

EFFICIENT BIOLOGICAL CONSTRUCTION OF REPETITIVE POLYPEPTIDES FOR INTERCONNECT APPLICATIONS BY BLOCK COPOLYMERIZATION

Seiichiro Higashiya¹, Silvana C. Ngo¹, Kenneth S. Bousman¹, Xiaolin Jin¹,
John T. Welch^{1*}, Richard P. Cunningham²,
Eric T. Eisenbraun³, Robert E. Geer³, and Alain E. Kaloyeros³

¹ Department of Chemistry, ² Department of Biology, ³ School of Nanosciences and Nanoengineering, University at Albany – SUNY
1400 Washington Ave, Albany, NY 12222
Fax: 518-442-3462, jwelch@uamail.albany.edu.

Introduction

Self-assembly is an especially attractive strategy for the assembly of microstructures for microelectronics in the era of giga to tera scale integration. For such nanoscale self-assembly, preparation of precisely controlled building blocks is crucial. We are currently focusing on nanoscale molecular interconnect construction using precisely arrayed aromatic moieties. We chose aromatic amino acid residues for charge transport and β -sheet repetitive polypeptides as scaffolds for the aromatic arrays. These arrays can be biologically prepared in a monodisperse manner, controlling the sequence, and configuration, using environmentally friendly, cost-effective, and scalable processes.

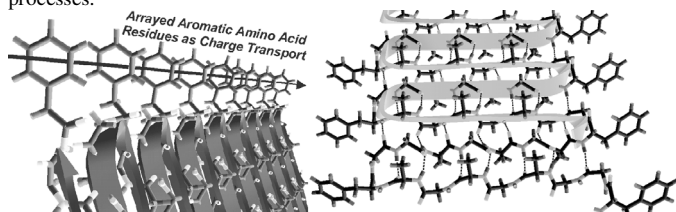


Figure 1. Modeled structures of arrayed aromatic amino acid (phenylalanine) residues for nanoscale molecular interconnects

Since the preparation of repetitive β -sheet polypeptides from artificial repetitive DNA prepared by head-to-tail polymerization has been reported,^{1,2} initially we focused on reproducing and modifying these results. However, due to the inavailability of the vectors, especially, the cloning vectors that have unique *BanI* site were derived from pSY937 originally prepared by Ferrari *et al.*,³ we could not reproduce the results by the same methods.

However, many methods for the construction of repetitive DNA sequences have been reported for the production of the corresponding polypeptides.^{4,5} In view of the recent advances in the construction, expression, and purification procedures, we herein describe an improved and versatile method for library construction of repetitive DNA coding sequences for the corresponding polypeptides for molecular interconnect applications. Utilizing multimerization/block copolymerization in the presence of adapters containing appropriate recognition sites of type II and IIs restriction endonucleases, for respectively, cloning and regeneration of assembled DNA units, repetitive coding sequences were successfully constructed in a reproducible and predictive manner without using special cloning vectors. This construction method also suppresses the intramolecular cyclization of multimers that is problematic when longer sequences are constructed.^{4a}

Experimental

pUC18 were obtained from Bayou Biolabs (Harahan, LA) or Amersham Biosciences (Piscataway, NJ). pET-28a, BLR(DE3)pLysS, and Benzonase nuclease were obtained from Novagen, Inc. (Madison, WI). All other restriction endonucleases and modifying enzymes were purchased from New England Biolabs (Beverly, MA). Histidine-tagged polypeptides were detected using SuperSignal West HisProbe kit.

Most of the common protocols were adapted from literature.⁶ Plasmids, DNA fragments separated by agarose gel electrophoresis, and preparative amounts of polyhistidine-tagged soluble repetitive polypeptides were purified using QIAprep Spin Miniprep Kit, QIAquick Gel Extraction Kit, and Ni-NTA Spin Column (Qiagen Inc., Valencia, CA), respectively. EtOH precipitation was performed with 80% ethanol in the presence of 0.2 M NaCl.

One letter abbreviation of amino acids: A, alanine; G, glycine; F, phenylalanine; Y, tyrosine; H, histidine; K, lysine; E, glutamic acid.

