FRAGLER: a fragment recycler application enabling rapid and scalable modular DNA assembly

David Öling[1], Jordi Chi[1], Tora Edström[1], Bo Peng[1], Fredrik Karlsson[1], Petr Volkov[1] and Robert Roth[1]

1. Discovery Sciences, Biopharmaceuticals R&D, AstraZeneca, 43183 Gothenburg, Sweden

Abstract

Rapid and flexible plasmid construct generation at scale is one of the most limiting first steps in drug discovery projects. These hurdles can partly be overcome by adopting modular DNA design principles, automated sequence fragmentation and plasmid assembly. To this end we have designed a robust, multi- module *golden gate* based cloning platform for construct generation with a wide range of applications. To minimize timelines and cost for complex constructs, we developed a software tool named FRAGLER (FRAGment recycLER) that performs codon optimization, multiple sequence alignment and automated generation of fragments for recycling. To highlight the flexibility and robustness of the platform, we fragmented multiple SarsCoV2 spike protein sequences as well as generated protein reagents in a rapid, scalable and cost efficient manner.

Introduction

- Construct generation for protein expression and cell line engineering is one of the main bottlenecks for progression of early drug discovery projects
- Over the last decade numerous modular Golden Gate based (Type IIS restriction enzyme) cloning kits have been published. While these kits are modular in nature, none of them address the long timelines for *de novo* synthesis of long coding sequences
- Here we describe an automated plasmid generation platform and software application termed FRAGLER **(FRAGment recycLER)**
- **Integration of FRAGLER in Benchling enables an algorithmic search and recycling of pre-existing DNA fragments.**

Methods

- All assemblies were done according to standard procedures using either *BsaI* or *BsmBI* master mix (NEB #E1601L/E1602L) with a 60x protocol: 37C 5min, 16℃ 5min and final digestion at 60℃ for 5 minutes.
- Assemblies were performed with equimolar DNA fragment concentrations (synthesized, cloned and codon optimized by GeneArt) in a concentration of 50-100 fmol/fragment.

Conclusions

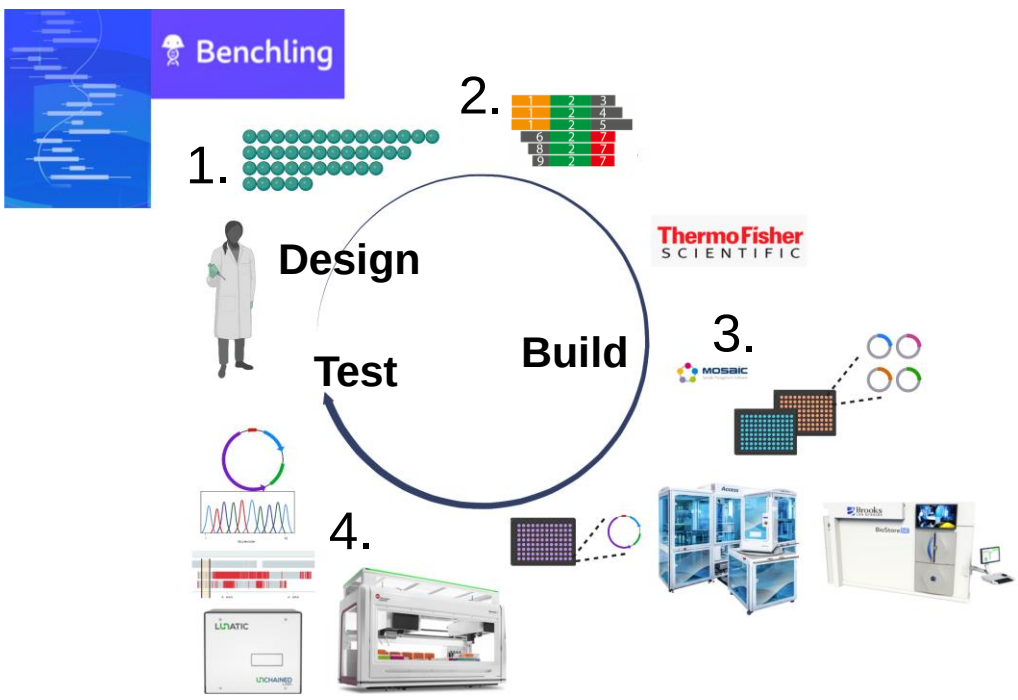
- We have built an automated plasmid generation platform powered by **Benchling** and an in-house built fragmentation algorithm termed **FRAGLER**
- FRAGLER was used for rapid generation of multiple long and complex SARSCoV2 spike protein constructs for expression optimization We further demonstrate the value of FRAGLER by fragmenting 30 Spike protein sequences resulting in 55,3% (63k) **nucleotide recycling** rate generating a significant cost-reduction in a sustainable manner
- Taken together, FRAGLER enables rapid, cost efficient and scalable plasmid construct generation with **3 week cycle times**

References

1) FRAGLER: a fragment recycler application enabling rapid and scalable modular DNA assembly

