DNA Cloning

A brief overview of cloning and methods

For use with IDT gene fragments:

- eBlocks™ Gene Fragments
- gBlocks™ Gene Fragments
- gBlocks HiFi Gene Fragments

For Research Use Only. Not for diagnostic procedures. | Version 1

Table of contents

INTRODUCTION: RECOMBINANT DNA TECHNOLOGY

DNA cloning is the cornerstone of biotechnology. The cloning process comprises multiple steps from generating recombinant DNA and vector constructs, to introducing them into host cells. As such, recombinant DNA technology is the foundation for a range of synthetic biology products like vaccines, biologics, enzymes, new crop varieties, microbial strains for bioremediation, research models, biodegradable materials, and biofuels, among others.

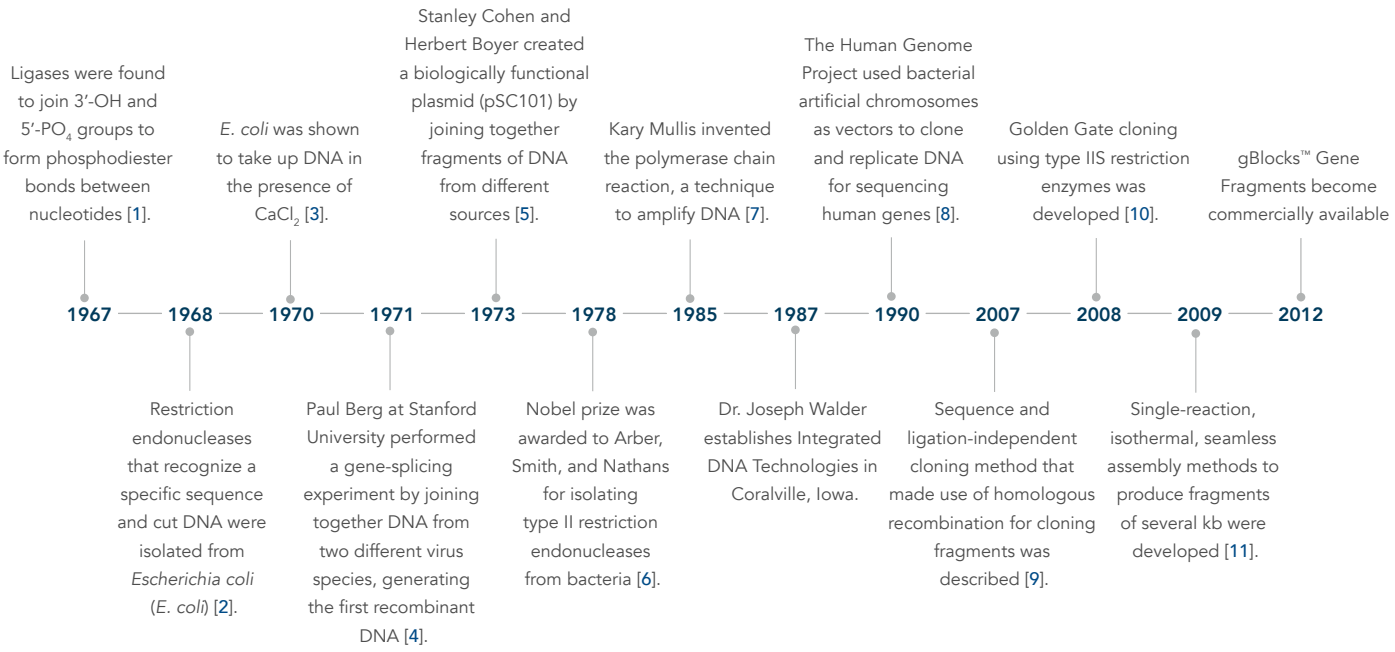
The core of the cloning process is to generate recombinant DNA. For this, DNA fragments from different sources are joined together and introduced into a DNA-carrying vehicle, a so-called vector. Vectors in DNA cloning include plasmids, attenuated viruses or bacteriophages, or larger constructs such as artificial chromosomes. Finally, the resulting recombinant DNA construct is transformed into host cells which use their own cellular machinery to replicate and express the construct.

In this handbook, we will introduce some primary cloning techniques and offer some practical tips to help you plan your experiments.

First, let's quickly review of the history of DNA cloning.

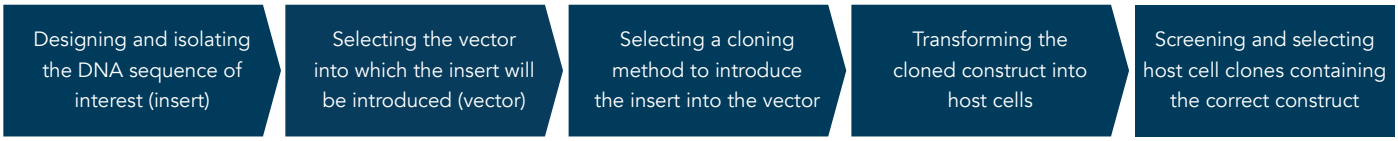
A BRIEF HISTORY OF DNA CLONING

Before advanced synthetic biology products were created, the ability to combine DNA from different sources was first described in the 1970s. Since then, several commercially available technologies evolved that made manipulating DNA easier. Key milestones in the development of recombinant DNA technologies are captured below:



THE DNA CLONING WORKFLOW

Molecular DNA cloning generates constructs that can be used to create new biologically derived products. The basic workflow for DNA cloning consists of the following steps:



Each experimental step requires diligent planning and design. The following section explores each step in more detail before introducing the basic techniques to assemble DNA constructs. With this handbook, we further offer practical tips to plan cloning workflows.

Designing and isolating the DNA sequence of interest (insert)

To start a DNA cloning experiment, identify and design the DNA sequence or gene(s) of interest to clone; this is the insert. This DNA sequence can be artificially synthesized or amplified by polymerase chain reaction (PCR) from an existing template, such as genomic DNA, RNA, or another plasmid.

During the design step, consider adding elements such as restriction enzyme sites, phosphate groups, or complementary sequence overlaps to the DNA sequence. Determine if a single sequence will be cloned or multiple sequences assembled to create a larger construct. As such, the design process depends on the chosen cloning and DNA synthesis technique. To help with the DNA synthesis process, Integrated DNA Technologies Inc. (IDT) manufactures different types of synthetic single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) such as gBlocks™ or eBlocks™ Gene Fragments.

Selecting the vector into which the insert will be introduced (vector)

Vectors are circular DNA molecules into which the DNA sequence of interest can be cloned. A cloning vector typically contains an origin site to initiate plasmid replication within bacterial cells, a multiple cloning site into which sequences of interest are inserted using restriction enzymes, and an antibiotic resistance gene to select for host cells containing the vector.

Depending on the objective of the experiment and the host organism, the insert can be integrated into different vectors. To express the gene product within host cells, use a plasmid expression vector. Expression vectors often contain host-specific regulatory elements such as promoter and terminator sequences that drive protein expression. Other types of vectors include viral (bacteriophage), bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs), and mammalian cell-specific vectors [12]. Vendors providing different types of cloning and expression vectors also have resources and recommendations on choosing optimal systems while IDT also provides DNA sequences already in bacterial vectors.

Selecting a cloning method to introduce the insert into the vector

Several methods can be used to insert DNA fragments or genes of interest into vectors. The most common approaches rely on restriction enzymes to bind and cut DNA fragments at specific recognition sites. After this restriction digest, ligases link the insert with the vector. Other methods such as site-specific recombination and seamless cloning are also widely used. The next section outlines commonly used techniques to introduce DNA fragments into vectors and design considerations for each of them.

Transforming the cloned construct into host cells

Once the desired DNA fragments have been assembled into a recombinant DNA construct, it is transformed into host cells. Some bacterial and other microbial strains are naturally "competent" or have been developed to be "competent". These strains are capable of intaking foreign DNA and generating multiple copies of these vectors.

The process of transformation introduces the vector into the host organism, followed by plating of the transformed cells on solid agar growth media. Antibiotic resistance genes in the recombinant vectors allow only the transformed cells to survive on the growth media containing the antibiotic. Any cells that did not uptake the recombinant vector lack the antibiotic resistance and thus will not survive.

Screening and selecting host cell clones containing the correct construct

Individual bacterial colonies on agar plates represent a genetically clonal population. These should be screened to ensure that the recombinant vector is present in the cell and the DNA fragment of interest was inserted in the correct location and orientation within the vector. Common colony screening methods include cutting with restriction enzymes, PCR, or sequencing.

