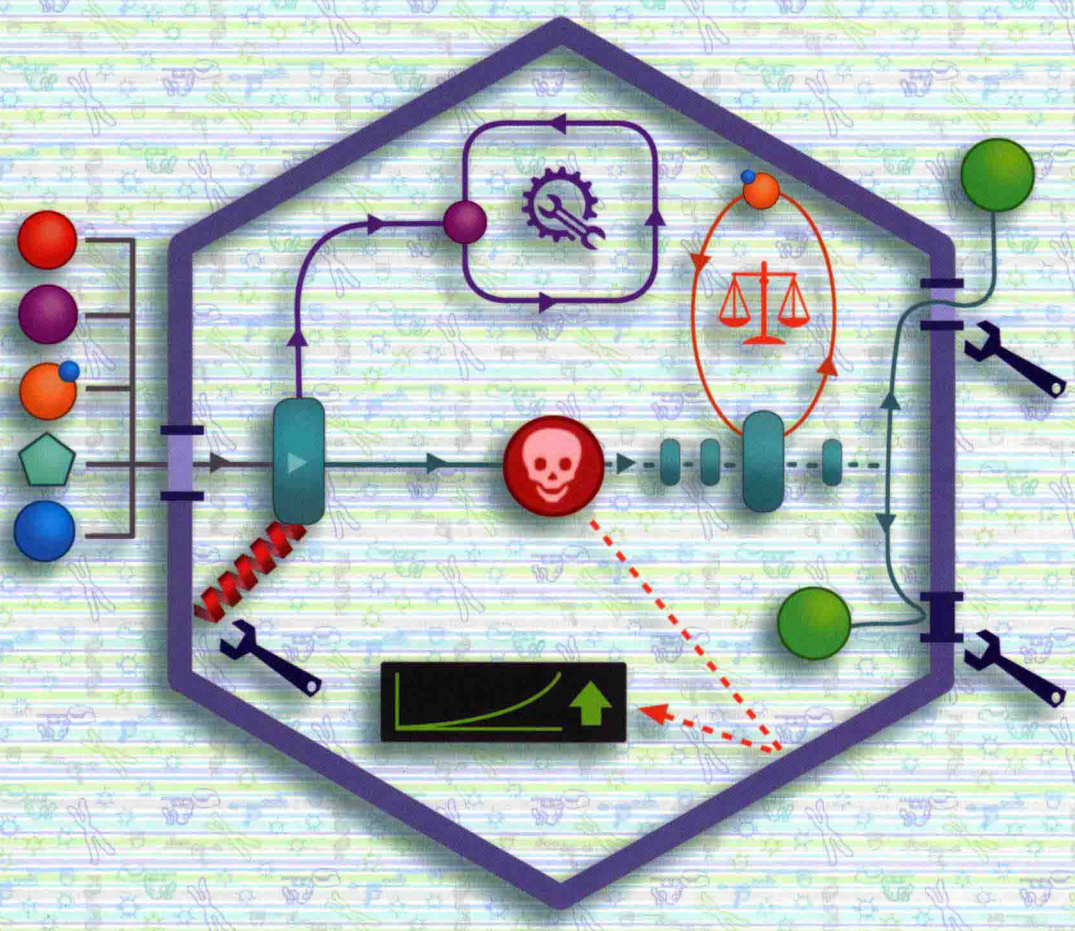
Advances in Molecular Biology and Medicine

Robert A. Meyers

Synthetic Biology

Volume 1



WILEY-VCH

Synthetic Biology

Advances in Molecular Biology and Medicine

Edited by
Robert A. Meyers

Volume 1

WILEY-VCH
Verlag GmbH & Co. KGaA

The Editor

Dr. Robert A. Meyers
Editor-in-Chief
Ramtech Limited
122, Escalle Lane
Larkspur, CA 94939
USA

Cover

Design and customization of microcompartments. © Martin J. Warren. Detail from figure 12 in chapter 13 "Synthetic Biology in Metabolic Engineering: From Complex Biochemical Pathways to Compartmentalized Metabolic Processes – a Vitamin Connection".

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

Print ISBN: 978-3-527-33482-7
ePDF ISBN: 978-3-527-66855-7
ePub ISBN: 978-3-527-66856-4
Mobi ISBN: 978-3-527-66857-1

Cover Design Adam Design, Weinheim, Germany
Typesetting Laserwords Private Limited, Chennai, India
Printing and Binding Markono Print Media Pte Ltd, Singapore

Printed on acid-free paper

Synthetic Biology

Edited by
Robert A. Meyers

Related Titles

Meyers, R.A. (ed.)