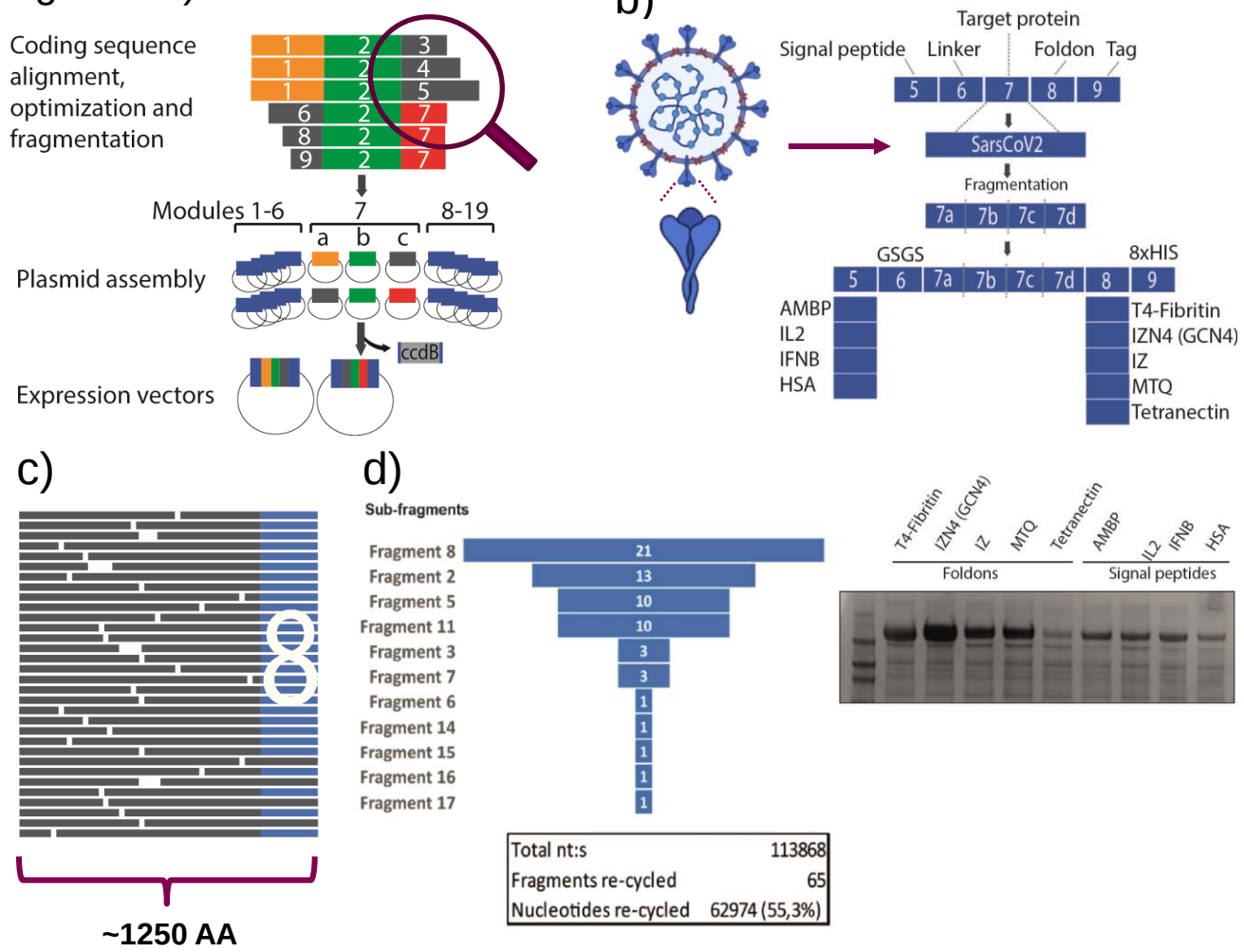
Results

Figure 1



Automated plasmid generation workflow. 1) Multiple Amino acid sequences are requested in a web-portal. 2) Sequences are bioinformatically processed **(FRAGLER)**[1] and synthesized. 3) In-house DNA fragments from the Biostore are combined with *de novo* synthesized fragments and assembled using an ECHO 655T integrated on an Access system. 4) The assembly mixture is cloned and single colony derived plasmids extracted on a Biomek i7 with integrated colony picker. Plasmids are validated by Sanger sequencing. The **cycle time is 3 weeks.** Data is tracked and documented in **Benchling**.

Figure 2 a)



2. FRAGLER enables rapid generation of SarsCoV2 expression constructs. a) FRAGLER performs AA sequence alignment to find overlapping regions, codon optimization, and fragmentation to increase DNA production success. FRAGLER runs in Benchling to perform an algorithmic search to identify pre-existing DNA fragments. b) Corona virus spike protein schematic representation. The spike protein coding sequence is fragmented into 4 submodules of 7 (a, b, c and d) to reduce production timelines. These are combined with varying signal peptides (module 5) and trimerization domains (module 8) (top). Coomassie-stained gel from a small-scale purification of HIS-tagged spike protein constructs and expression in Hek293 cells (bottom). C) Schematic representation of an alignment and fragmentation of SARS2 spike protein (PODTC2) top 30 hits from Uniprot (90% identity). d) Graphical representation of fragments that are recycled (top,). Fragment 8 is the most frequently re-cycled fragment. Total number of recycled nucleotides and fragments (bottom).

Supported by AstraZeneca

REMOVE IF NOT PERMITTED