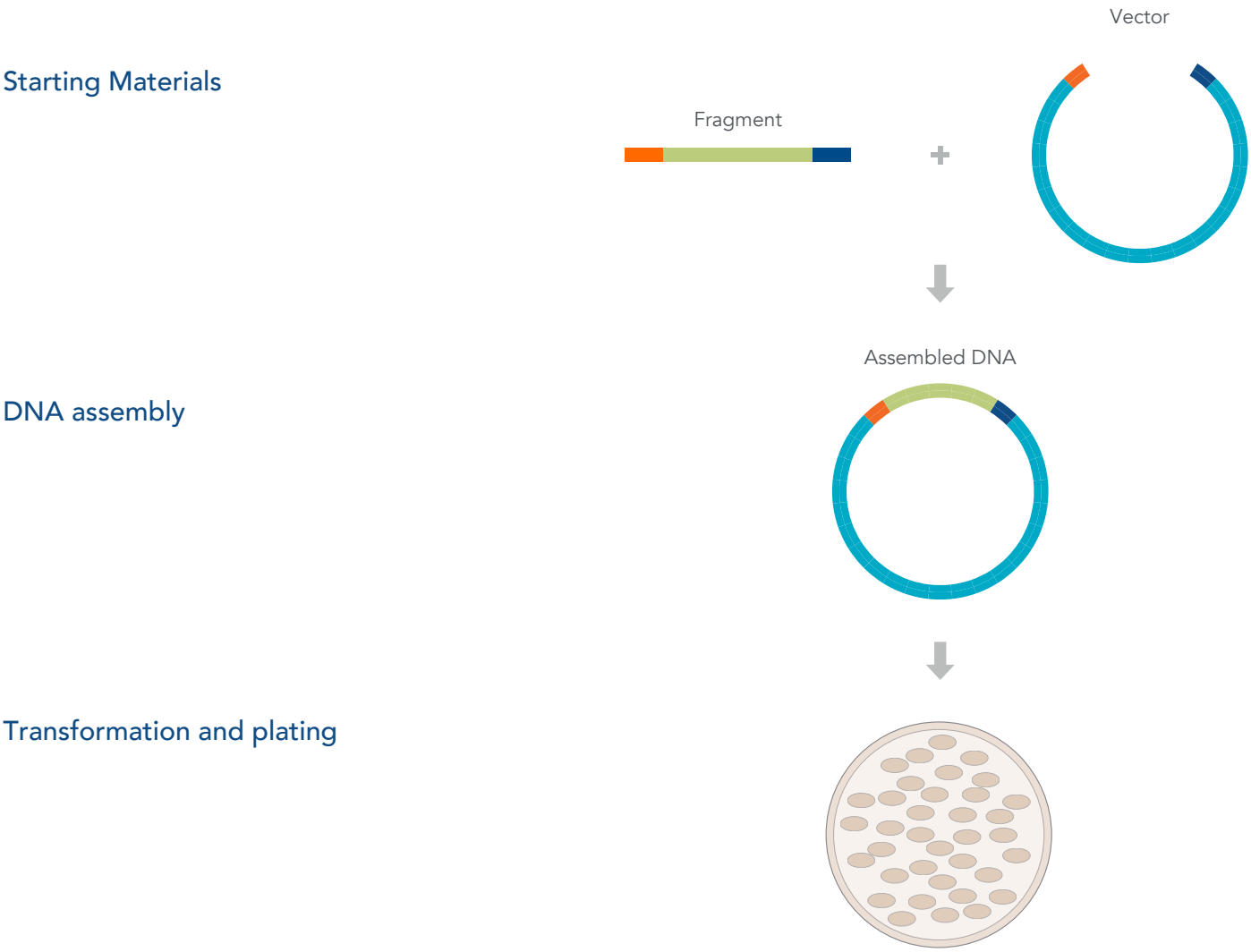


Figure 1. An overview of cloning: A standard workflow for DNA assembly, transformation, and plating. The DNA fragment is ligated into a vector to form an assembled construct, transformed, and plated on a solid agar growth medium.

CLONING METHODS—DIFFERENT WAYS TO CREATE A RECOMBINANT VECTOR CONSTRUCT

To insert the desired DNA fragment into a vector of choice, different methods can be used. Some make use of restriction enzymes that recognize specific base pair sequences and cut both DNA strands. Cutting both the vector and the insert with the same enzymes creates compatible ends that are joined using DNA ligase.

Other methods do not require restriction enzymes, but rather optimal design of the insert DNA. The choice of cloning method depends on the protocol, specification of the research experiment, and required cloning efficiency. Consider whether a single insert will be cloned, or multiple fragments assembled into a larger construct. Not all methods for cloning single constructs are compatible with multi-fragment assembly. By understanding each method in more detail, you will be able to choose the optimal cloning method for your application.

1. Blunt-end cloning

Blunt-end cloning uses restriction enzymes that generate blunt ends on both the DNA fragment and the vector. Alternatively, synthetic double-stranded DNA or PCR products with 5′-PO_4 ends can be added directly into blunt-end assembly reactions. However, in this case, it is essential to remove the phosphate groups from the vector, as these can otherwise cause the empty vector to ligate without an insert. To prevent re-ligation, vectors can be dephosphorylated by calf intestinal alkaline phosphatase, or other available phosphatases. In the last step, specific ligases such as the T4 ligase join the blunt-ended insert and vector by covalently linking their 3′-OH and 5′-PO_4 ends.

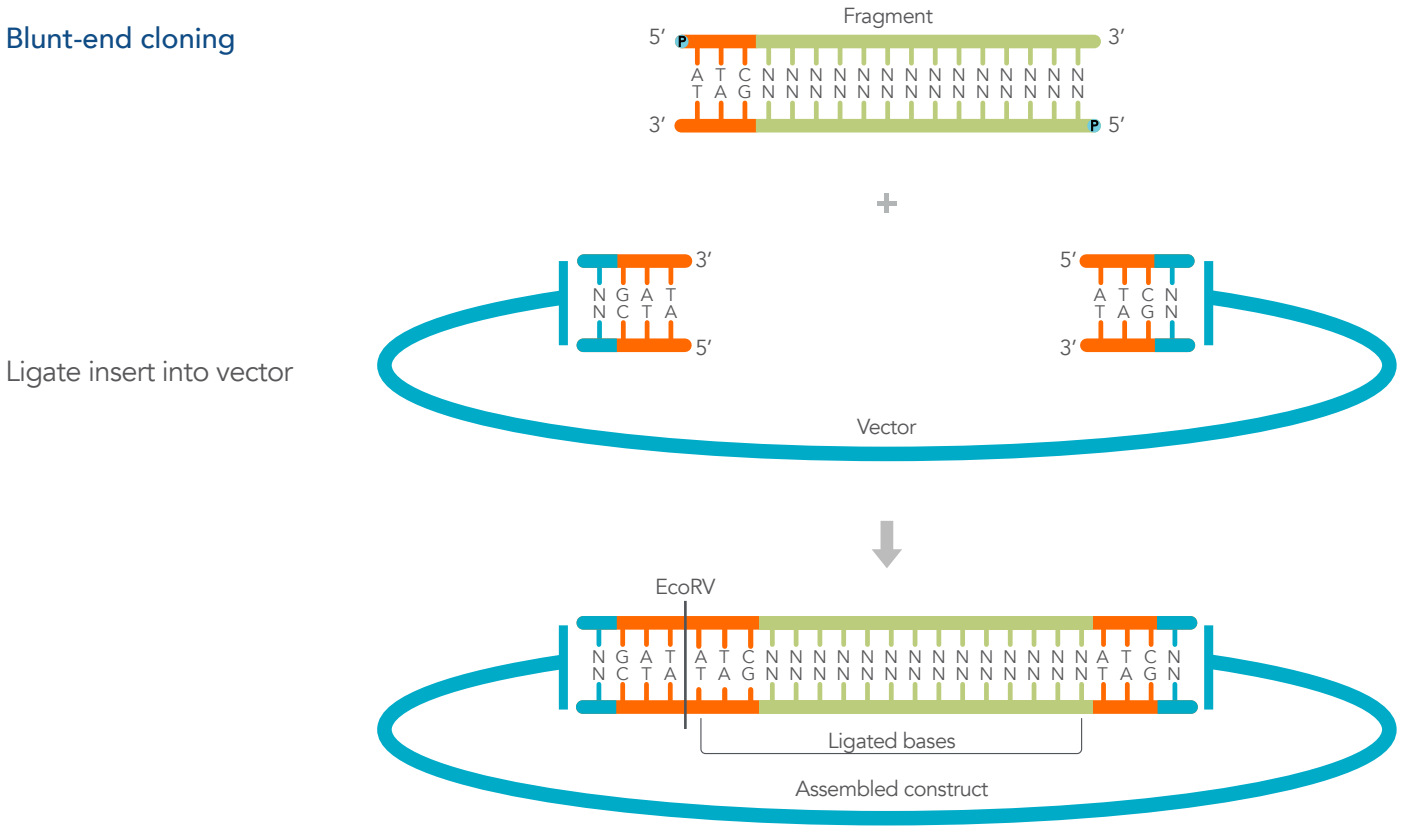


Figure 2. An overview of blunt-end cloning. The DNA fragment is ligated into a vector using a T4 ligase which covalently links the 3′-OH and 5′-PO_4 ends; however, the vector does not contain a 5′-PO_4 group, preventing the possibility of re-ligation and the formation of an empty construct with no DNA fragment inserted. A newly single EcoRV site is created on one end of the assembled construct so that a possible re-cut can occur, as shown in the figure.