Encyclopedia of Molecular Cell Biology and Molecular Medicine

2007
Online ISBN: 978-3-527-60090-8
www.meyers-emcbmm.com

Schmidt, M. (ed.)

Synthetic Biology
Industrial and Environmental Applications

2012
ISBN: 978-3-527-33183-3, also available in digital formats

Luisi, P.P. (ed.)

Chemical Synthetic Biology

2011
ISBN: 978-0-470-71397-6, also available in digital formats

Kahl, G.

The Dictionary of Genomics, Transcriptomics and Proteomics
5th Edition

2015
ISBN: 978-3-527-32852-9, also available in digital formats

Preface

Synthetic Biology is an emerging field that integrates all biological subsystems with the objective of engineering, constructing or modifying biological functions, organelles, cellular structures, simple cells and creating whole novel organisms with designed properties. This is accomplished by application of engineering principles such as hierarchical design, modular parts, isolation of unrelated functions and standard interfaces as well as systems biology and all aspects of molecular biology. Synthetic Biology extends genetic engineering to focus on whole systems rather than individual genes and primary gene products, and will ultimately provide diagnostic tools, novel methods for production of therapeutics and strategies for treatment of diseases. Our compendium is written for university undergraduates, graduate students, faculty and investigators at research institutes. Our Board of 11 Nobel Laureates approved our overall approach and content and our selection of articles was validated and enhanced by four reviewers from major research institutions. There are 23 peer reviewed articles at a length of over 700 pages and as such is the largest in depth, up to date treatment presently available.

The 23 detailed articles organized into six sections: **Biological Basis; Modeling; Modular Parts and Circuits; Synthetic Genomes; Diseases and Therapeutics; and Chemicals Production**. In addition, there is an introductory article entitled Synthetic Biology. The **Biological Basis** section defines key areas that support synthetic biology approaches including Emergence of the First Cells (Protocells), Regulation of Gene Expression, the Interactome and also Microbiomes; the **Modeling** section provides mathematical and computer programming expertise including Dynamics of Biomolecular Networks, Computer Simulation of the Cell and the SynBioSS Designer Modeling Suite, which support the following section on **Modular Parts and Circuits**. This section covers key practical applications and advances such as Synthetic Gene Networks, DNA Origami Nanobots, RNAi Synthetic Logic Circuits for Sensing, Information Processing and Actuation, Synthetic Hybrid Biosensors and Synthetic Biology in Metabolic Engineering. The **Synthetic Genomes** section includes articles on Minimal Gene-Set Machinery; Production of the Mitochondrial Genome and Chromosomal DNA Segments, and Synthetic Genetic Polymers Functioning to Store and Propagate Information. As part of the introductory article, Sanjay Vashee, of the J. Craig Venter Institute, covers the recent work on the formation of a bacterial cell that is controlled by a chemically synthesized genome. This work was accomplished at the Institute where the

author is located. The next two sections cover important end use applications of synthetic biology. **Diseases and Therapeutics** covers synthetic biological improvements in use of stem cells in regenerative medicine and new approaches for vaccine development utilizing synthetic biology and finally the final section on **Chemicals Production** includes synthetic biology approaches for production of diols, biofuels and antibiotics.

We are pleased that many of the major synthetic biology research institutions participated in preparation of our book, e.g. the Departments of Biological Engineering, the Synthetic Biology Center of the Department of Electrical Engineering & Computer Science, the MIT Microbiology Program, and the MIT Computational and Systems Biology Program all of the Massachusetts Institute of Technology; the ETH Zurich, the National Institutes of Health, and the J. Craig Venter Institute.

Our team hopes that you, the reader, will benefit from our hard work – finding the content useful in your research and education. We wish to thank our Managing Editor, Sarah Mellor as well as our Executive Editor, Gregor Cicchetti for both their advice and hard work in the course of this project.

Larkspur, California, January 2015

Robert A. Meyers
RAMTECH Limited

Contents

1
Synthetic Biology: Implications and Uses*

Sanjay Vashee, Mikkel A. Algire, Michael G. Montague, and Michele S. Garfinkel
J. Craig Venter Institute, Synthetic Biology, 9704 Medical Center Drive, Rockville, MD 20850, USA

* This chapter has previously been published in: Meyers, R.A. (Ed.) Systems Biology, 2012, 978-3-527-32607-5.

