Synthetic Proteome

Cooperative Research and Development Final Report

CRADA Number: CRD-16-605

NREL Technical Contact: Yannick Bomble

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-2700-76734
April 2020

Synthetic Proteome

Cooperative Research and Development Final Report

CRADA Number: CRD-16-605

NREL Technical Contact: Yannick Bomble

Suggested Citation
Bomble, Yannick. 2020. *Synthetic Proteome: Cooperative Research and Development Final Report, CRADA Number CRD-16-605.* Golden, CO: National Renewable Energy Laboratory. NREL/TP-2700-76374. https://www.nrel.gov/docs/fy20osti/76374.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-2700-76734
April 2020

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Bioenergy Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

Cooperative Research and Development Final Report

Report Date: 2/27/20

In accordance with Requirements set forth in the terms of the CRADA agreement, this document is the final CRADA report, including a list of Subject Inventions, to be forwarded to the DOE Office of Science and Technical Information as part of the commitment to the public to demonstrate results of federally funded research.

Parties to the Agreement: FedIMPACT, LLC

CRADA number: CRD-16-00605

CRADA Title: Synthetic Proteome

Joint Work Statement Funding Table showing DOE commitment: No NREL Shared Resources

Estimated Costs	NREL Shared Resources a/k/a Government In-Kind
Year 1	$.00
TOTALS	$.00

Abstract of CRADA Work:
There are several areas of interest with regard to advancing renewable energy technology and increasing Participant's use of renewable energy. Participant would like to collectively work with the National Renewable Energy Laboratory on a variety of projects as outlined in part by the statement of work below.

Summary of Research Results:

Phase 1. Design, assembly and testing of basic prototype Synthetic Proteome

- Tasks 1.1-1.2: Linker, Fok1a and Fok1b proteins isolated, stable and ready for assembly
- Tasks 1.3-1.4: Functioning Synthetic Proteome
- Tasks 1.5-1.6: Testing and demonstration of functional Fok1 dimer/DNA synthetic proteome complex

Phase 2. Design assembly and testing of multi-protein prototype Synthetic Proteome

- Tasks 2.1-2.2: Proteins isolated, stable and ready for assembly
- Task 2.3: CAS-9 fusion proteins isolated, stable and ready for assembly
- Task 2.4: linkable sgRNA isolated, stable and ready for assembly
- Task 2.5: Fusion proteins and nucleic acid assembled into Synthetic Proteome
- Task 2.6: Primary testing of CRISPR type Synthetic Proteome compare to literature
- Task 2.7: Demonstrate stable functioning CRISPR-type reactions

Phase 1. Design, assembly and testing of basic prototype Synthetic Proteome

Tasks 1.1-1.2: Linker, Fok1a and Fok1b proteins isolated, stable and ready for assembly

Summary

In this task, we tried to construct and express three kinds of proteins to be ready for tasks 1.3-1.4: Linker (20Lmod), Fok1a (wild type), and FokIb (Fok1 mutant; D483A and R487A). All three proteins were successfully expressed and purified at predicted size. The purified Fok1a exhibits the expected cleavage activity on DNA substrate. However, the Fok1b did not show any cleavage activity, as expected. We successfully completed these tasks for the subsequent experiments to construct a functional synthetic proteome.

Procedures

Cloning and construction of expression strains

DNA sequences for Linker (20Lmod), Fok1a (wild type), and FokIb (Fok1 mutant; D483A and R487A) were codon optimized and cloned into a pET22b(+) vector (GenScript, Piscataway, NJ). The sequence for a hexa-histidine tag was placed at the C-terminus of the constructs. *E. coli* BL21(DE3) LysY/Iq (NEB) was used for all protein expression experiments. Recombinant strains were grown in LB broth supplemented with kanamycin (50 μg/mL). Cultures were induced at 15°C with 0.2 mM IPTG when OD600=0.6. Cultures were centrifuged at 5,000xg for 15 min when OD600≥1.2. The Fok1a and Fok1b mutants were successfully expressed only co-expression with M.FokI.

