Efficient production of mussel adhesive proteins using *e. coli* and crescentus *caulobacter*

Nora Yucel, Daniel Choi, Jim Einholtz, Parijata Mackey, Rob McConeghy, Damon Wang, Soren Yeadon: University of Chicago

Abstract

Our research group is interested in engineering E. coli and caulobacter to express recombinant mefp-5, fp-151 and which are found naturally in mussels (Mytilus edulis and Mytilus galloprovincialis). These mussel foot proteins are strong bioadhesives and powerful anti-biofouling agents, with applications for biomaterials and biomedical research . We aim to produce results that will achieve the initial goals of genetic engineering, as well as further the conceptual goals of synthetic biology. The final goal of this research is to produce 5 biological Systems that meet the specifications of the Standard Registry of Biological Parts: 1) Expression of recombinant MAPs 2) secretion of MAPs 3) Surface display of MAPs 4) expression of tyrosinase and 5) concomitant production of tyrosinase with MAPs.

Introduction to synthetic biology

Synthetic biology is a novel approach to engineering biology, focusing on organizing genetic information into "standard parts," comparable to interchangeable parts in manufacturing industries such as automobiles or electronics. The fundamental goal of synthetic biology is the organization of genetic information into standardized regions of code, that will create the functional equivalent of an open-source biological programming language. As opposed to ad hoc genetic engineering methods, that do not necessarily emphasize the organization of genetic data, synthetic biology aims to move towards a scalable engineering framework.

The basic ideas underlying synthetic biology are really quite simple:

1. Automated (DNA) Construction
2. Standards of Abstraction
Abstraction is the **process of generalization by reducing the information content of a concept or an observable phenomenon, typically in order to retain only information which is relevant for a particular purpose**. It is a mechanism and practice to reduce and factor out details so that one may focus on a few concepts at a time. For example, abstracting "a well-worn, bouncy basketball" to simply "a ball" retains only the information on the attributes and behavior of a general ball. **Abstractions allow us to hide complexity, and utilize a greater amount of parts in a simpler design**.

Registry of Standard Biological Parts

In the summer of 2004, the MIT Registry of Standard Biological Parts was created to support the design and construction of biological systems through interchangeable parts. It records and indexes biological parts that are currently being built and offering synthesis and assembly services to construct new parts, devices, and systems. Starting with about 100 basic parts it now has 700 available, and 2000 defined parts

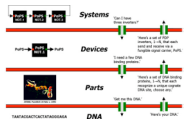
According to the Registry descriptions the parts in it are not simply segments of DNA, they are functional units. Those functions are being specified and parameters measured. Proper interoperation of a family of devices will depend on compatible parameters. The Registry stresses the need for composable system design and standards of measurement for gene-expression based systems.

Assembly of parts into devices and systems is performed using traditional cloning techniques with a common set of restriction sites that allow easy composition of composite devices that, in turn, can themselves be used as parts. Simultaneous parallel assembly lets us build many biological systems quickly.

PoPs
Or Polymerase Per Second: PoPS measure the rate of transcription at the boundaries of a part. Seen as the new "standard unit of measurement" the Registry is in the process of characterizing all parts in terms of PoPS

iGEM
The Registry hosts a competition for undergraduates to challenge undergraduate students to put to use the idea that biological engineering can be made reliable through the use of standardized, well-documented parts (iGEM). The results from this summer's research will be submitted this year as the new University of Chicago iGEM

Example: the "Repressilator"

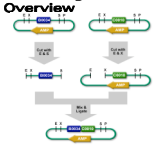


The latest (working) definitions from the Registry of Standard Biological parts are as follow:

DNA = low level DNA sequence ... a string of A,T,G,C's
Part = RBS, CDS, promoter, terminator; a piece of DNA in BioBricks format that has a specific function.
Device = inverter; something else with a higher level function; composed of multiple parts
System = ring oscillator; composed of many devices