Synthetic Biology: Advances in Molecular Biology and Medicine
First Edition. Edited by Robert A. Meyers.

Keywords

Synthetic biology
An effort to construct biological systems, which may include entire biosynthetic pathways, synthetic organelles and cellular structures, and whole organisms, that have medical, industrial, and scientific applications. This is achieved via the application of engineering principles, such as hierarchical design, modular reusable parts, the isolation of unrelated functions, and standard interfaces.

Synthetic cell
A cell that is controlled solely by a genome that was assembled from chemically synthesized pieces of DNA.

DNA assembly
The building of larger DNA fragments from smaller DNA fragments.

Circuits
A collection of various modular component parts that responds to an input signal that is then relayed to produce an output signal.

Compartmentalization
The spatial sequestering of substrates, intermediates, products, enzymes, and activities.

Synthetic biology is an effort to construct and engineer biological systems, ranging from individual genetic elements, to biosynthetic pathways, to whole organisms. The results of these engineering efforts can be of great value to human interests such as medicine and industry. In this chapter, advances in DNA assembly technologies are reviewed, and how these advanced DNA assembly technologies, in conjunction with the application of engineering principles such as modular parts, have facilitated the rational engineering of organisms to obtain desired functions or to understand complex cellular behavior, are highlighted. The recent creation of a synthetic cell is also described. Finally, the societal concerns posed by synthetic biology are discussed.

1 Introduction

The field of Synthetic Biology can be considered more as an engineering discipline, and less as an empirical science. Efforts to create artificial life systems, both in biochemical systems [1] and in software environments [2], may also be considered as Synthetic Biology, though these are beyond the scope of this chapter. Synthetic Biology is viewed as the effort to construct and engineer biological systems of value to human interests. Such

efforts can range in scopes far larger than the traditional genetic engineering of genes, to include the engineering of entire biosynthetic pathways complete with the regulation of the genes in that pathway [3, 4], synthetic organelles and cellular structures [5], whole organisms [6–9], and even ecosystems [10–13]. Synthetic Biology has the ambition to apply classical engineering principles such as hierarchical design, modular reusable parts, the isolation of unrelated functions, and standard interfaces. The empirical fields that correlate to Synthetic Biology are Systems Biology, Genetics, and Molecular Biology.

Synthetic Biology is not a new field, but rather extends back into prehistory. For example, it has been determined that the process of engineering maize – a highly optimized domestic agricultural crop plant – from the wild grass teosinte began over 9000 years ago [14]. The method used by the pre-Columbian cultivators of teosinte was simple artificial selection which, as such, is very slow. However, with the discovery of laws of inheritance and natural selection [15–17], and the suggestion that DNA was the chemical medium of inheritance [18], the scene was set to engineer a living system in a far more direct and rapid manner. A prominent example of this is the Dupont *Escherichia coli* strain used for the production of 1,3 propanediol, in which case an entire biosynthetic pathway has been added to *E. coli*, and the metabolism of the bacterium substantially altered to allow for a majority of the carbon feedstock (glycerol) to be converted into the economically valuable chemical 1,3 propanediol [6, 8, 9]. This feat, which was begun prior to the development of most of the Synthetic Biology techniques reviewed in this chapter, took many years and substantial investment to achieve. Yet, with recent advances in the field, such bioengineering projects will become faster to develop, easier to operate, and also much more ambitious.

During recent years, Synthetic Biology has progressed in a manner which is very different from those of other engineering disciplines. This is because, unlike architecture or software engineering, there is already a reservoir of highly sophisticated and complex functional parts to be found in Nature, and consequently most efforts in Synthetic Biology have been focused on harnessing that natural resource base. In general, two basic approaches have been undertaken to achieve this feat. The first approach has been to engineer natural organisms so as to incorporate recombinant pathways and other such desirable attributes. This method has the advantage of not requiring the capability to build – nor require an understanding of – massive biological systems such as genomes and metabolisms. However, it does have the disadvantage of being undefined; that is, whilst certain genes of the organism might be the result of human intervention, most of the genome remains wild-type, and is neither subject to human control nor necessarily operating within the limits of human knowledge.