Protein purification and running the SDS-PAGE gel.

The frozen cell pellets were thawed at room temperature with equal volume of buffer A (50 mM Tris pH 7.5, 100 mM NaCl, 10 mM imidazole, 0.1 mM thiamine pyrophosphate (TPP), 0.5 mM dithiothreitol (DTT) and 1 mM $MgCl_2$) and lysed with lysozyme and sonication. One mg/mL lysozyme (Hampton Research, Aliso Viejo, CA), 1.0 U/mL Pierce Universal Nuclease (Thermo Scientific, Rockford, IL) and EDTA-free protease inhibitor (Thermo Scientific, Rockford, IL) accordingly to manufacturer instructions were added in the lysis mixture and incubated for 30 min at room temperature with occasional vortexing. Sonication was done at room temperature for 2

min using a Branson 5510 water bath sonicator (Branson Ultrasonics Corporation, Danbury, CT). Cell debris was removed by centrifugation at 15,000×g for 15 min. The supernatant was loaded onto an eight mL HisPur Cobalt column (Thermo Scientific, Rockford, IL) using an Akta FPLC system (GE Life Sciences, Piscataway, NJ) with buffer A (50 mM Tris pH 7.5, 100 mM NaCl, 20 mM imidazole, 0.1 mM TPP, 0.5 mM DTT, and 1 mM $MgCl_2$). After loading and washing the unbound proteins from the column, protein samples were eluted using 100% of Buffer B (50 mM Tris pH 7.5, 100 mM NaCl, 250 mM imidazole, 0.1 mM TPP, 0.5 mM DTT, and 1 mM MgCl2). Final purification was performed by size-exclusion chromatography using a HiLoad Superdex 200 (26/60) column (GE Healthcare, Piscataway, New Jersey, USA) in buffer C (20 mM Tris pH 7.5, 100 mM NaCl, 0.1 mM TPP, 0.5 mM DTT, and 1 mM $MgCl_2$). Purified proteins were analyzed by SDS-PAGE using a 4−12% NuPAGE Bis-Tris Gel (Invitrogen, Carlsbad, CA) run at 150 V for 55 min in MOPS SDS buffer. The gel was stained with Colloidal Blue (Invitrogen, Carlsbad, CA) to visualize proteins which are stained blue.

FokI enzymatic assay

The DNA substrate used for this task were generated by PCR amplification using Q5 High-Fidelity DNA polymerase (NEB) with DCB229 and DCB230 primers. The cleavage activity of FokI variants was assayed by incubating 1 µg of protein at 37°C in FokI reaction buffer (pH 7.9 ; 20mM Tris-acetate, 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM DTT, 100 mg/ml BSA) for 20 mins. The sizes of cleavage products were analyzed by running the samples on 1.0% agarose gel in 100V.

Results

The three proteins, Linker, Fok1a and Fok1b, were successfully expressed, purified and exhibited the predicted size (Fig.1). The Fok1a exhibited very specific cleavage activity, but Fok1b was not able to cleave the DNA substrate, as expected (Fig. 2). And also the 20Lmod for scaffoldin was successfully expressed and ready to be used assembly of synthetic proteome complex.

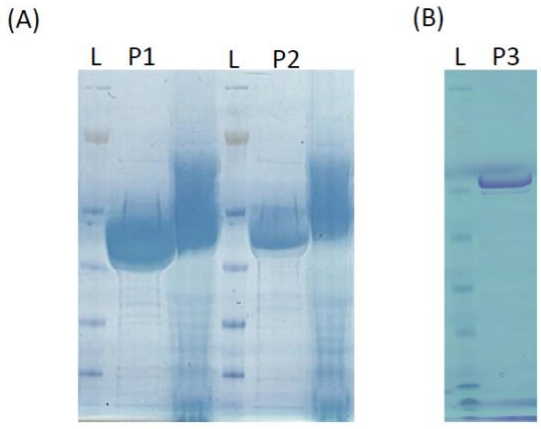


Fig.1 SDS-PAGE gel (4 to 12% gradient) stained with Colloidal Blue showing. L, MW standards (BioRad); P1, Fok1a; P2, Fok1b; P3; linker.