Design of cloning adapters for oligomerized units construction and block synthesis (A1) and expression (A2). Adapter A1 consists of two pairs of complementary oligonucleotides for oligomerized unit construction, **1** and **2**, **3** and **4** (**Figure 2**) were synthesized, respectively. Adapter A2 also consists of two pairs of oligonucleotides for expression, 5'-GA TCC TAT GTT TGC GGC CGC AAA TAT TCT CGC GAT CCG ATG G (**5**), 5'-GC ACC CAT CGG ATC GCG AGA ATA TTT GCG GCC GCA AAC ATA G (**6**), 5'-GT GCC TAA TAA CCC GGG G (**7**), and 5'-AA TTC CCC GGG TTA TTA G (**8**) were designed to produce the same coding sequence as reported.¹

Oligonucleotides **2**, **3**, **6**, and **7** were 5'-phosphorylated by T4 DNA kinase, respectively, and the enzyme was deactivated at 80 °C. The same amounts of **1**, **4**, **5**, and **8** were added to each solution and the mixtures were annealed, respectively, then parts of **1-4** and **5-8** solutions were mixed to prepare 1 μ M stock solution of A1 and A2, respectively.

Design of monomer units coding polypeptides. Oligonucleotides **9** and **10** which generate a duplex with asymmetric *BanI* sticky ends and coding for the monomeric repeat of an oligopeptide, GAGAGAGF (**8F**) were synthesized according to literature.¹ Oligonucleotides **11** and **12** coding GAGAGAGY GAGAGAGF (**16YF**) that give alternate tyrosine and phenylalanine residues were derived from **5** and **6** as 5'-GT GCC GGG GCT GGC GCG GGA TAT GGT GCT GGA GCA GGA GCC GGG TTT G and 5'-GC ACC AAA CCC GGC TCC TGC TCC AGC ACC ATA TCC CGC GCC AGC CCC G, respectively. In a similar manner, 5'-GT GCC GGT GCA GGA GCT GGT CAT G (**13**) and 5'-GC ACC ATG ACC AGC TCC TGC ACC G (**14**) for GAGAGAGH (**8H**), 5'-GT GCC GGA GCT GGT GCT GGC TAT G (**15**) and 5'-GC ACC ATA GCC AGC ACC AGC TCC G (**16**) for GAGAGAGY (**8Y**), 5'-GT GCC GGC GCT GGA GCA GGT AAA G (**17**) and 5'-GC ACC TTT ACC TGC TCC AGC GCC G (**18**) coding for GAGAGAGK (**8K**), and 5'-GT GCC GGC GCA GGT GCT GGT GAA G (**19**) and 5'-GC ACC TTC ACC AGC ACC TGC GCC G (**20**) for GAGAGAGE (**8E**) were synthesized, respectively. **8K** and **8E** coding sequences contain additional *Ngo*MIV cleavage sites (GCCGGC) to facilitate the subsequent oligomer selection. Each pair of complementary oligonucleotides was 5'-phosphorylated by T4 DNA kinase then annealed from 80 °C to an ambient temperature to prepare 20 μ M duplex solutions.

Construction of repetitive polypeptide coding sequences. Each of recipient vectors, pUC18 for the oligomer unit construction or pET-28a for expression was digested by *Bam*HI and *Eco*RI in NEBuffer *Eco*RI and BSA then the small fragments were removed by gel electrophoresis. The desired recipient plasmid band was excised from the gel, purified using QIAquick Gel Extraction Kit, (precipitated with EtOH, and dissolved in TE) to prepare each of 100 ng/ μ L stock solution.

The monomer-coding DNA solution (1~2 μ L) and the adapter solution for oligomer unit construction A1 or expression A2 at a ratio of 10~100:1 (0.2~2 μ L) depending on the desired degree of oligomerization, were mixed with PEG8000 (final concentration at 4%), and H₂O (up to 9 μ L) and annealed at 45 °C for 5 min. The T4 DNA ligase (1 μ L) and the supplied ligation buffer (1 μ L) were added and allowed to promote oligomerization at 4 °C for overnight or longer. The resultant oligomer population bearing adapters at the both ends was separated on agarose gel. Oligomers of the desired length were excised from the gel, purified, and precipitated with EtOH in a 0.6 mL microcentrifuge tube.

To the tube were added an appropriate recipient vector solution (2 μ L), PEG8000 (final concentration at 4%), T4 polynucleotide kinase (0.25 μ L), kination buffer (0.5 μ L), and H₂O (up to 5 μ L) then the mixture was incubated at 37 °C for 30 min. T4 DNA ligase (0.25 μ L) was added at room temperature and the mixture was incubated at 4 °C overnight or longer. The ligation mixture (1 to 2 μ L) was used to transform XL-1 Blue (20 to 40 μ L, Stratagene Inc., La Jolla, CA) by electroporation (13 kV/cm) using Ecoli Pulser (Bio-Rad Laboratories, Hercules, CA). The transformants were selected on LB agar plate with an appropriate selection marker (100 μ g/mL ampicillin, IPTG, and X-gal for recombinant pUC18 or 50 μ g/mL kanamycin for recombinant pET-28a) and incubated at 37 °C. Upon colonie selection and plasmid digestion by the restriction enzymes (*Bam*HI for A1 or *Bam*HI and *Eco*RI for A2) electrophoresis of the cleavage products facilitated selection of plasmids with the appropriate inserts for sequencing.