The alternate approach is to use functional components, originally "mined" from Nature, such as promoters, aptamers, protein–protein interaction domains, terminators, or ribosome-binding sites. These functional components can be cataloged and then used to compose larger defined constructions of genes, pathways, and even whole genomes. Because all the components of a synthetic biological system that are constructed in such an approach have precisely defined properties, a high degree of predictive control over the final product is afforded. Indeed,

this more defined approach has become synonymous with advances in Synthetic Biology.

Historically, one of the main limitations in following this defined approach relates to the knowledge of these natural systems that serve as a source of parts. As biology has been characterized, both new components for synthetic biology – and also new tools to utilize those components – have become available. In turn, a new Synthetic Biology capability has driven greater advances in the understanding of biology. This has been most obvious in the development of tools for the synthesis and manipulation of DNA, the first of which were developed via the discovery of restriction endonucleases, DNA ligases, and the creation of recombinant DNA molecules [19–24]. These tools allowed the development of the recombinant DNA cloning and expression techniques that ultimately made possible the exploitation of enzymes with desired activities and properties. Notably, the development of the polymerase chain reaction (PCR) greatly increased the ability to amplify and manipulate DNA [25–27]. The subsequent combination of recombinant DNA cloning techniques with the PCR allowed a much greater exploitation of natural biological components, such as thermostable DNA polymerases, and this in turn made the PCR more robust and practical. Today, the PCR has become an indispensable tool for biology. Thus, knowledge of natural systems has led to an improved technology for exploiting those systems, which has in turn provided an improved knowledge of biological systems in a double feedback loop, thus improving both scientific understanding and technological capabilities. As shown in Fig. 1, as knowledge of the fundamental principles of biology have continued to grow, it has been possible to take a more defined engineering approach.

In the past, advances in synthetic biology have been bounded by the capacity to assemble and modify DNA, as well as knowledge of biological parts and circuits that such DNA might encode. Correspondingly, these two areas are directly addressed in the following sections, with details of synthetic biological pathways, synthetic genomes, synthetic organelles, and even synthetic organisms provided as

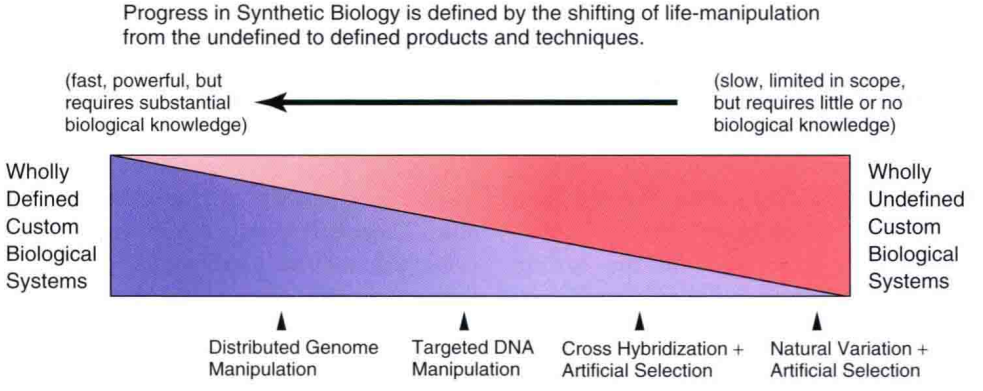


Fig. 1 Increasingly defined biological engineering. A schematic of how biological engineering has emphasized a more defined and rational design as it has advanced, based on a greater knowledge of natural biological systems.

examples of the capabilities that Synthetic Biology has already begun to deliver. These new capabilities create, in turn, new societal challenges, and these are also discussed.

2 DNA Assembly and Modification

As discussed above, one way to view Synthetic Biology is as an engineering discipline aimed at manipulating cellular systems to produce a *de novo*-designed function that does not exist in the natural organism. As with other engineering fields, however, Synthetic Biology is dependent on the tools and techniques available.

Organisms carry out a variety of reactions aimed at their self-sustenance and self-replication, with such reactions being carried out by the proteins and RNAs encoded in the organism's genome. As the sequence of the DNA directs production of the proteins and RNAs of an organism, control of the cellular DNA therefore allows for an ability to direct the functions of a cell. Based on this principle, many of the basic tools utilized in Synthetic Biology are aimed at producing defined sequences of DNA molecules, and easily manipulating the DNA content of an organism. The DNA molecules necessary for Synthetic Biology purposes can vary greatly in length, from individual DNA parts, genes and plasmids (containing tens to thousands of base pairs) to biosynthetic pathways and genetic circuits (thousands to millions of base pairs) to synthesizing whole genomes (viral and bacterial).