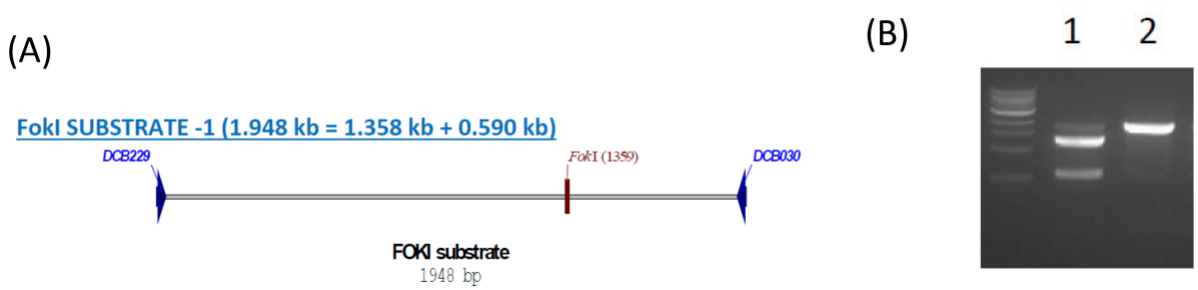


Fig.2 Comparison of DNA cleavage by FokIa (wild type) and Fok1b (D483A, R487A; Dimerization-defective FokI variant). (A) The DNA substrate used for this cleavage assay. (B) The agarose gel running after cleavage reaction for 20 mins as described at procedure section.

Tasks 1.3-1.4: Functioning Synthetic Proteome

Summary

We tried to construct four kinds of synthetic fusion proteins, which are Fok1a-AcDoc, Fok1a-BcDoc, Fok1b-BcDoc, and Fok1a-CtDoc, to complete the synthetic proteome for Phase 1. All of four synthetic fusion proteins were successfully expressed at predicted size. Both Fok1a-BcDoc and Fok1a-CtDoc exhibit the expected cleavage activity on DNA substrates. However, the Fok1a-AcDoc lacked any cleavage activity. The Fok1b-BcDoc did not exhibit any FokI cleavage activity, as expected. We were able to successfully express and purify two functional FokI-Doc synthetic proteins, Fok1a-BcDoc and Fok1b-CtDoc, respectively. We decided to move to next task to test and demonstrate functional Fok1 dimer proteome complex with Linker (20Lmod).

Procedures

Cloning and construction of expression strains

DNA sequences for Fok1a-AcDoc, Fok1a-BcDoc, Fok1b-BcDoc, and Fok1a-CtDoc were codon optimized and cloned into a pET22b(+) vector (GenScript, Piscataway, NJ). The sequence for a hexa-histidine tag was placed at the C-terminus of the constructs. *E. coli* BL21(DE3) LysY/Iq (NEB) was used for all protein expression experiments. The culture condition for *E. coli* to express the proteins are performed as described at Task 1.1 – 1.2.

Protein purification and running the SDS-PAGE gel.

The procedures are performed as described at Task 1.1 – 1.2.

FokI enzymatic assay

The procedures are performed as described at Task 1.1 – 1.2.

Results

To complete production of functioning synthetic proteome, the four kinds of synthetic proteins were newly designed including Fok1a-AcDoc, Fok1a-BcDoc, Fok1b-BcDoc, and Fok1a-CtDoc (Fig. 3). These proteins were successfully expressed, purified, and exhibited the predicted size (Fig. 4). The functionality of these proteins was confirmed by DNA cleavage assay for Fok1 activity. We were able to confirm that Fok1a-BcDoc and Fok1a-CtDoc were expressed as functionally active form. However, the Fok1a-AcDoc did not show any cleavage activity on our substrate. Our negative control, which contains FokI, did not exhibit any FokI cleavage activity as expected (Fig.6). We were able to get two functional Fok1 derivatives along with Dockerin.