Applications and limitations

- Blunt-end cloning is a fast and simple ligation-based cloning method for which DNA fragments/PCR products can be easily prepared.
- For sequences that are challenging to clone using recombination-based cloning techniques or when the desired restriction sites cannot be added, blunt-end cloning can be a suitable alternative.
- Blunt-end cloning is not directional. Ligation of fragments with phosphate groups on both ends can occur in either orientation.
- The inability to direct the orientation of the insert makes it unsuitable for assembling multiple fragments in a single reaction.
- Blunt-end cloning efficiencies decrease due to re-ligation of empty vectors if the vectors have not been dephosphorylated.

Design considerations

- Synthetic fragments for blunt-end cloning can be designed with 5′-PO_4 ends, or phosphates can be added by T4 polynucleotide kinase.
- When amplifying inserts from template DNA, PCR primers can be phosphorylated.
- If vectors do not have suitable restriction sites to generate blunt ends, they can be amplified or linearized by PCR. As 5′-PO_4 vector ends may re-ligate, ensure that the vector is dephosphorylated.

2. TA cloning

Cloning with Taq polymerase using TA-overhangs relies on the inherent property of the Taq polymerase enzyme to add deoxyadenosine (A) to the 3′-OH ends of double-stranded DNA fragments. Linearized cloning vectors with 3′ thymidine (T) overhangs designed to be compatible with TA cloning are commercially available. Alternatively, plasmid vectors can be first treated with a blunt-end restriction enzyme, followed by terminal transferase and dideoxythymidine triphosphate to generate single T-overhangs for TA cloning [12]. The insert then joins with the vector through A-T complementation and DNA ligase enzymatically links the two fragments generating the recombinant construct. Note that this leaves a minimal scar of an additional AT-pair in the final construct that needs to be considered in downstream applications.

TA cloning

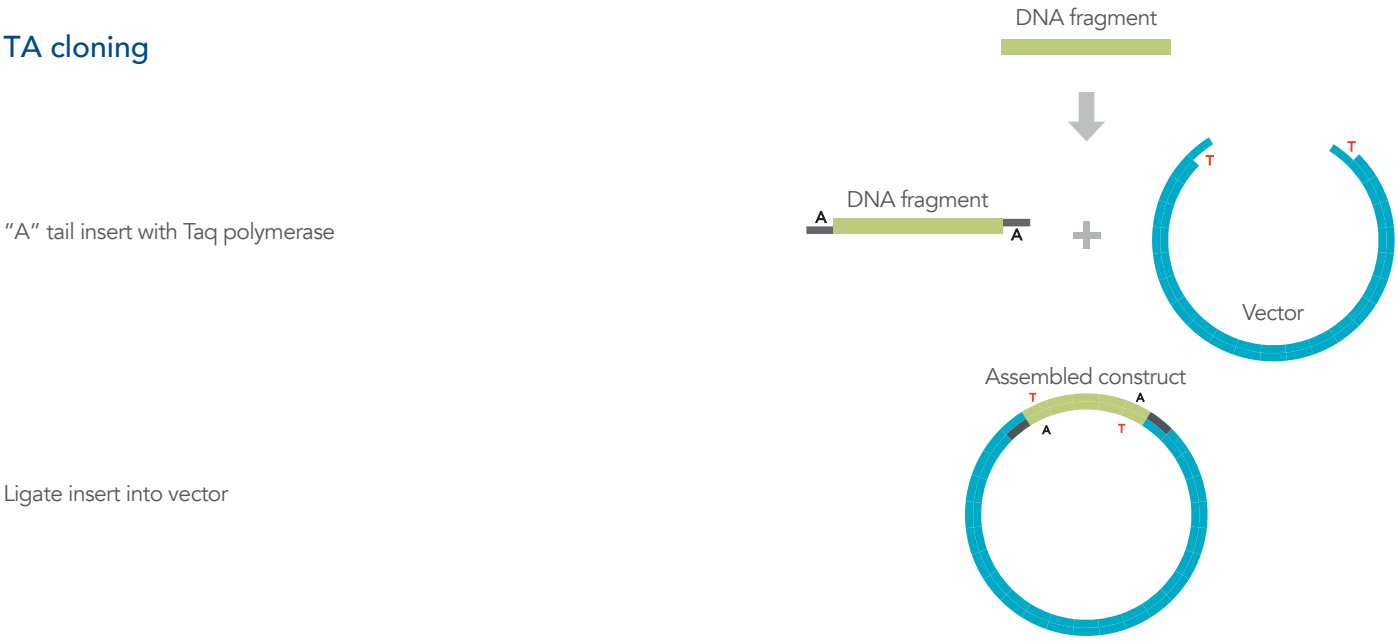


Figure 3. An overview of TA cloning using Taq polymerase and recombinant DNA technology to form an assembled construct. Taq polymerase inserts a deoxyadenosine triphosphate (A) tail to the 3′-ends of the DNA fragment. The DNA fragment with an attached A-tail gets ligated into the vector containing 3′ thymidine (T) overhangs to form an assembled construct.

Applications and limitations

- TA cloning is fast and thus suitable for quick and simple applications. The technique does not require additional modifications beyond the enzymatic addition of T and A bases. Use TA cloning when having unique restriction enzyme sites within the insert is challenging for cloning, or residual bases from enzyme sites might interfere with the expression of the construct.
- TA cloning relies on complementary T and A bases on the ends of both vector and insert. Hence, there is a theoretical 50% chance of the insert ligating in reverse orientation (actual frequencies vary). This lowers the efficiency of cloning, requiring a larger number of colonies to be screened to identify clones containing the DNA fragment in the desired direction.
- TA cloning is not ideal for assembling multiple fragments into a construct.

Design considerations

- When using synthetic DNA fragments for TA cloning, sequences can be ordered without any modifications. Once received, A-overhangs can be added to the fragments enzymatically via incubation with Taq polymerase.
- For PCR amplified fragments, A-overhangs can be designed into amplifying primers.

3. Restriction enzyme cloning

Type II restriction enzymes recognize specific DNA sequences. They then cut within this recognition site at shifted sites in both DNA strands. This generates single-stranded overhangs, so-called sticky ends. These are complementary, thus allowing a cut end to join with a fragment that has been cleaved by the same or a compatible enzyme. Hence, this method provides specificity over blunt-end cloning since only compatible ends join. Additionally, two different restriction enzymes with unique restriction sites can be used to cut both insert and vector. Like this, the insert will join with the vector only in one orientation allowing directional cloning and improving cloning efficiency. Finally, DNA ligase covalently links the 3′-OH and 5′-PO_4 ends of insert and vector resulting in a recombinant vector with an insert cloned in the intended orientation.

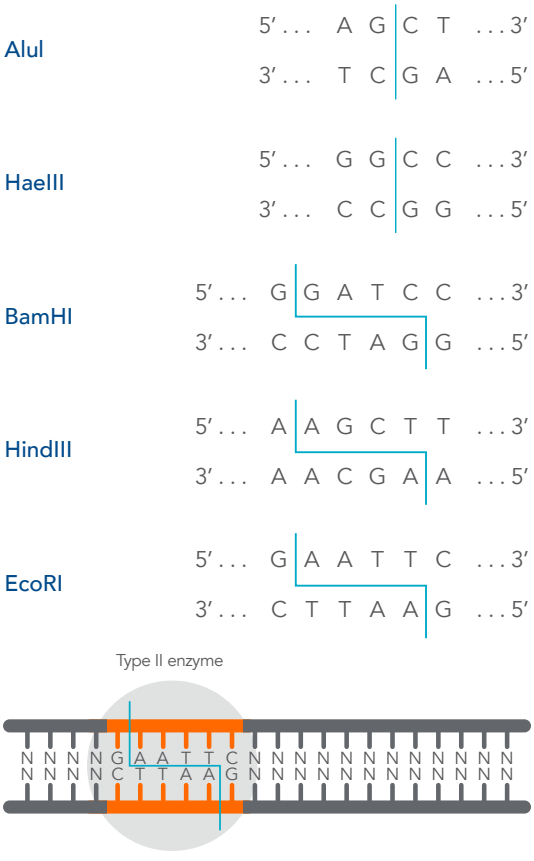


Figure 4. Restriction enzymes make incisions in the DNA strands. Type II restriction enzymes recognize and make incisions on specific sites in the DNA strands, which results in blunt or sticky ends. Enzymes, such as AluI and HaeIII, make incisions in the DNA strands, thus resulting in unpaired nucleotides, referred to as blunt ends. Restriction enzymes, such as BamHI, HindIII, and EcoRI, create incisions in the DNA strands that result in a short nucleotide overhang, referred to as sticky ends.