Creating a biobrick

Overview



The BioBrick is flanked by restriction sites, comprised of the BioBrick Prefix and Suffix, between the EcoRI and XbaI cutting sites on the left and SpeI and PstI on the right. For the assembly shown, the blue part is cut out of its plasmid with the enzymes EcoRI and SpeI. The resulting fragment is called the insert because it will be inserted into the plasmid containing the other part.
In a separate reaction, a gap is cut in the plasmid containing the green part using EcoRI and XbaI. Using gel electrophoresis, the insert for the blue part and the cut plasmid containing the green part are purified and the unwanted fragments discarded. The purified insert and cut plasmid are mixed under the right conditions to allow the E sticky ends to come together and the S sticky ends to come together with the X sticky ends. Once this happens, the DNA backbone is re-ligated and the resulting blue-green part can be transformed into E.coli cells. These cells may be grown to produce as much of blue.green part as desired.

mefp-5 for e. coli

E. coli-Optimized Protein Coding Region:

```
1 atg aag ctt agc tgc atc gtt ctg gtt ctg
31 ttt ctg gtc acc ctg gcg gcg tat agc gat
61 gtg ggt tct tcg tct tcc gaa gag tat aaa
91 ggc ggt tac tac cca ggt aac gct tac cac
121 tac cac agc ggc ggt tct tac cac ggc agc
151 ggt tac cac ggc ggt tac aaa ggg aag tat
181 tac ggc aaa ggc aaa aaa tac tac tac aaa
211 tac aag aaa tcc ggc aaa tac aaa tac ctg
241 aaa aaa ggt cgt aaa tat cac cgt aaa ggt
271 tat aag tac tac ggt ggt agt agc
```

Biobrick Prefix:
GAATTCGCGGCCGCTTCTAG [**ATG**
Remaining CDS]
EcoRI NotI XbaI

Biobrick Suffix:
TACTAGTAGCGGCCGCTGCAGAGATCT
SpeI NotI PstI BglII

Final Sequence for Completed BioBrick:

```
GAATTCGCGGCCGCTTCTAGGCATGAAGCTTAGCTGCATCGTTC
TGGTTCTGTTTCTGGTCACCCTGGCGGCGTATAGCGATGTGGGT
TCTTCGTCTTCCGAAGAGTATAAAGGCGGTTACTACCCAGGTAA
CGCTTACCACTACCACAGCGGCGGTTCTACCACGGCAGCGGT
TACCACGGCGGTTACAAAGGAAGTATTACGGCAAAGCCAAAAA
ATACTACTACAAATACAAGAACTCCGGCAAATACAAATACCTGAA
AAAGGTCGTAAATATCACCGTAAGGTTATAAGTACTACGGTG
GTAGTAGCTACTAGTAGCGGCCGCTCCAGGATCT
```

mefp-5 for caulobacter

Caulobacter-Optimized Protein Coding Region:

```
1 atg aag ctg tcg tgc atc gtc ctg gtc ctc
31 ttc ctg gtc acg ctg gcg gcg tac tcc gac
61 gtg ggt tcg tcg tcc ggc gag tac aag
91 ggc ggc tac tac ccg ggt aac ctg tac cac
121 tac cac agc ggt ggc tct tac cac ggc agc
151 ggc tac cac ggt ggc tac aag ggc aag tac
181 tac ggt aag ggc aag aag tac tac tac aag
211 tac aag aac tcg ggc aag tac aac tac ctg
241 aag aag ggc cgc aag tac cat cgc aag ggc
271 tat aag tac tac ggc ggc tcg tcg
```

Biobrick Prefix:
GAATTCGCGGCCGCTTCTAG [**ATG** Remaining CDS]
EcoRI NotI XbaI

Biobrick Suffix:
TACTAGTAGCGGCCGCTGCAGAGATC
T
SpeI NotI PstI BglII

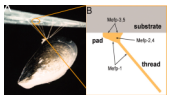
Final Sequence:

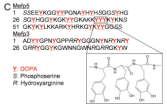
```
GAATTCGCGGCCGCTTCTAGGCATGAAGCTGTCGTGCAT
CGTCCTGGTCCTCTTCCTGGTCACGCTGGCGGCGTACTCC
GACGTGGGCTCGTCCTCGTCGGAGGAGTATAAGGGCGGC
TATTACCCAGGCAATGCCTATCACTATCATAGTGGAGGCT
CTTACCACGGCAGCGGCTATCACGGTGGCTACAAGGGCA
AGTACTACGGTAAGGCAAGAAATATTACTATAAGTACAA
GAACTCGGGCAAGTACAAATATCTCAAGAAGGCCCGCAAA
TACCATCGCAAGGGCTATAAGTACTACGGCGGCTCGTCG
ACTAGTAGCGGCCGCTCCAGGATCT
```

Goals

We are using multiple approaches and two organisms to accomplish our goal of mussel adhesive protein (MAP) production.The first is expression of MAPs through transformation of recombinant plasmid DNA. Our second track is secretion, through co-expression of recombinant a pathogenic secretion system in E. coli and manipulation of the inherent secretion mechanism in caulobacter crescentus, a DESCRIPTION. Our third approach is directed at producing adhesive E. coli, rather than pure "glue." To do this, we will engineer overexpression of MAPs and direct self-destruction via apoptosis. In addition, we hope to utilize available biobricks supplied by iGEM to engineer surface display of full-length MAPs and/or model short peptides.

Mussel adhesive proteins





Purified protein

E. Coli

1. Expression of recombinant mefp-5 biobrick in E. coli.
   - Ligate our mefp-5 bio-brick DNA into a simple IPTG-inducible plasmid for expression in the cytoplasm. We will treat both the crude lysate and purified protein, as obtained through a his-tag affinity column, with tyrosinase to induce conversion of tyrosine to DOPA. This procedure was carried out *y Hwang* et al with Mgfp-5, the analogous protein in Mytilius galloprovincialis, with limited efficiency. As such, this will serve as a control rather than a representative method of MAP production.
   - We will also try expression of longer length fp-151, a recombinant protein consisting of mefp-5 flanked on either side by 6 repeats of the decapeptide repeat from mefp-1. This construct has been demonstrated to have better expression efficiency and adhesive ability.
2. Expression of hexapeptide and decapeptide repeats from mefp-1.
   - The hexa and decapeptide consensus repeats Ala-Lys-Pro-Thr-Tyr-Ala and Ala-Lys-Pro-Ser-Tyr-Pro-Pro-Thr-Tyr-Lys make up over 80% of mefp-1, the latter appearing more than 80 times within the protein. Their use as model peptides is well documented and they have been shown (along with other repeating sequences) to possess similar adhesive properties to full length MAPs.

Hexapeptide repeat: Co-expression with recombinant secretion system taken from salmonella pathogenecity island II.

- *Wilson* et al describe the transfer of the salmonella pathogenecity island II, including type III secretion system onto a single plasmid and subsequent functional expression in non-pathogenic E. coli. Salmoella pathogenecity island II (or SPI-2) is a type III secretion system used by Salmonella typhirium to direct secretion of pathogenic proteins into the extracellular media in a single step. Following expression of the SPI-2 secretion system, we will direct secretion of various MAPs, including mefp-5, fp-151.

Caulobacter crescentus

Caulobacter crescentus is a Gram-negative bacteria easily optimized for secretion. Adapted to surviving under a minimal nutrient environment, it is composed of two parts: a long tubular stalk which secretes adhesive proteins, allowing it to attach to surfaces, and a "swarmer cell" which searches for nutrients while affixed. With help from Dr. Sean Crosson, we will direct secretion of recombinant mefp-5, fp-151 and model peptides bearing the C-terminal caulobacter secretion tag . Under nutrient-poor conditions, the proteins are secreted as a "scum" at the top of the media, which we will extract and purify.