Block copolymerization of oligomerized multimer units. *E. coli* colonies found to possess recombinant pUC18 harboring the desired oligomeric units with A1 adapters were inoculated into 2xYT (4 to 6 mL) containing 100 μ g/mL ampicillin and cultured at 37 °C for 15~20 h. The

plasmids were harvested then digested by *BsaI*. The resulting oligomeric units were separated by agarose gel electrophoresis, purified, and precipitated with EtOH. In the case of block copolymerization, an equal amount of each plasmid containing the desired oligomeric unit was digested by *BsaI* and the resulting fragments were combined. The purified fragments of oligomeric units were used for further oligomerization in the presence of the appropriate adapters as previously described.

Expression and purification of repetitive polypeptides. *E. coli* colonies found to possess recombinant pET-28a containing the desired repetitive coding sequences with A2 adapters were inoculated in 2xYT (6 mL) containing kanamycin (50 µg/mL) and cultured at 37 °C for 15~20 h. The plasmids were harvested then used to chemically transform an expression host, BLR(DE3)pLysS. The transformants were selected on LB agar plates containing kanamycin (50 µg/mL), chloramphenicol (34 µg/mL), and 1% glucose. Selected colonies were inoculated into 2xYT (2 to 5 mL) for preliminary 5 to 50 mL scale expression or 200 mL for 4 L scale expression) containing kanamycin (50 µg/mL), chloramphenicol (34 µg/mL), and 1% glucose. The overnight precultures were used to inoculate desired amounts of expression culture (2xYT containing kanamycin (50 µg/mL) and chloramphenicol (34 µg/mL). Expression of polypeptides was initiated between 1.0 to 2.0 OD by the addition of IPTG (1 mM) and was induced for 3 to 4 h with aeration. The expression products were analyzed by SDS-PAGE and the polyhistidine-tagged polypeptides were detected using SuperSignal West HisProbe kit (Pierce, Rockford, IL).

8F and 16YF polypeptides were purified according to the literature protocol² with the following modifications: Benzonase was used instead of Dnase I and RNase A and three additional washes with 1% Triton-X100 were employed. **32YEHK** polypeptides were purified using Ni-NTA Spin Column.

Results and Discussion

Construction of repetitive polypeptide coding sequences.

Recognition sites for the restriction endonucleases in the adapter oligonucleotides A1 for oligomerized unit construction are shown in **Figure 2**.

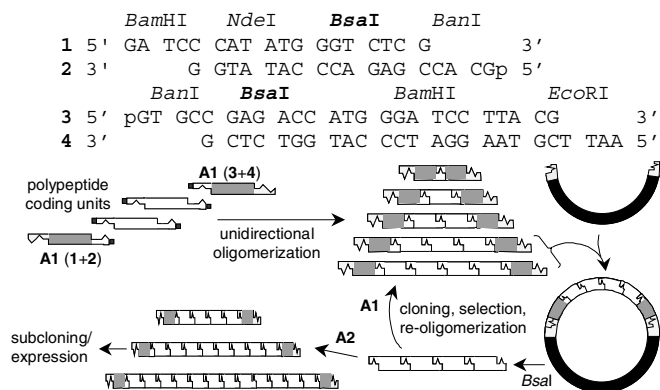


Figure 2. Recognition sites of restriction endonucleases in adapter A1 and construction of coding sequences for repetitive polypeptides

The duplexes produce asymmetric 4 bp *BanI* sticky ends to adapt monomeric and oligomerized units in a unidirectional manner. The 5'-OH of the *BanI* sides (2 and 3) were phosphorylated, respectively. The remaining ends form, respectively, *BamHI* and *EcoRI* sticky ends for cloning into the multicloning site of pUC18. The 5'-OH moieties of 1 and 4 were left unphosphorylated to prevent the adapter-adaptor ligation. The *BanI* sides of expression adapter A2 were 5'-phosphorylated in the same way. The adapter was mixed with monomeric or oligomerized units and ligated in the presence of PEG8000.⁶ Once the ligation between the units and adapters occurred, the subsequent intramolecular self-cyclization that predominates when longer fragments are subjected for oligomerization,^{4a} is completely suppressed facilitating the exhaustive ligation and simultaneous construction of oligomer libraries of longer repeats in an efficient manner. Oligomerized units with adapters on both ends were separated by agarose gel electrophoresis. The desired fragments were 5'-phosphorylated and cloned into recipient plasmids.

The recipient plasmids were linearized by double digestion using *BamHI* and *EcoRI* and the shorter fragments were removed to promote the efficient formation of the desired recombinant plasmids.