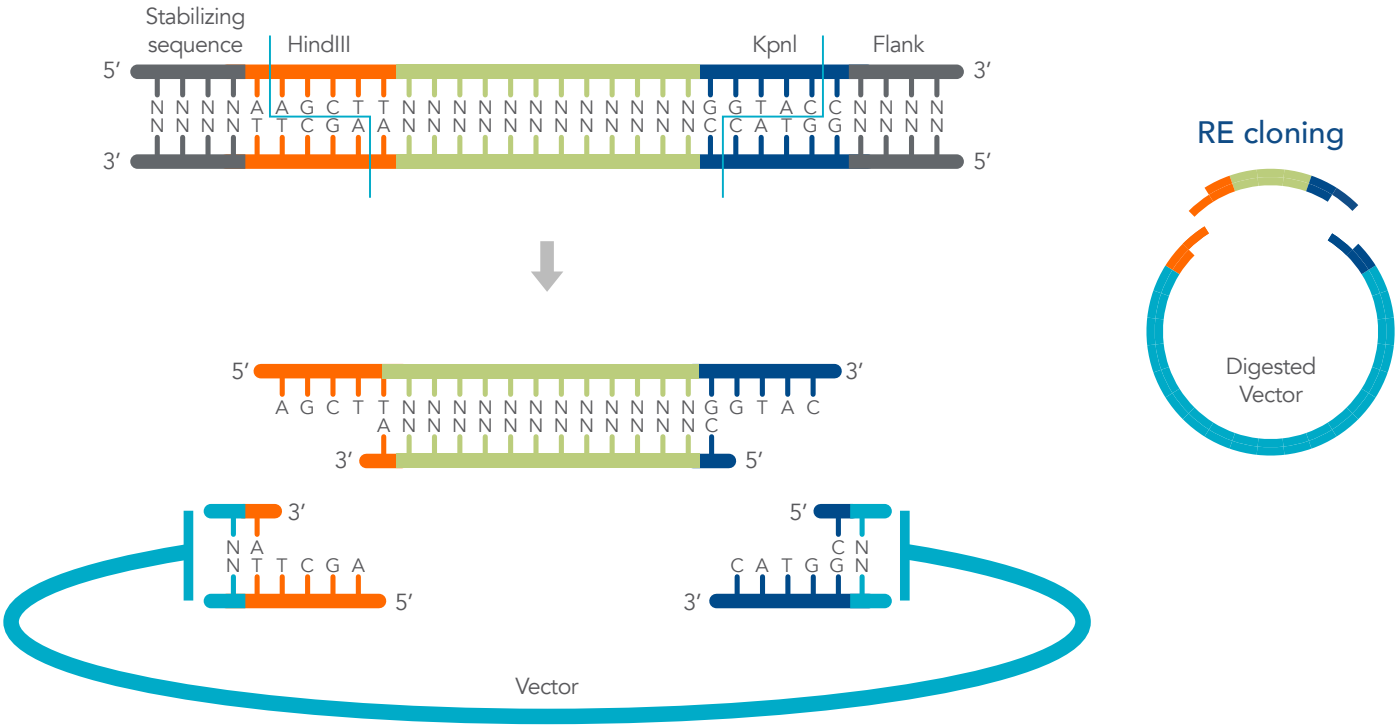


Figure 5. Restriction enzyme digestion, ligation, and DNA base pair complementation result in a recombinant DNA vector. The figure shows that restriction enzymes HindIII and KpnI have unique recognition sites in the DNA sequence that creates a short nucleotide sequence overhang. The resulting digested DNA target and the nucleotide sequence of the vector are oriented in one direction so that complementary DNA base pairing can occur. The DNA target is ligated to the vector using DNA ligase.

Applications and limitations

- Sequences with unique type II restriction sites (or those designed to contain unique sites) can be most effectively cloned using this method. An advantage of restriction enzyme cloning over TA and blunt-cloning methods is directional cloning of the insert. Cloning efficiencies are thus higher.
- Restriction sites used for cloning must only appear at the end of the insert and only in the multiple cloning site or polylinker section of the vector. Sometimes, it can be difficult to identify unique enzyme sites or incorporate them in the fragment and vector design. In such cases, consider an alternative cloning method.
- Consider the overhangs or residual bases of the enzyme cleavage sites when designing inserts. These 'scars' left from restriction enzyme sites could interfere with the reading frame when expressing the DNA sequence into protein.
- While most enzymes leave longer overhangs, those resulting in shorter overhangs (1–2 bases) may cause the vector to reanneal, possibly lowering cloning efficiency.
- For constructs requiring the assembly of multiple fragments, other methods such as Golden Gate assembly or seamless cloning might be better suited.

Design considerations

- For this method, design DNA fragments such as gBlocks™ or eBlocks™ Gene Fragments and include recognition sites for restriction enzymes. These sites must be extended by a few bases (4–5) so that the restriction enzymes can bind more efficiently. These 4–5 'flanking' bases can be random and will improve sequence recognition and DNA cleavage.
- With codon optimization, alternative codons can be introduced to result in including or excluding specific restriction sites. If there is flexibility in redesigning the original sequence (while preserving the amino acid sequence) then consider this option to include unique restriction sites in fragments.
- For PCR-amplified fragments, restriction sites can be added to the 5′ end of primers. The sites will thus be incorporated into the inserts. Consider incorporating two different restriction sites on either end of the insert to impart directionality and improve cloning efficiency.

4. Golden Gate cloning

Cloning with type IIS restriction enzymes using the Golden Gate method is also based on complementing sticky ends but employs type IIS restriction enzymes. In comparison to type II restriction enzymes, which cut DNA strands within their recognition site, type IIS restriction enzymes cut outside of their recognition site. As a result, the overhangs produced after restriction digest are independent of the recognition site and unique to the sequence of the DNA fragments. Hence, after ligating and joining the fragments, no residual bases of the recognition sequence—a so-called scar—will be present in the cloned construct.

Different type IIS restriction enzymes cut the DNA at different locations outside of their recognition site. This produces overhang sequences of different lengths. For example, a type IIS enzyme cleaves the sense strand 1 base downstream of the recognition site (towards the 3′ end) and the antisense strand 5 bases towards the same direction. This leaves 3′ overhangs of 4 bases in length. Hence, the DNA fragment to be joined with must contain overlaps of exactly these 4 base pairs in the same direction. Using this cloning strategy, multiple fragments can be assembled as one insert into a vector in a single ligation reaction. For this, design DNA fragments with type IIS recognition sites at the terminal ends and overlaps with the adjacent fragments.

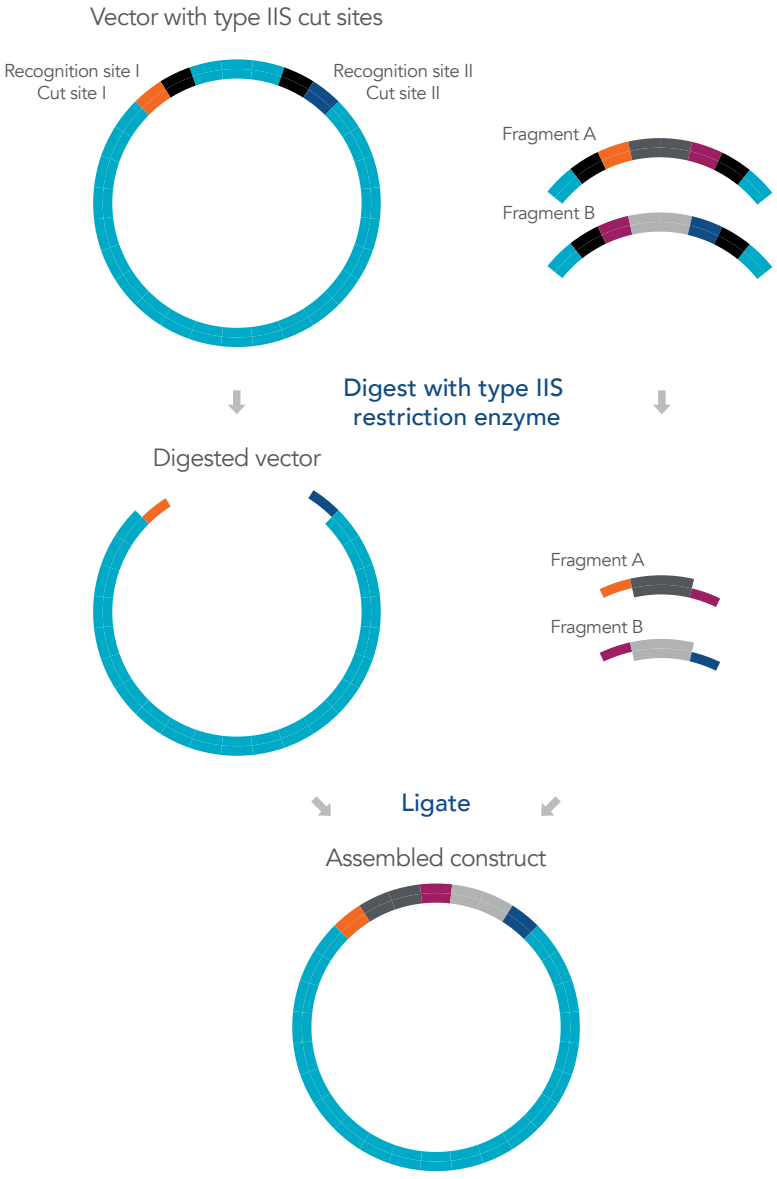


Figure 6. An overview of the Golden Gate cloning method. The Golden Gate cloning method inserts more than one target DNA sequence into a vector. The vector and the DNA fragments, designated as fragments A and B, contain recognition sites (black) and cut sites (orange and navy blue). Both vector and DNA fragments are digested with type IIS restriction enzymes, which generate a short nucleotide sequence overhang. The digested vector and the DNA fragments do not include the recognition sites, as shown in the figure. The overhang regions (designated as orange and navy blue in the figure) ligate to form an assembled construct.