Adhesive bacteria

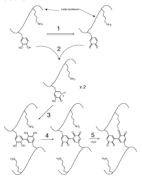
1. Self-lyse model
   - Ligate our mefp-5 bio-brick DNA into a simple IPTG-inducible plasmid for expression in the cytoplasm. We will treat both the crude lysate and purified protein, as obtained through a his-tag affinity column, with tyrosinase to induce conversion of tyrosine to DOPA.
   - Biobrick part BBa_I745006
2. Surface display of model peptides
   - Part BBa_I728500

A device for **CPX Terminal Surface Display Protein with Polystyrene-Binding Peptide**
Circular-Permutated OmpX gene called CPX, used to display surface proteins. CPX is linear with both the N and C termini displayed extracellularly. The insert in the iGEM part is a polystyrene binding peptide, flanked by restriction sites, so we can swap in our own peptides at will

DOPA

DOPA is a key residue in all MAPs. They are contained in every type of mussel-foot protein, with concentrations greatest in proteins at the plaque-surface interface. Mefp-5 and mefp-3 show the highest levels (30 mol% and 20 mol% respectively, while mefp-1, another model MAP stands at 13 mol%.

Multiple factors account for the "stickiness" of DOPA. Small cations like Mg++ stimulate cross-linking via chelate-mediated or covalent coupling. However, this is only one of many pathways DOPA can follow. Studies with DOPA-containing polypeptides show that high oxidation, leading to the formation of ortho-quinone DOPA, forms cross-links. DOPA is also very sensitive to pH, rapidly converting to it's ortho-quinone form in alkaline conditions.

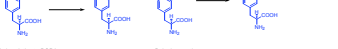


"Proposed pathway of peptidyl-DOPA-based cross-linking in plaque precursor proteins (adapted from Burzio and Waite, 2000). In step 1, dopa is converted to dopa-orthoquinone by catalysis or auto-oxidation. Dopa-quinone reacts with dopa by a reverse dismutation process (2) to form two semiquinone free radicals. The free radicals couple through a ring- /ring c-c bond (3). McDowell *et al.* (1999) have proposed a 5, 5' coupling based on solid-state 13C NMR analysis of plaques. Di-dopa can easily be reoxidized once or twice to a diquinone (4). Such oxidation steps are suggested by mass spectrometry (Burzio and Waite, 2000). Involvement of lysine in Schiff base additions does not drive peptide cross-linking, but may occur after di-dopa cross-links have been formed (5). This is based on losses of lysine in oxidized peptides (Burzio and Waite, 2000) "

Waite, Herb. "Adhesion a la moule," Integrative and Comparative Biology 2002 42(6):1172-1180

Tyrosinase

Tyrosinase or tyrosine hydroxylase is an oxygen transferring, copper metalloprotein that can act on diphenols and monophenols. It acts on phenols to catalyze orthohydroxylation to catechols, and further catalyzes dehydrogenations of catechols to orthoquinones. Accordingly, it specifically hydroxylates tyrosine to produce 3,4-dihydroxy-L-phenylalanine or L-DOPA as well as its orthoquinone form.



Hydroxylation to DOPA Dehydrogenation

This enzyme is a key enzyme in melanin synthesis, and is found in a variety of fungi, animal and plant species.Since DOPA is a non-native peptide for E. coli and caulobacter, and MAPs obtained from both species must be modified to fully activate the adhesive, we will use commercial tyrosinase to modify our MAPs.

1. Preliminary experiments: We will first test efficiency of our commercial tyrosinase with synthesized short-peptides
2. Modification of purified protein: Add tyrosinase to purified protein or crude cell lysate to fully activate MAPs
3. Co-expression of tyrosinase: Ultimately, we hope to express tyrosinase in conjunction with MAPs, in order for completely autonomous activated MAP production.

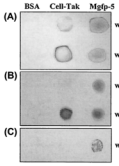
Methods

Modification

- Modification of tyrosine residues. To investigate adhesion, purified recombinant fp-151, Mefp-5, and the synthesized 6 and 10-residue length peptides are resolved in 5% acetic acid buffer to prevent auto-oxidation of DOPA residues to o-quinone under basic pH conditions will be modified with tyrosinase to convert tyrosine residues into DOPA.
  - As a positive control, Cell-Tak is a commercially available naturally extracted M. edulis mussel adhesive protein mixture of Mefp-1 and Mefp-2, that already contains DOPA residues in 5% acetic acid buffer.