For many years, gene cloning and DNA assembly were dominated by the use of restriction endonucleases and DNA ligases [19–24]. While some well-designed restriction enzyme-based methods are still commonly in use [28, 29], these methods are gradually being superseded by the development of very rapid, more robust and less limited DNA assembly techniques. For example, although BioBricks were originally designed to be assembled with a restriction enzyme/ligation method [30], more recently an *in vitro* homologous recombination was adapted to increase the flexibility and speed of BioBrick assembly [31].

The starting materials for DNA sequence construction may include chemically synthesized DNA oligonucleotides (oligos), natural DNA fragments, PCR products, or a combination of all three sources. Defined short single-stranded oligos have been commercially available as a commodity for many years, usually for use as primers in PCR, mutagenesis, and sequencing reactions. Since chemically synthesized DNA oligos are of a user-defined sequence, this allows for a nucleotide level control of gene sequences and even entire genome sequences – a firm requirement when designing new functions in organisms. The idea of synthesizing genes from DNA oligos is not new; previously, oligos have been used to synthesize genes such as the alanine tRNA from yeast [32] and the human leukocyte interferon gene [33]. Likewise, the gene encoding a mammalian hormone, somatostatin, was synthesized and expressed in *E. coli* [34]. It is only recently that the lower costs of oligonucleotide synthesis and DNA sequencing have been combined to allow the development of more cost-effective and rapid methods for assembling groups of oligos into synthetic pieces of DNA or genes [35–39]. Today, several commercial gene synthesis companies exist that are able to produce custom genes/DNA at an accessible cost per base pair, although such costs

can quickly become prohibitive if numerous different DNAs are required. The cost prohibition of large-scale gene synthesis can, in part, be overcome by utilizing natural DNA fragments and PCR products in the DNA assembly for those sections of DNA that do not need to be created synthetically.

Typically, the process of constructing DNA is hierarchical (Figs 2 and 3). Briefly, groups of smaller DNAs (single- or double-stranded) are mixed and assembled into larger DNA pieces. Figure 2 shows double-stranded DNAs being assembled into a larger construct, but the process can begin with single-stranded oligos as the substrates. These larger pieces (subassemblies) are then grouped and assembled. These steps can be repeated until the final full-length DNA construct is obtained, whether it is a gene or genome. The DNA pieces to be assembled must have homologous overlapping ends, the overlaps being important because the DNA assembly techniques utilize homologous recombination. For example, if three pieces of DNA (A, B, and C) are to be assembled into a single DNA molecule, then one end of piece A must have an overlap with piece B, and the other end of B must overlap piece C. This configuration will result in a linear DNA molecule, A–B–C. In order to generate a circle from these pieces, the end of C must overlap piece A. The assembly of DNA into a circle is most often achieved with a DNA piece which contains sequences that enable the final construct to be cloned into a desired host (e.g., *E. coli*, *Saccharomyces cerevisiae* or *Bacillus subtilis*) [37, 38, 40–42]. The DNA homologous recombination reaction can be carried out completely *in vitro* in one step, either with an enzyme mix [37] or by using the PCR [35]. The reaction can also be performed *in vivo* by utilizing the natural homologous recombination activity of an organism, such as *S. cerevisiae* (yeast) and *B. subtilis* [7, 40–44]. Occasionally, a method will include an *in vitro* step to perform a partial reaction (DNA chewback/DNA annealing/DNA extension), and an *in vivo* step to complete the reaction (DNA repair) [38, 45, 46].

The various *in vitro* homologous recombination methods used to assemble double-stranded DNA or single-stranded oligos share the same general mechanism (Fig. 2). The nucleotides are first removed from one strand of the overlapping ends of the adjacent double-stranded DNAs, thus creating single-stranded ends of the DNA (Step 1). This process is analogous to a restriction enzyme digestion creating complementary sticky ends of DNA, except that the single stranded ends are typically 20–60 nucleotides long. Depending on the method used, the nucleotides can be removed by applying an exonuclease activity from either the 5′ or 3′ ends of the DNA [37, 38]. The creation of single-stranded overhangs (Step 1) is not necessary if oligos are used as the starting material for DNA assembly, because the oligos are already single-stranded. The single-stranded DNA overhangs are complementary to each other on adjacent molecules, and thus are able to anneal (Step 2). If the ends of the molecule being constructed are complementary, then the final construct will be circular (this is the normal method used when DNA is being assembled into a cloning or expression vector). The final step to the reaction is repair of the DNA. In the *in vitro* reaction, a DNA polymerase is used to fill the gaps, while a DNA ligase seals the nicks so as to create the larger assembled DNA molecule. The DNA repair activity of *E. coli* can be utilized to complete the reaction after the annealing step