Applications and limitations

- Golden Gate cloning is scarless and can be used for cloning single inserts as well as assembling large constructs originating from multiple fragments.
- This method is suitable for assembling multiple DNA fragments or fragments that contain internal structural complexities.
- GC extremities, repeats, or secondary structures in the overlap regions could be problematic.
- Like other restriction-enzyme based methods, for Golden Gate assembly to work, make sure the type IIS enzyme recognition site is uniquely present in the cloning site. If the recognition site occurs anywhere else in the insert or the vector, it will need to be removed to prevent undesired cuts.
- Overhangs with similar sequences (e.g., only 1 base pair difference) may re-ligate the vector unintentionally and lower cloning efficiency.

Design considerations

- When designing DNA fragment(s) for Golden Gate cloning, include the cut sequence of the type IIS enzyme in the terminal regions of the insert. Ideally, sequences containing 4-base overhangs have the highest accuracy. Add the restriction enzyme recognition site terminal to the cut site. Make sure the recognition and cut sites are also added to the vector so that the overlap joins together.
- Note, that some enzymes require a spacer sequence between the recognition site and the cut site. For restriction enzymes that cut the sense strand one base pair outside of the recognition site, include a one base pair spacer nucleotide after the recognition sequence and before the cut sequence. For enzymes that cut after two base pairs, 2 base pair spacer nucleotides need to be included, and so on.
- Also, a 4–5 base pair flanking sequence facilitates the recruitment of enzymes to the DNA template. Hence, add such a sequence to both terminal ends of the fragment to increase reaction efficiency.

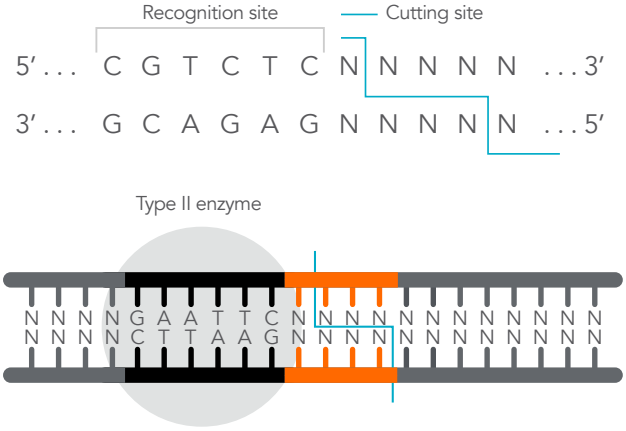


Figure 7. The restriction enzyme's recognition site differs from its cutting site. A Type II restriction enzyme cuts the non-specific nucleotide sequence away from its recognition site in the DNA molecule.

5. Seamless cloning

There are several different commercialized master mixes used for seamless cloning methods, such as InFusion™, which do not require restriction enzyme digestion. Instead, this method requires inserts to contain overlaps of 15–40 bases at their terminal ends. These overlaps need to be complementary to the insertion sites on the vector. Integrate the overlaps into the DNA fragment sequence in the design step or add the overlap by PCR.

The reaction requires the activity of three enzymes: an exonuclease enzyme digests the 5′ ends of the fragments when incubated at 50°C. This lays open the 3′ single-stranded overhangs. The temperature facilitates annealing of complementary sequences of the vector and insert. Next, DNA polymerase adds bases to fill gaps and DNA ligase forms covalent bonds between the complementary regions of insert and vector.

Prepare starting material

Linear vector and DNA insert fragments with 15–40 bp overlapping ends

Single-tube reaction

DNA fragments and enzyme master mix

5′ exonuclease

DNA polymerase

DNA ligase

Incubate

Single step

15–60 minutes at 50°C

Transformation and plating

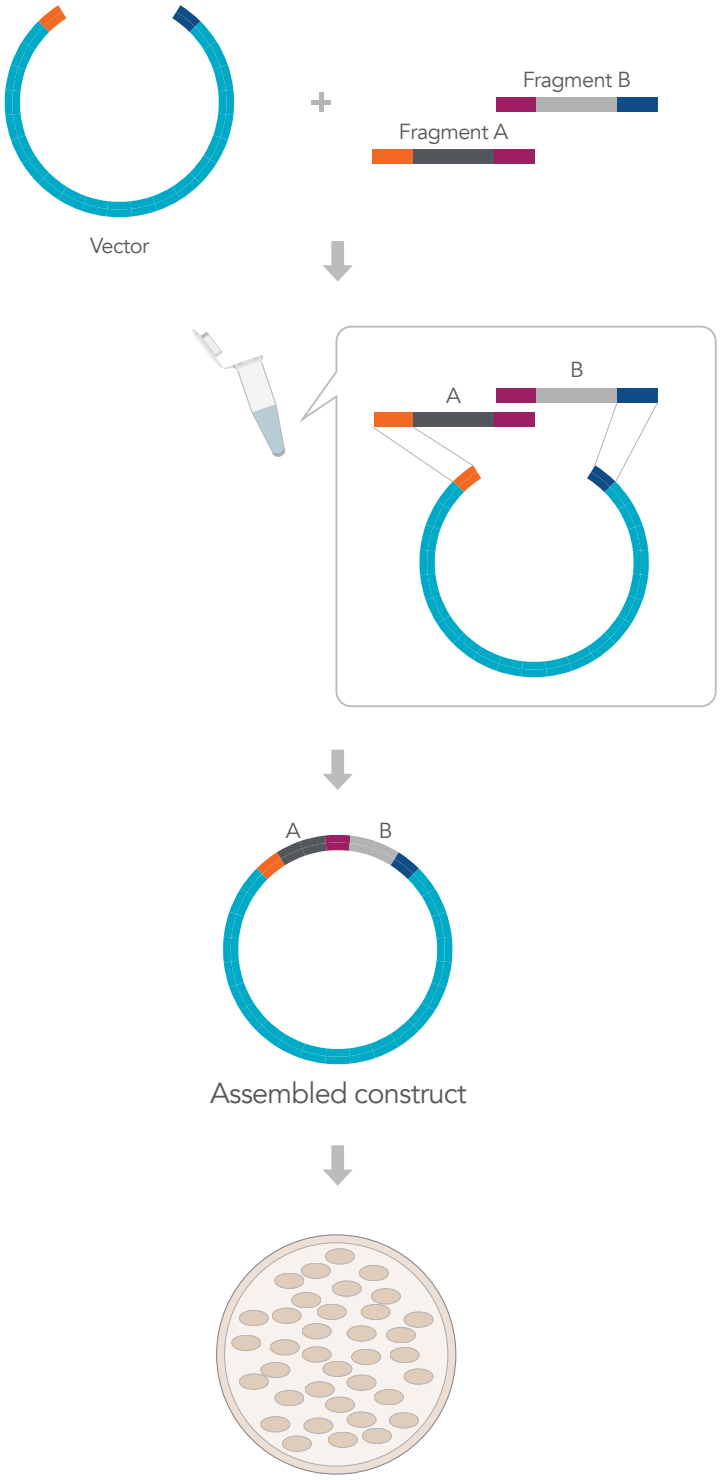


Figure 8. An overview of the seamless cloning method. Like the Golden Gate cloning method, this method incorporates cloning of more than one DNA fragment into a vector. Although this method does not require restriction enzyme digestion, the vector and the DNA fragments require insertion of a 15–40 base pair sequence overhang (depicted as orange and blue in the figure). Instead, the DNA fragments and an enzyme master mix containing 5′exonuclease, DNA polymerase, and DNA ligase are combined in a single-tube reaction. After an incubation period, the assembled construct is transformed into a host organism then plated on a solid agar growth medium.

Applications and limitations

- Like the Golden Gate method, seamless cloning is especially useful for assembling large constructs, and multiple fragments with high rates of efficiency.
- Seamless cloning is a robust and convenient method that does not require restriction sites and leaves no scars in the DNA sequence.
- With appropriate sequence overlaps, it can be used to clone single fragments or even assemble larger constructs efficiently.
- Seamless cloning is sensitive to sequence complexities since it requires overlaps of up to 40 bases between adjacent fragments. Complexities such as GC extremes, secondary structures, or repeats in the overlapping regions can become problematic when trying to join fragments.

Design considerations

- Fragments to be assembled must have double-stranded sequence overlaps of 15–40 base pairs with adjacent fragments. These overlapping regions could be on the vector and/or with other DNA fragments to be assembled.
- Evaluate the melting temperature of the overlapping regions and refer to a cloning kit for guidance. Consider adjusting the melting temperature by increasing or decreasing the length of overlaps (within 15–40 base pairs).
- Tandem repeats in the overlap sequences can result in misaligning and incorrect assembly. Check to ensure that the overlaps are free of repeats.
- GC extremities, repeats, or secondary structures in the overlap regions could be problematic.

6. Gateway cloning

Gateway™ cloning uses a recombination-based cloning technique that allows shuffling DNA fragments between vectors. Commercialized by Invitrogen/Thermo Fisher Scientific [13] this cloning technique is based on a bacteriophage lambda site-specific recombination system. This approach uses enzymes to mediate the recombination of DNA sequences flanked by homologous attachment (*att*) sites.