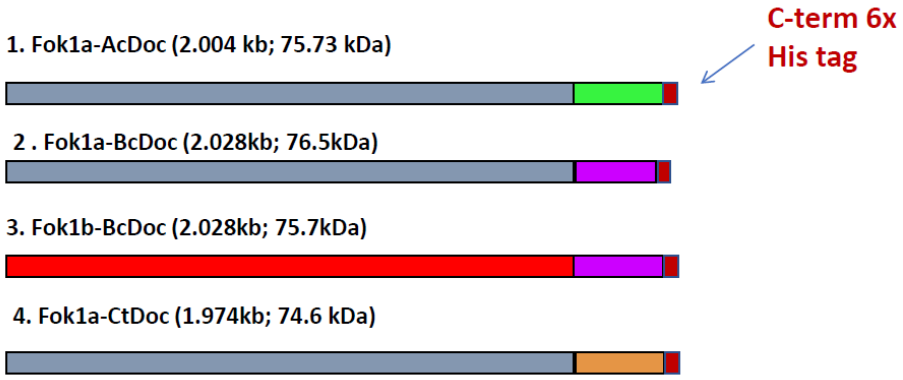


Fig. 4. A diagram of four kinds of newly designed Fok1 derivatives. 1. Fok1a-AcDoc; 2. Fok1a-BcDoc; 3. Fok1b-BcDoc, and 4. Fok1a-CtDoc. Fok1a (wild type; gray) and Fok1b (mutant; red).

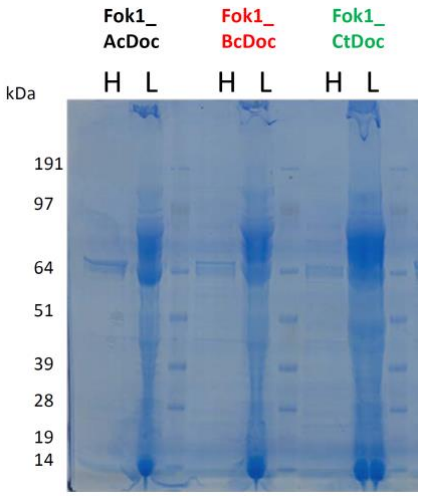


Fig. 5. SDS-PAGE gel (4 to 12% gradient) stained with Colloidal Blue showing. MW standards (BioRad) is indicated. The samples are indicated on the top of the gel image. Lane H, His-tag purified protein; Lane L, Cell Lysate.

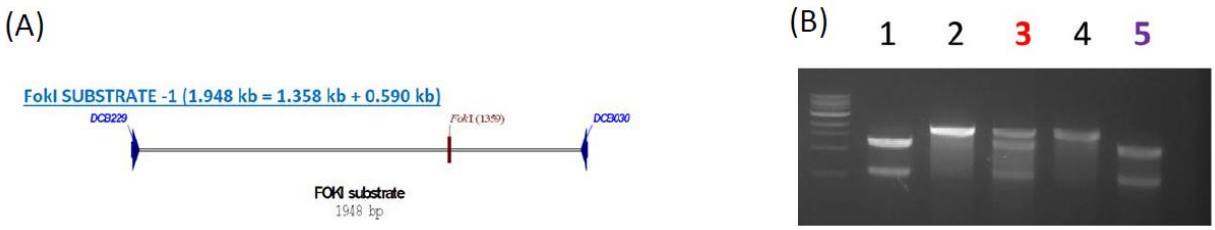


Fig. 6. Comparison of DNA cleavage by various FokI derivatives. (A) The DNA substrate used for this cleavage assay. (B) The agarose gel running after cleavage reaction. Lane 1, Fok1a; Lane 2, Fok1a-AcDoc; Lane 3, Fok1a-BcDoc; Lane 4, Fok1b-BcDoc; Lane 5, Fok1a-CtDoc.

