower and we had a difficult time isolating and thus characterizing any protein that may have been present. The final linear precursor, MC-4, had very little expression overall in both expression vectors and, even after chitin binding purification, SDS-PAGE bands were extremely faint. I even analyzed other MC-4-positive clones from the original ligation plating and the SDS-PAGE results were consistent with the original faint bands observed.

With successful expression of MCoTI-II from two different linear precursors, my next step was to begin thinking about optimization of *in vivo* protein biosynthesis. There

are a number of different parameters that could be tuned in an attempt to optimize expression, degree of intein cleavage, etc. and these include: induction time, inducer concentration, and host cell line. My final contribution was transforming the plasmid containing the MC-6-intein-CBD fusion (in both expression vectors) into another host cell line that we think will potentially promote higher expression levels, and then subsequently performing a direct expression study with the original cell line used. I think my work has contributed a considerable amount to a MCoTI-II biosynthesis publication that my adviser hopes to have out by the end of this year. I have shown that we can successfully express an active folded protein in the *E. coli* cytoplasm and have also laid the groundwork for direct comparison between different linear precursors and different conditions for expression. These were both very big milestones in the project's big picture. Some of the remaining work to be done prior to publication includes obtaining the cyclotide's trypsin inhibition kinetics, as well as analyzing direct expression studies in which select parameters are tuned (see above) to determine the best possible conditions to achieve the highest degree of *in vivo* intein cleavage and the highest possible levels of expressed active protein.

- Impact on academic/career planning

My internship experience didn't particularly influence my academic or career planning. I am already two years into my graduate studies and therefore my academic plans were essentially set before I came to the lab. Further, I think I'd ultimately like to enter academia after I receive my degree so that I can have the opportunity both to do research and to teach. I think one major positive of this experience is my having the

knowledge of another interesting area within my field and the ability to potentially apply some of the concepts I've learned about to future projects. Furthermore, I have now had the opportunity to complete a co-op in industry and an internship in a national laboratory. Taken with my graduate studies, I have had the chance to see how three very different types of organizations work, which is beneficial in helping choose different options in my career path.

**

This summer, I attended a number of scientific seminars hosted by LLNL that covered a wide array of topics. I think the biggest benefit from attending these seminars was observing the most useful ways to give an effective presentation. One speaker may have been very good at slide layout while another excelled with his slide transitions, but ultimately, I was able to determine an approach in which I could combine the positive aspects from nearly every speaker in order to give an effective presentation of my own in the future. LLNL also had a weekly DHS required talk in which various projects within the lab were discussed. Topics included border security, a nuclear program overview, and small-scale PCR, but most times I felt as if a large part of the intern audience was left confused by the technical details discussed in the talks. I thought that the broad topics such as border security were by far the most valuable since everyone could more or less comprehend the material and also form opinions that led to interesting dialogue about ideas that most of us generally never have an opportunity to discuss. For this reason, I also thought that the DHS talk in which Maureen McCarthy spoke to LLNL and Sandia interns was above average because, although she was still theoretically giving a talk, it felt more like a discussion forum on an assortment of broad but important topics.

Unsurprisingly, it was more engaging to be a part of talks that led to interesting questions and debate rather than ones where only a handful of interns could grasp the material or where the attention to detail began to outweigh the big picture.

• Ideas for future DHS Research

I think a few main areas of research that could receive more attention are: border security, treating persons afflicted with bioterror agents, and high-throughput screening procedures at airports.

I think that the DHS talk on border security this summer really opened my eyes (and others as well) to the lack of an efficient and reliable means to effectively secure our borders. We learned that most of our southern border is unprotected and those small areas that are being observed are normally by just 1-2 officers who are not fully equipped to deal with a lot of the potential encounters that could occur. In addition, we were told that due to the mainly rural nature of the northern border, it becomes, in essence, impossible to secure. I feel that significant research should be committed to improve upon what appears to be extreme shortcomings in securing our borders. This could mean simply increasing manpower along our borders, looking into more effective preventative means, or changing the approach that is taken with those who wish to illegally enter the country. For instance, we heard that in most cases, all it takes is persistence by those who want to cross the border since most times, if caught, they are simply sent back to their country only to basically try again and again. By whatever means, I think this is definitely an area that should receive more research attention; both hard science-based and social science-based.

Additionally, it appeared to me that a lot of research is being done in order to potentially screen for bioterror agents in large areas (there were multiple talks this summer about small-scale PCR for instance), but I think we should also be devoting more resources into treating those who do potentially become affected by a bio-agent. I can't remember any seminars or DHS talks that were concerned with developing vaccines or antibodies or therapeutics towards bioterror agents, although my particular project this summer could be potentially applied towards these means in the long run. It seemed, based on the seminars and talks I attended this summer, that there was much more focus on identifying or screening for the presence of agents relative to actually treating them if necessary. But again, I can only say this based on the ten-week sample of talks I was able to attend.

Finally, in light of recent current events, I think a stronger focus should be placed on large-scale high throughput screening not only of bio-agents, but also other potential destructive or debilitating forces as well, specifically in airports. Research could be placed into optimizing a process to screen passenger carry-on items such as water bottles or shampoo bottles, for example, to test for explosive materials. I think the public would be in favor of a high-throughput screening process that would, in a best case scenario, ensure their safety but also not cause extremely long airport delays. Extensive research must be placed creating some sort of screening protocol that could ultimately encompass a broad range of materials – not just liquids for example. Only with said process in tact, can we truly secure our commercial flights, since they seem to continue to be targeted.

In summary, I think that DHS should focus more attention towards securing our nation's borders (both land and sea), developing treatments against the wide scope of

bioterrorist agents that exist, and securing our commercial flights against an ever expanding repertoire of threats.

References

[1] Kimura et al., *Angewandte Chemie*, **2006**, 45, 973-976

[2] Avrutina et al., *Biological Chemistry*, **2005**, 386, 1301-1306

[3] Camarero et al., *Bioorganic & Medicinal Chemistry*, **2001**, 2479-2484

[4] Kimura et al., *Protein & Peptide Letters*, **2005**, 789-794