After selection and sequencing, plasmids harboring the oligomeric units with A1 adapter were excised by the type II restriction endonuclease *BsaI*. The recognition sites were earlier introduced in A1 to regenerate the oligomeric units with asymmetric *BanI* sticky ends for further multimerization/block copolymerization.

The 13 repeats of **8F (8Fx13)** and 4 repeats of **16YF (16YFx4)** were successfully constructed from the monomer units **8F** and **16YF** with A1, respectively. The oligomerized units were then subcloned into an IPTG inducible expression vector pET-28a with A2 adapter. The respective plasmids harboring **8Fx13**, **x27**, **x39** and **16YFx4**, **x8**, **x12** were obtained. The plasmids were used to transform a recombinase deficient expression host, BLR(DE3)pLysS and **8Fx13**¹² and **16YFx12** were expressed in mass quantities, purified, and cleaved by CNBr to obtain the desired repetitive polypeptides.² The products gave satisfactory amino acid and sequencing analyses results even though they aggregated strongly as seen in SDS-PAGE gel¹ and were difficult to handle due to their hydrophobic nature.

Next, we changed the pendant aromatic amino acids (**Figure 1**, left model) to repeats of tyrosine (Y), glutamic acid (E), histidine (H) and lysine (K), *i.e.* **8Y8E8H8K (32YEHK)** to facilitate the β -sheet formation and solubility. The electrostatic interaction of the charged and hydrophilic E and K side chains appearing in every 2 β -turns can occupy closely contacting positions. In addition aromatic interactions of alternate electron-rich Y and -poor protonated H residues on the opposite sides of the β -turns are also enhanced.

For the construction of repetitive coding sequences such as **32YEHK**, a mixture of **8Y** and **8E**, and **8H** and **8K** coding oligonucleotides were ligated, respectively, and the subsequent recombinant plasmids harboring dimerized units, **16YH** and **16HK** were selected with the help of *NgomI* digestion. The cleavage sites had been previously introduced to **8E** and **8K** coding sequences. Each of the desired dimerized units was obtained by the digestion with *BsaI* and were then dimerized to form **32YEHK**. A selected plasmid harboring desired **32YEHK** was used for the further oligomerization and a plasmid harboring 7 repeats of **32YEHK (32YEHKx7)** was selected. The **32YEHKx7** unit was subcloned into pET-28a and plasmids harboring **32YEHKx14**, **x21**, and **x28** were identified in 48 colonies that were used for transformation of BLR(DE3)pLysS. Following expression and purification the polypeptides showed a vastly diminished propensity to aggregation and were identified by SDS-PAGE.

Conclusions

Several repetitive coding sequences were successfully constructed by a new oligomerization technique. The oligomerized units were formed in the presence of adapters containing appropriately introduced type II restriction endonuclease recognition sites and the corresponding repetitive polypeptides were successfully expressed.

Acknowledgement. We greatly acknowledge Prof. Charles P. Scholes and Dr. Evelina Loghin for their kind help. This work was supported by MARCO Interconnect Focus Center and the Focus Center-NY.

References

- (1) Yoshikawa, E.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *Macromolecules* **1994**, *27*, 5471-5475. Preparation of pMD-3a: Dougherty, M. J. 1993, Ph.D. thesis, University of Massachusetts, Amherst.
- (2) Cantor, E. J.; Atkins, E. D. T.; Cooper, S. J.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *J. Biochem.* **1997**, *122*, 217-225. Preparation of pECL.1: Cantor, E. J. 1997, Ph.D. thesis, University of Massachusetts, Amherst.
- (3) Ferrari, F. A.; Richardson, C.; Chambers, J.; Causey, S. C.; Pollock, T. J.; Capello, J.; Crissman, J. W. US patent 5,243,038 (1993).
- (4) a) Won, J.-I.; Barron, A. E. *Macromolecules* **2002**, *35*, 8281-8287. b) Goeden-Wood, N. L.; Conticello, V. P.; Muller, S. J.; Keasling, J. D. *Biomacromolecules* **2002**, *3*, 874-879. c) Meyer, D. E.; Chiloti, A. *Biomacromolecules* **2002**, *3*, 357-367 and references cited in the papers.
- (5) a) Hwang, J.; Hsu, C.-T.; Ting, C. J.; JP patent 2001321183 (2001). *Chem. Abstr.* 135:370626. b) Hwang, J.; Hsu, C.-T.; Ting, C. J.; EP patent 1,090,944 (2001). *Chem. Abstr.* 134:294520.
- (6) Sambrook, J.; Russel, D. W. *Molecular Cloning 3rd Ed.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2001.