Tasks 1.5-1.6: Testing and demonstration of functional Fok1 dimer/DNA synthetic proteome complex

Summary

In this task, we tried binding assays using three kinds of proteins, including Fok1a-BcDoc, Fok1a-CtDoc and Linker (scaffolding 20Lmod), to demonstrate the assembly of functional Fok1 dimer/DNA synthetic proteome complex. The cohesin-dockerin covalent interaction and its assembly to functional synthetic proteome complex was evaluated by native PAGE gel analysis. We were able to detect the binding interactions between Fok1 derivatives and Sca20Lmod. The cleavage activity assay indicated that the synthetic proteome complex still contains the Fok1 cleavage activity. We believe that the Phase I was successfully completed.

Procedures

Protein expression.

E. coli SG1146a (*clp::cat, ompT,*Δ*(lon*) strain was used for all protein expression experiments in this task to improve protein stability during cell growth and purification step. The culture condition for *E. coli* to express the proteins are performed as described at Task 1.1 – 1.2.

Binding assay

His-tag purified proteins (Fok1a-BcDoc and Fok1a-CtDoc) were mixed in equal molar amounts (1.31 pmol) with engineered scaffolds (20Lmod) to form the synthetic proteomes in an optimized buffer (20 mM Tris, 100 mM NaCl, 5 mM CaCl2 and 0.02% sodium azide at pH 7.4.) and allowed to incubate at cold room for 2 hours.

Protein purification and running the native gel.

The protein purification procedures are performed as described at Task 1.1 – 1.2. The native PAGE gel (3 to 12%) was used, and the gel running performed at cold room.

FokI enzymatic assay

The procedures are performed as described at Task 1.1 – 1.2.

Results

We employed a new *E. coli* expression strain, which is an all protease deficient strain, to improve the expression and purification of Fok1a-BcDoc and Fok1a-CtDoc proteins. We were able to purify large amount and highly pure proteins (Fig. 7). To test and demonstrate the assembly of functional Fok1 dimer/DNA synthetic proteome complexes, we performed the binding assay using three proteins as described in the procedures. The diagram of our binding assay is shown in Fig. 8. Although, there was no strong binding interaction between cohesin and dockerin, we could detect the visible assembled synthetic proteome complex (Fig. 9). The functionality of the assembled synthetic proteome complexes was confirmed by DNA cleavage assay for Fok1 activity. The activity assay was performed using the synthetic proteome (Sca20L-FokI_BcDoc-FokI_CtDoc) after 2 hour binding reaction. We were able to confirm the assembly for functionally active assembly of proteome complex (Fig. 10).

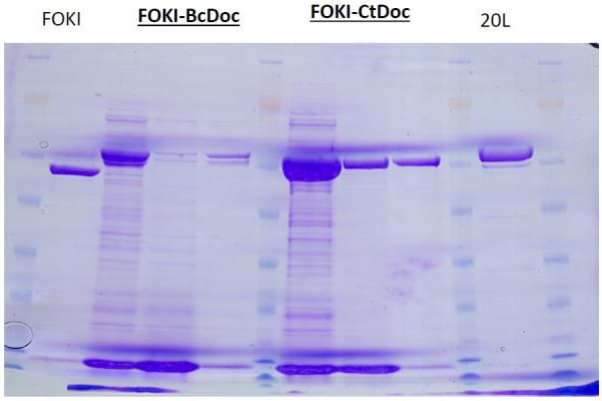


Fig. 7. SDS-PAGE gel (4 to 12% gradient) stained with Colloidal Blue showing. MW standards (BioRad) is shown between samples. The samples are indicated on the top of the gel image.