By ligation or PCR, the DNA fragment is first inserted into a vector so that it contains *att*L sites on both flanking ends. The *att*L sites of this so-called entry clone are homologous to the *att*R sites on the destination vector. This contains a negative selection marker *ccd*B flanked by the *att*R sites while *AmpR* serves as a positive selection marker. In the so-called LR recombination reaction, LR Clonase™ mediates the recombination between the *att*L and *att*R sites. Thus, the DNA fragment in the entry clone swaps place with the *ccd*B gene in the destination vector. As a result, the destination vector becomes an expression clone containing the DNA fragment flanked by *att*B sites and the entry clone becomes a donor vector containing the *ccd*B gene flanked by *att*P sites.

Note, that Gateway cloning reaction is reversible. LR Clonase exchanges the DNA fragment with *ccd*B, and thus moves the fragment from an entry vector to the destination vector. In the reverse BP reaction, the flanking *att*B sites on the expression clone can recombine with the *att*P sites of the donor vector. Like this, the DNA fragment can be excised from the expression clone and into the entry vector, recreating the entry clone. Antibiotic resistance selection genes on the vectors and the toxin-producing *ccd*B gene assure that the final living host bacteria harbor the entry clone with the desired DNA fragment.

Figure 9. An overview of the Gateway™ cloning method. This recombination-based cloning technique is based on a bacteriophage lambda site-specific recombination system and mediates the shuffling of DNA fragments between vectors. The vectors have two homologous (*att*) sites that flank the DNA fragment. The destination vector contains a negative marker (*ccdB*) and a positive marker (*AmpR*), and an antibiotic resistance gene to facilitate the target gene uptake. This reversible reaction utilizes LR clonase for the forward reaction and BP clonase for the reverse reaction.

Applications and limitations

- Gateway cloning is a fast and efficient technique to move DNA fragments between vectors in a ligase-independent manner. It allows the transfer into any destination vector with *att*R sites.
- The presence of a negative selection marker (*ccd*B) improves cloning efficiency. Constructs that do not contain the DNA fragment carry the *ccd*B instead. Since its gene product is toxic to host cells, only transformants with the inserted construct will be viable.
- This method is limited by its requirement for specific vectors containing the *att* sites. If the expression clone or the entry clone is not suitable for the experimental design, then consider an alternative cloning method.
- Gateway cloning cannot be used to assemble multiple fragments due to the presence of *att* sites. When annealing fragments from different vectors, the intervening *att* sites will disrupt the frame. Therefore, this method is best suited for single fragment cloning.

Design considerations

- Synthetic DNA fragments or PCR products to be used with the Gateway method require the recombinase-specific *att* sites flanking the sequence of interest. Add the *att*L sites and *att*B sites to the DNA fragment to result in the entry and expression clones, respectively.
- When designing the insert, consider how the flanking *att* sites will impact the translational reading frame.

COMPARISON OF CLONING METHODS

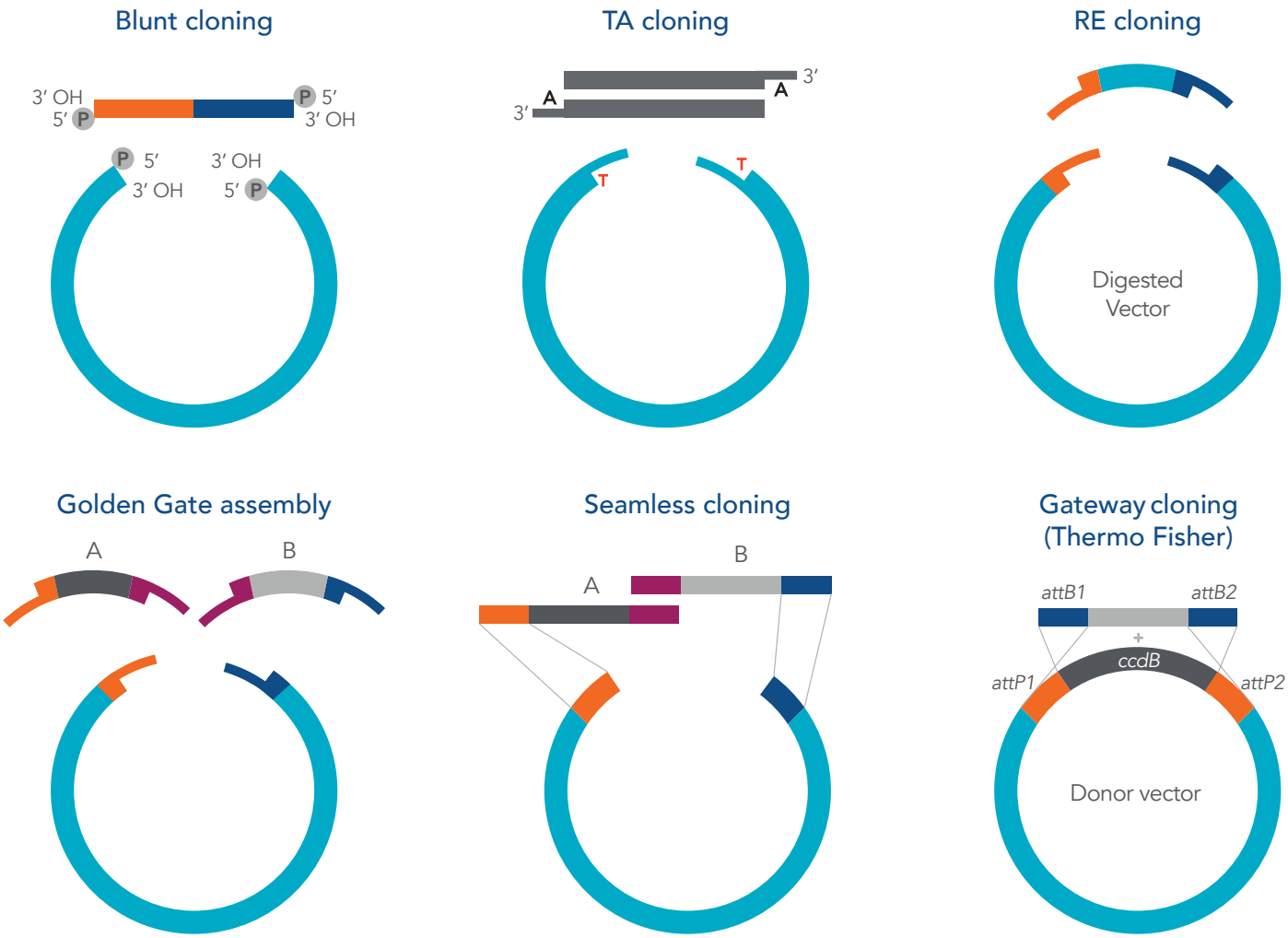


Figure 10. A comparison of cloning methods. The sections in this handbook discuss the applications and limitations of the various recombinant DNA assembly methods. Blunt-end cloning requires restriction enzymes to create blunt-ends in the DNA strands. In contrast, TA cloning requires Taq polymerase to insert deoxyadenosine (A) and thymine (T) nucleotide overhangs, creating sticky-ends in the DNA strands and vector. Similarly, restriction enzyme (RE) cloning creates a short nucleotide overhang in the DNA strands and the vector. The Golden Gate and the seamless methods assemble constructs containing more than one DNA fragment. The Gateway cloning (Thermo Fisher Scientific) method mediates the transfer of the DNA fragment into the destination vector by the homologous (*att*) sites. The presence of a negative selection marker (*ccd*B) improves cloning efficiency.

Table 1. Summary of cloning methods.

Assembly method	Pro	Con
Blunt-end cloning TA cloning	• Fast and easy • No extensive preparation or design steps	• Cloning is not directional • Often requires more screening • Not suitable for assembling larger constructs
Type II restriction enzyme cloning (sticky ends)	• Directional cloning with high efficiency • Unique enzyme sites ensure correct ligation • Restriction enzymes and buffers are widely available	• Requires restriction sites in both insert and vector • Leaves "scars" or residual bases in the fragment
Type IIS restriction enzyme cloning (Golden Gate Assembly)	• Scarless cloning as type IIS enzymes cleave outside of the enzyme recognition site • Sticky ends for high efficiency cloning and ligation • Method for assembling multiple fragments	• Requires restriction sites in both insert and vector • May result in non-specific ligation if overhangs of fragments are similar (only differ by a base or two)
Seamless cloning	• Fast cloning of single or multiple fragments in a single assembly reaction • Does not require restriction sites • High efficiency	• Requires overlaps of 15–40 base pairs between sequences of fragments, inserts and vectors
Gateway cloning	• Rapid exchange of inserts between vectors • Large selection of Gateway compatible plasmids available	• Requires *att* sites at flanking ends of fragment, so it leaves "scars" • Cannot be used for assembling multiple fragments

SYNTHETIC FRAGMENTS FOR CLONING AND ASSEMBLY

As convenient alternatives to PCR products, synthetic, double-stranded DNA fragments can be used in cloning and gene assembly applications. Flexible customization allows users to design and validate the synthetic fragments *in silico* while also adapting the fragments to the specific requirements of the cloning method. Hence, synthesizing DNA sequences is convenient and faster and does not require amplifying or assembling shorter regions from several templates.