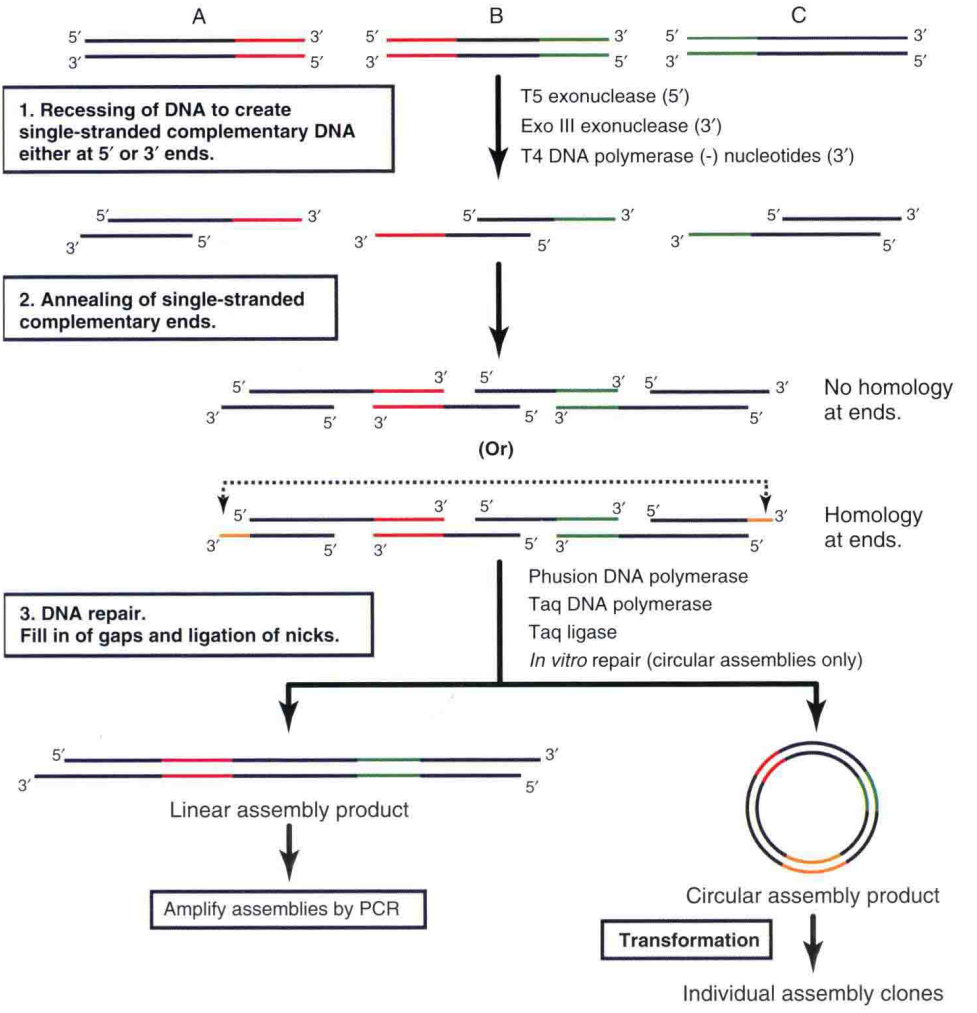


Fig. 2 Schematic depicting *in vitro* homologous recombination DNA assembly. First, nucleotides are removed from either the 5′ or the 3′ ends of the DNA pieces (5′ removal depicted). This step can be performed by several enzymes. The newly exposed single-stranded homologous ends (red, green, or yellow regions) on the adjacent pieces are complementary, and can anneal. Providing homology at the ends of the DNA pieces will result in the assembly of a circular DNA molecule. Following the annealing step, the DNA is repaired by filling in the gaps with a DNA polymerase and sealing any nicks with DNA ligase. A linear assembly product can be amplified by using the PCR and used in further assembly reactions. The circular assembly products are transformed into the appropriate host in order to isolated individual clones with the final assembled molecule. The assemblies can also be repaired *in vivo* after transformation by the native activities of *E. coli*.