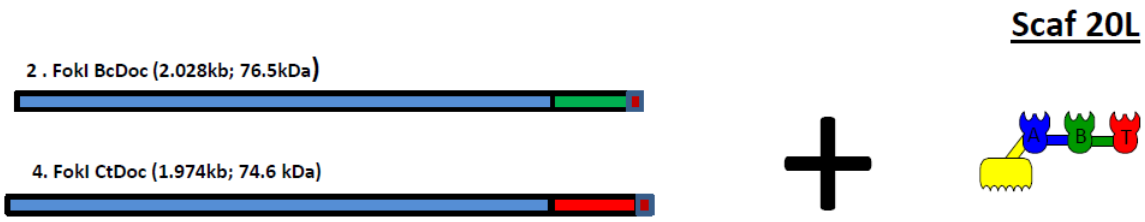
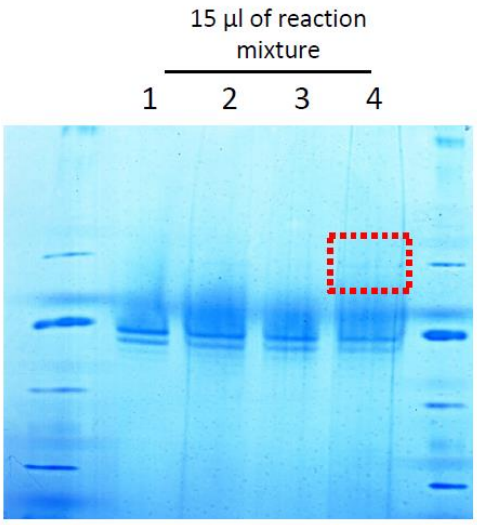


Fig. 8. A diagram of the binding assay employing three kinds of synthetic proteins. The cohesion and dockerin are color matched. BcDockerin and BcCohesin, green color; CtDockerin and CtCohesin, red color.



Lane 1 : Scaf 20L only
Lane 2 : Scaf 20L + FokI_BcDoc
Lane 3 : Scaf 20L + FokI_CtDoc
Lane 4 : Scaf 20L + FokI_BcDoc + FokI_CtDoc
L : Protein size Ladder

Fig. 9. Native PAGE gel (3 to 12% gradient) stained with Colloidal Blue showing. MW standards (BioRad) is shown at left panel. The samples are indicated. 15 µl of reaction mixture was analyzed. The predicted assembled protein complex assembled protein complex was highlighted in red box.

Lane 1 : Scaf 20L only
Lane 2 : Scaf 20L + FokI_BcDoc
Lane 3 : Scaf 20L + FokI_CtDoc
Lane 4 : Scaf 20L + FokI_BcDoc + FokI_CtDoc

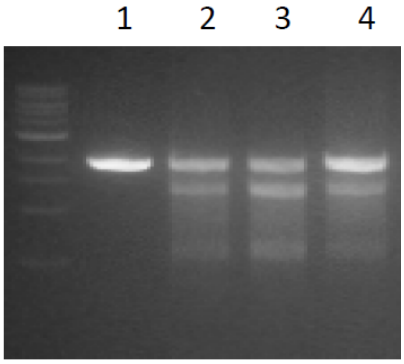


Fig. 10. Comparison of DNA cleavage by assembled synthetic proteome complex. The samples are indicated at left side.

Phase 2. Design assembly and testing of multi-protein prototype Synthetic Proteome