For example, long fragments of up to 3 kb can accommodate protein coding regions, regulatory elements, biological parts, antibody fragments, genomic regions, and more. However, certain DNA sequences such as promoter regions might be challenging to manufacture synthetically, in which case we suggest to use a vector containing the desired elements. Otherwise, multiple fragments can be designed and assembled to generate gene circuits or constructs for expressing large proteins. Additionally, variable bases can be incorporated into sequences to generate diversity for screening and discovery applications. Use bioinformatics tools for experimental design, construction, validation, and revisions *in silico*.

IDT offers synthetic DNA fragments in the form of gBlocks Gene Fragments and eBlocks Gene Fragments. These are customizable and flexible enough for any cloning method, as well as compatible with all cloning applications. See this white paper to learn more. Double-stranded DNA fragments are delivered in several convenient options, as summarized in Table 2.

Table 2. Characteristics of synthetic DNA fragments.

	eBlocks Gene Fragments	gBlocks Gene Fragments	gBlocks HiFi Gene Fragments
Length (base pairs)	300–900	125–3000	1000–3000
Deliverable	dsDNA fragment	dsDNA fragment	dsDNA fragment
Yields	200 ng	250-1000 ng	1000 ng
Quality control	CE	CE/MS	NGS
Median error rate	1:5000	1:5000	1:12,000
Formats	96-well or 384-well plates	Tubes, 96-well plates	Tubes
Estimated TAT (business days)	1–3	2–8*	6–10
Applications	High-throughput screening and antibody discovery	Gene construction and controls for qPCR, NGS	Pathway design and large construct assembly

* TAT = turnaround time; applicable to tube orders; plate orders have a 10–15 day TAT.

IDT gene fragments are manufactured using high-fidelity Ultramer™ DNA oligomers with error correction measures to provide products for cloning and gene assembly applications. An internal study showed that a correct clone can be found by screening 2–4 colonies. Based on this internal study, Table 3 extrapolates how many colonies would need to be screened for a 90% chance of finding a correct clone. Values represent the results from an in-house experiment that used a seamless cloning method. The rate of finding the correct clone in four or fewer colonies was >90% for fragments of 500 to 2500 base pairs.

Table 3. The approximate number of colonies to screen for a 90% chance of getting a correct clone.

Sequence length (bp)	eBlocks Gene Fragments	gBlocks Gene Fragments	gBlocks HiFi Gene Fragments
500	2	2	N/A
900	3	3	
1500	N/A		2
2000		4	
2500			

GUIDELINES TO DESIGN SYNTHETIC DNA FRAGMENTS

DNA sequence manipulation software can help in planning and experimental design. Tools for annotating sequences, generating vector maps, analyzing GC content, structural motifs, melting temperatures, restriction enzyme sites, and visualizing cloned constructs are available. To help with the design and customization of synthetic DNA fragments, IDT developed the SciTools™ suite of **web tools**. These tools assess custom DNA sequences for complexities and optimize codon usage based on a host organism.

Sequence complexities

Challenging sequence characteristics—also called complexities—interfere with the synthesis, assembly, or even the end use of DNA fragments. Because of this, our complexity screening tool checks input sequences throughout the ordering pipeline for three types of complexities: (1) extremes in GC content (2) repeats, and (3) secondary structures. The tool points out problem regions within a DNA sequence and provides suggestions for its improvement.

Each product type has a unique scoring algorithm with acceptance cutoffs summarized in Table 4. Sequences with scores that fall into the denied category must be redesigned to resolve complexities. Due to differences in manufacturing, gBlocks HiFi Gene Fragments have a higher complexity cut-off; as a result, if a large gBlock (over 1000 bp) fails complexity screening as a gBlock Gene Fragment it may still pass complexity screening as a gBlocks HiFi Gene Fragment.

Editing a sequence due to complexities can be challenging based on the downstream application or the nature of the DNA that is being used. Here we outline three strategies to reduce or resolve sequence complexities Table 4. These strategies can be used based on the location and type of the complexity, as well as the downstream application of the DNA fragment. It may be necessary to use multiple strategies to get a passing sequence. Finally, IDT's Custom Gene product line can accommodate higher complexity scores for sequences down to 25 bases, but will be delivered in a simple plasmid backbone. If the strategies outlined below do not work this may be an option for ordering.

Table 4. Summary of complexity score cutoffs for each product type.

	eBlocks Gene Fragments	gBlocks Gene Fragments	gBlocks HiFi Gene Fragments
Accepted- Low Complexity	Less than 7	Less than 7	Less than 7
Accepted- Moderate Complexity	Between 7 and 15	Between 7 and 10	Between 7 and 19
Denied- High Complexity	Greater than 15	Greater than 10	Greater than 19

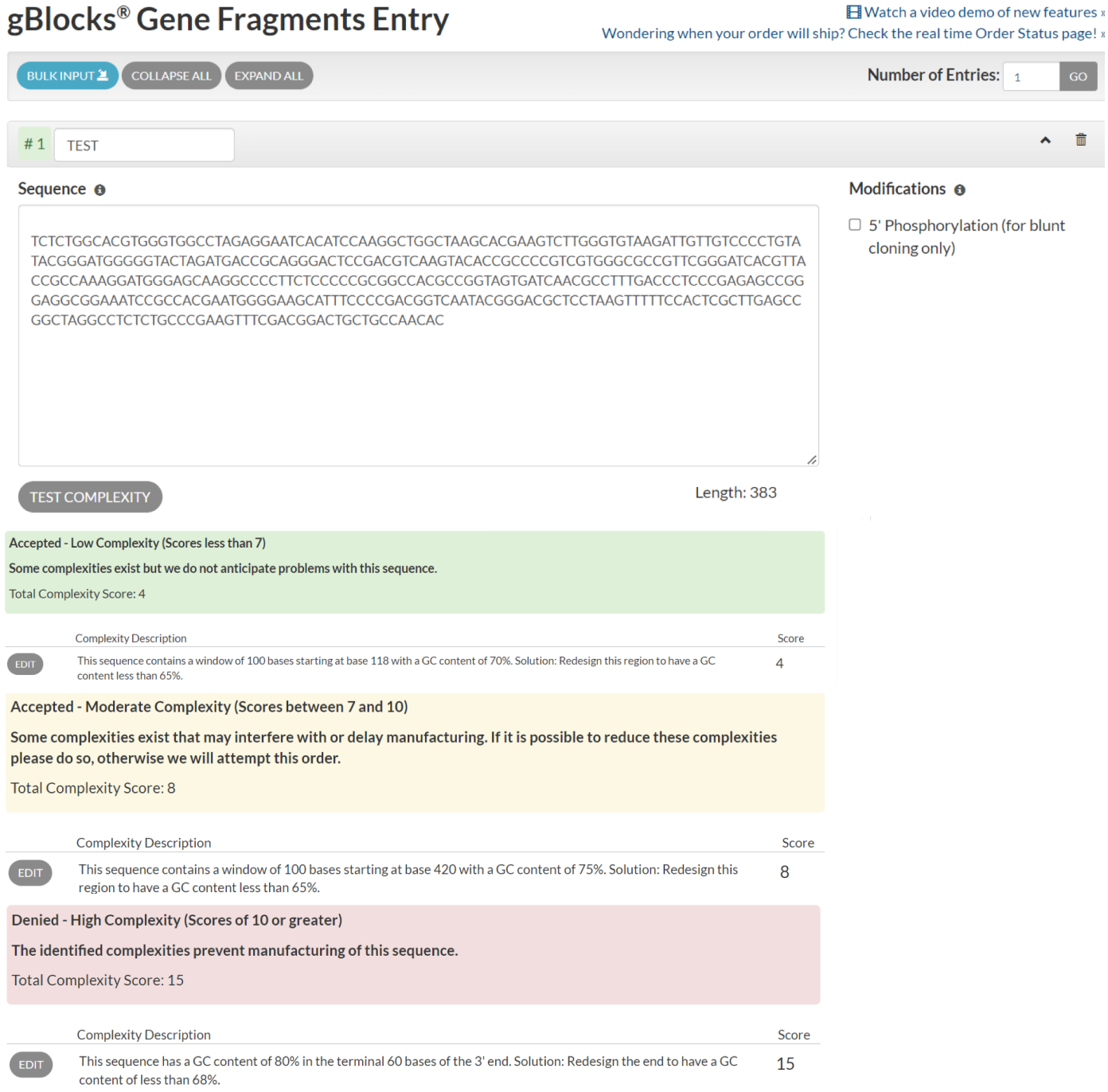


Figure 11. The ordering tool screens the complexity of the sequence using a proprietary algorithm that assesses the sequence based on GC content, repetitive sequences, and the formation of secondary structures. A sequence with a low complexity is an accepted sequence (green box). When the tool reports a moderate complexity (yellow box),this indicates that the complexities present may interfere or delay manufacturing. When there are moderate complexities, it is recommended to resolve the complexities. A high complexity (red box) indicates that the sequence is too complex to manufacture and will need to be redesigned prior to ordering.