Characterization of MAPs

- SDS-PAGE and Western blot analysis. Samples will be resuspended in protein sample buffer (0.5 M Tris-HCl [pH 6.8], 10% glycerol, 5% sodium dodecyl sulfate [SDS], 5% ß-mercaptoethanol, 0.25 % bromophenol blue) and heated to 100°C for 5 min. After centrifugation for 1 min, proteins will be separated by SDS-polyacrylamide gel electrophoresis (PAGE) (15% [wt/vol] polyacrylamide) and then detected using Coomassie blue staining (Bio-Rad) or silver staining (Bio-Rad), or Western blotted. The membrane will be scanned, and its image analyzed using gel-analyzing software.
- Investigation of coating of various substances. The ability of recombinant Mefp-3,5 and Mgfp-5 to coat the following surfaces will be investigated: glass slide, poly(methyl methacrylate) plate, polystyrene plate, commercial silicone-based antifouling agent (SigmaGlide; Sigma Coatings)-coated slide, and aluminum plate. Each material surface is to be cleaned by washing with distilled water several times and drying with nitrogen gas. A 10-µl drop of 1.44-mg/ml protein solution will be added to each material surface and incubated in a humid environment for 12 h at 25°C. After being dried, each surface shall be washed thoroughly with deionized water for 2 h with shaking. The coating by each protein will be visualized using Coomassie blue staining.
  - As a control we will use commercial Cell-tak, as well as an aqueous solution of dopamine (commercial), pH 8.5.



Sample data from coomassie-blue staining of Mgfp-5 expressed by E. coli, as obtained by *Hwang*, et al. Methods will (initially) be as described here for our recombinant protein.

Figure 3:
Recombinant Mgfp-5 coating of a glass slide (A), a poly(methyl methacrylate) plate (B), and a commercial silicone-based antifouling agent SigmaGlide-coated slide (C). BSA was used as a negative control, and commercial Cell-Tak was used as a positive control. Modification of each protein sample was performed using 10 U of tyrosinase per µl for 6 h at room temperature. A 10-µl drop of 1.44-mg/ml protein sample was placed on each surface, the surfaces were incubated for 12 h at 25 °C, and the slides were washed with deionized water for 2 h. Coated proteins were visualized using Coomassie blue staining.

Future directions

Currently, most of our work is in the planning stage. However, once we are able to achieve some form of MAP production—be it through simple expression and cell lysing, secretion or surface display, our second goal will be to express concomitantly express tyrosinase. We will first attempt to express tyrosinase within the cell, then aim at secretion or a cell-lyse model. Once we have gathered preliminary information about reaction kinetics between tyrosine and tyrosine through studies with mefp-1 hexa- and deca-peptide repeats, we will attempt to engineer one of the following systems:

1. Concomitant secretion of MAPs and tyrosinase
2. Expression of tyrosinase and MAPs followed by apoptosis
3. Surface display of MAPs with cytoplasmic production of tyrosinase followed by apoptosis.

Even further into the future, we are looking towards directed production of sticky plaques in the presence of some stimulus. For example, create a pathway that turns on production of MAPs/tyrosinase along with thread proteins in the presence of petroleum as a possible means of bioremediation, or as a bio "plug."

References

Waite, Herbert. Integrative and Comparative Biology 2002 42(6):1172-1180

Hwang, et al. Applied and Environmental Microbiology, June 2004.

Registry of Biological Parts. http://partsregistry.org/Main_Page

Team website at t http://www.freewebs.com/genehackers/

Acknowledgements

Thank you to everyone in the lab
Laura Stone and Thomas Graham
Dr. Schonbaum, Harper
PCBio/Beckman fellows and NIH