Tasks 2.1-2.2: Proteins isolated, stable and ready for assembly

Task 2.3: CAS-9 fusion proteins isolated, stable and ready for assembly

Summary

To develop the fusions for catalytically inactive Cas9 and Fok1 nuclease (fCas9) associated with scaffoldin, we made a huge effort to design proper constructs. It was the most critical part to perform the experiments for Phase 2. Our new architectures were carefully chosen based on the results of Phase 1 and the recently published paper (Guilinger *et.al.*, Nature Biotechnology 2014 DOI:10.1038/nbt.2909). The eight fusion architectures, which are fCas9 associated with various dockerins in various positions, were designed. To complete the phase 2 in efficient way, we

decided to pursue the Tasks 2.1, 2.2 and 2.3 at the same time. All of the proteins required for this phase are substantial in size; thus each gene synthesis take 4-8 weeks. We successfully constructed the *E. coli* expression strains. However, only three constructs (#5, 6, and 8) expressed as soluble forms among 8 constructs. Unfortunately, our experimental constructs with FokI (#1-4) are not successfully expressed at all. These kinds of large synthetic fusion proteins with several individual domains often cause the issue for proper protein folding, thus cannot be properly expressed. Based on these results, we decided not to pursue to complete the Phase 2.

Procedures

Cloning and construction of expression strains

DNA sequences for eight fusion constructs were codon optimized and cloned into a pET22b(+) vector (GenScript, Piscataway, NJ). The sequence for a hexa-histidine tag was placed at the C-terminus of the constructs. *E. coli* BL21(DE3) LysY/Iq (NEB) was used for all protein expression experiments. Recombinant strains were grown in LB broth supplemented with kanamycin (50 μg/mL). Cultures were induced at various temperatures with various amount of IPTG when OD600=0.6.

Protein purification and running the native gel.

The protein purification procedures are performed as described at Task 1.1 – 1.2.

FokI enzymatic assay

The procedures are performed as described at Task 1.1 – 1.2.

Results

The eight fusion architectures were carefully designed. Four of them (#4-8) designed as positive control without FokI based on the paper (Guilinger et.al., Nature Biotechnology 2014 DOI:10.1038/nbt.2909). The fusion architectures are described in Fig.1. The E. *coli* transformation were performed successfully using tightly regulated lysY/Iq cells (NEB) as a background strain. The #1-4 constructs were transformed with M.Fok1, and the positive control constructs transformed without M.Fok1 (Fig. 2). The only three constructs (#5, 6 and 8) were expressed in the Initial expression experiments (Fig. 3). We attempted to optimize the protein expression condition with various IPTG concentration from 0.05 to 1.0 mM. We also tried various induction temperatures from 15 to 30$_{o}$C. Unfortunately, constructs #1, 2, 3, 4, and 6 never expressed at any of the expression conditions.

Gene synthesis 1-4 (Constructs)

1. CtDoc-GGS (Linker1)-FokI-XTEN (Linker2)-dCas9-6x His (4983 bp, 1661 amino acid, 191 kDa)
2. BcDoc-GGS (Linker1)-FokI-XTEN (Linker2)-dCas9-6x His (5028 bp, 1676 amino acid, 200 kDa)
3. FokI-XTEN (Linker2)-dCas9-GGS (Linker1)-CtDoc-6x His (4983 bp, 1661 amino acid, 198 kDa)
4. FokI-XTEN (Linker2)-dCas9-GGS (Linker1)-BcDoc-6x His (5028 bp, 1676 amino acid, 200 kDa)

Gene synthesis 5-8 (Positive controls)

5. CtDoc--XTEN (Linker2)-Cas9-6x His (4377 bp, 1459 amino acid, 174 kDa)
6. BcDoc-XTEN (Linker2)-Cas9-6x His (4422 bp , 1474 amino acid, 176 kDa)
7. Cas9-XTEN (Linker2)-CtDoc-6x His (4377 bp, 1459 amino acid, 174 kDa)
8. Cas9-XTEN (Linker2)-BcDoc-6x His (4422 bp , 1474 amino acid, 176 kDa)

Fig.1 The eight fusion architectures designed for this study. The gene synthesis from 5-8 were included to be used as positive controls.

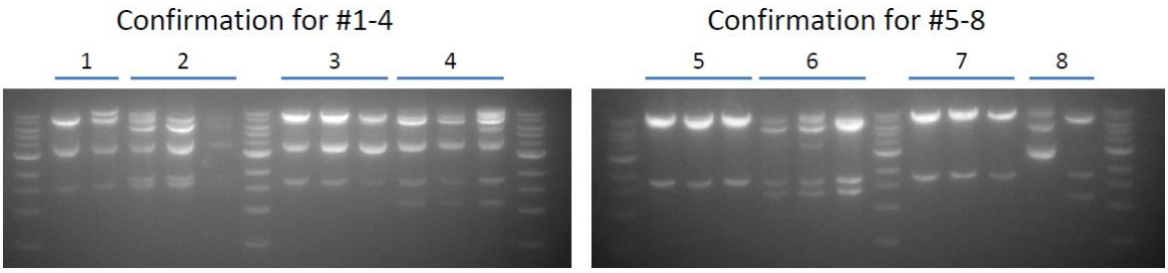


Fig. 2. Confirmation of *E. coli* expression strains. Expression plasmids for Gene synthesis (#1-4) were co-transformed with the M.FokI expression vector. However, gene synthesis (#5-8) expression vector were transformed without the M.FokI expression vector.

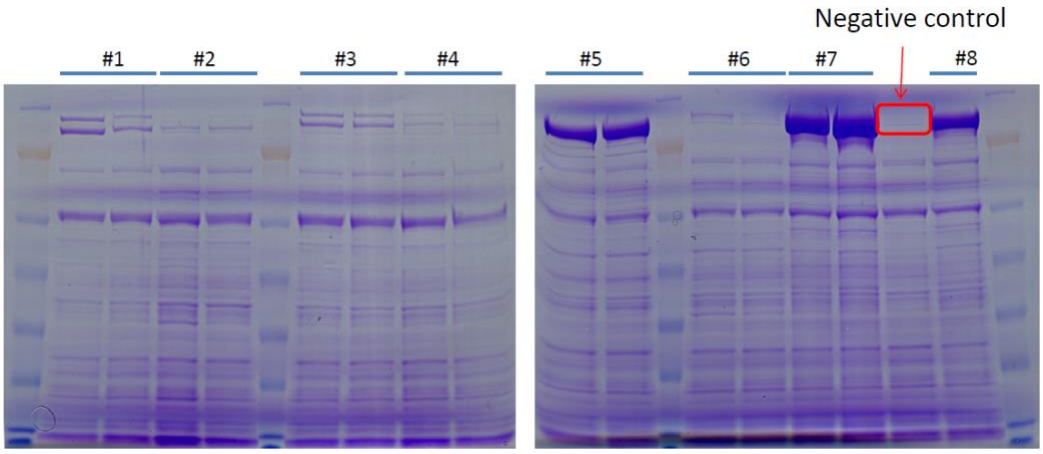


Fig. 3. SDS-PAGE gel (4 to 12% gradient) stained with Colloidal Blue showing. MW standards (BioRad) is indicated. The samples are indicated on the top of the gel image. The #5, 6 and 8 were expressed very well. However, other constructs are never expressed at any expression condition.

Unfortunately, our experimental constructs with FokI (#1-4) were not successfully expressed. These kinds of large synthetic fusion proteins with several individual domains often cause the issue for proper protein folding, thus cannot be expressed. Based on these results, we decided, with approval from Fed impact, not to pursue the completion of **Phase 2 (2.4-2.7)** but instead focus on obtaining these large fusion proteins. Unfortunately, after many trials, different constructs, and expression conditions, we could not obtain these constructs. Nevertheless, Fed impact was still able to see the promise of our approach.

Subject Inventions Listing:

None

ROI #:

None

Responsible Technical Contact at Alliance/NREL:

Yannick Bomble | Yannick.Bomble@nrel.gov

Name and Email Address of POC at Company:

Tammi Jantzen | tammi@fed-impact.com

DOE Program Office:

Energy Efficiency and Renewable Energy (EERE) Bioenergy Technologies Office (BETO**)**