Reduce complexity by adding a stabilizing sequence

Complexities at the terminal ends of fragments can cause manufacturing challenges. Add short stabilizing stretches of 20–60 random base pairs with approximately 50% GC content to the terminal ends. These GC-neutral sequences help normalize GC extremes and move secondary structures or repeats towards the interior of the sequence. While choosing stabilizing sequences, consider whether they follow good primer design principles. Hence, within the first and last 30 base pairs of the gBlocks Gene Fragment, avoid features that would lead to discard an oligo pair as a primer set. This may reduce complexities while preserving the original sequence.

There are some factors to consider before deciding to try using a stabilizing sequence to reduce the complexity of the DNA fragment. Stabilizing sequences are most effective at resolving complexities in the first and last 60 bases of DNA fragments, while they do not help with internal complexities. The length of the stabilizing sequence can be variable, however it is important to ensure that the stabilizing sequence does not alter the reading frame if the DNA fragment will be used to express protein. Finally, if the DNA fragments are used to assemble a larger construct, the stabilizing sequence may interrupt the correct assembly. In this case, the terminal stabilizing sequence can be removed with restriction enzymes. Consider adding restriction sites that can be used throughout the cloning process and that do not leave scars in the final construct.

The length of the stabilizing sequence can be variable. However, if the DNA fragment will be used to express proteins, ensure that the stabilizing sequence does not alter the reading frame.

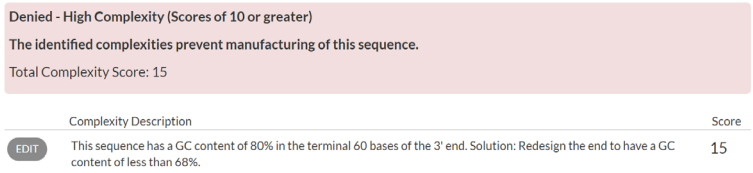
If the DNA fragments are used to assemble a larger construct, the stabilizing parts might interrupt the correct assembly. In this case, the terminal stabilizing sequences can be removed with restriction enzymes. Consider adding restriction sites that can be used throughout the cloning process and that do not leave scars in the final construct.

Here is an example of how stabilizing sequences can help reduce complexity scores:

Issue: GC content is too high in the 3′ end

Test Sequence 1:

GGCCTCGGTCACATTGGGGTACTCCTTGGCTCTTCCGCTTGGCCCGGTCTGTTGGGCCCCCGTAGCGTGAGTTTCGGCCCCGCGCTGCCCAGTGTCGGCCATTCTCATCGG
GGCCTCACTTCTGGATACCCCGACCTATTTTGACGGCACCATTGGCGGAAGTTGCTGGGCTTGTGCACCGTGGAGTCCTCCGATTACTCTTGTTGGTGTGGTATCGCTAAA
CTGCGCCGCGGAGCCTTATGGCATAGCCGTCCGCGGAGCATTCCGGTAGCGCTTATGGTCCATAGCACGTTCATCGCATCCGGGCGTGCGCTCTGTTTGACGACCCCTTGG
CGCAGAGGTGCTGGCCACGTGCTAAATTAAAGCGGCTGCACTGCTGTAAGGTCCGTCACGGAGACGGCGGCCCGGGGGGAGCACTAGCCCGCCGGCCCGTGCGGGAACACT
CTACATCGCTCCCGGACGGGCAGCGCTAGCCGTAGGTTCCACTGGTTCCGCCGTCGGCGCGGGTCCCGGCCGACGCGCGCCTCGTTGGGGGCCCGCAGT



Suggested modification: Add stabilizing sequence to 3′ end:

Test Sequence 1 (with stabilizing sequence):

GGCCTCGGTCACATTGGGGTACTCCTTGGCTCTTCCGCTTGGCCCGGTCTGTTGGGCCCCCGTAGCGTGAGTTTCGGCCCCGCGCTGCCCAGTGTCGGCCATTCTCATCGGG
GCCTCACTTCTGGATACCCCGACCTATTTTGACGGCACCATTGGCGGAAGTTGCTGGGCTTGTGCACCGTGGAGTCCTCCGATTACTCTTGTTGGTGTGGTATCGCTAAACT
GCGCCGCGGAGCCTTATGGCATAGCCGTCCGCGGAGCATTCCGGTAGCGCTTATGGTCCATAGCACGTTCATCGCATCCGGGCGTGCGCTCTGTTTGACGACCCCTTGGCGC
AGAGGTGCTGGCCACGTGCTAAATTAAAGCGGCTGCACTGCTGTAAGGTCCGTCACGGAGACGGCGGCCCGGGGGGAGCACTAGCCCGCCGGCCCGTGCGGGAACACTCTAC
ATCGCTCCCGGACGGGCAGCGCTAGCCGTAGGTTCCACTGGTTCCGCCGTCGGCGCGGGTCCCGGCCGACGCGCGCCTCGTTGGGGGCCCGCAGTTGCATGATCTACGTGCG
TCACATGCAGTACCACTAGCTCAGATTCAGTAGACCGCTGTTG

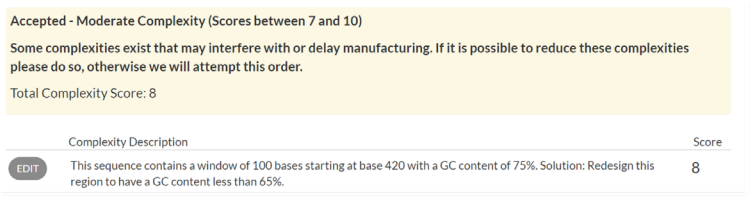


Figure 12. Example of reducing the complexity score of a gBlock Gene Fragment by adding a stabilizing. The original test sequence contained a region with high GC content in the 3′ end of the sequence. The complexity tool identified this complexity and denied the sequence. Once the stabilizing sequence (in red) was added to the sequence the complexity score was reduced to moderate. This sequence is acceptable for ordering.

Reduce complexity by splitting a sequence into sub-fragments

Complexities, especially in large DNA fragments with large repeats and secondary structures, are sometimes spaced apart by a longer sequence with no complexity issues. In this case, splitting up a sequence into shorter sub-fragments can separate and resolve the complexities. This is one way to preserve the characteristics of the sequence while reducing total complexities per sub-fragment. The sub-fragments can then be assembled using an assembly method such as seamless cloning or Golden Gate Assembly. The necessary elements for the assembly method of choice should be designed directly into the sub-fragments.

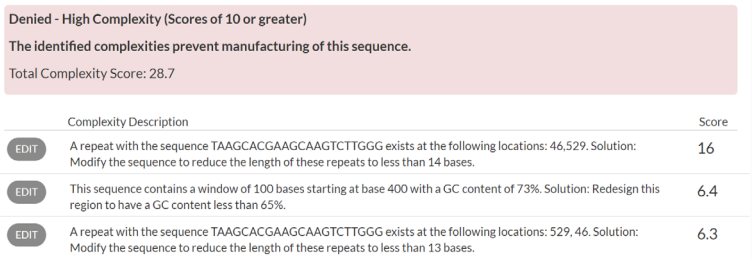
Creating sub-fragments is effective but has limitations and it might require several attempts to find the ideal region to split the sequence. Ideally, the sequence should not be split in the middle of a repeat, high/low GC region or a hairpin structure, as it can create new fragments with terminal ends that are problematic to manufacture.

Here is an example of how creating shorter sub-fragments can help reduce complexity scores:

Issue: Multiple complexity issues lead to high complexity score

Test Sequence 2:

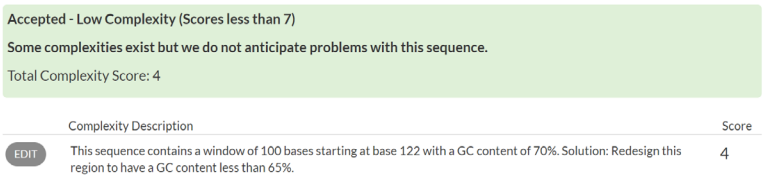
TCTCTGGCACGTGGGTGGCCTAGAGGAATCACATCCAAGGCTGGCTAAGCACGAAGCAAGTCTTGGGTGTAAGATTGTTGTCCCCTGTATACGGGATGGGGGTACTAGATGA
CCGCAGGGACTCCGACGTCAAGTACACCGCCCCGTCGTGGGCGCCGTTCGGGATCACGTTACCGCCAAAGGATGGGAGCAAGGCCCCTTCTCCCCCGCGGCCACGCCGGTAG
TGATCACGCCTTTGACCCTCCCGAGAGCCGGGAGGCGGAAATCCGCCACGAATGGGGAAGCATTTCCCCGACGGTCAATACGGGACGCTCCTAAGTTTTTCCACTCGCTTGA
GCCGGCTAGGCCTCTCTGCCCGAAGTTTCGACGGACTGCTGCCAACACCCAGGCATAGTTTTAGGGGGGTCACTCGGGGGCACCCGCAGCCAACTTGTCGGGTCCTGCCCGG
CTGGTCCTCGGGCTAATGCAGCGAATTGCCGAGGGCCCGGCCCCGCGCAACGGAACGTCTTTAGCTCCGGCAGGCAATTATAAGCACGAAGCAAGTCTTGGGAGGACAACGC
AAGCATGGCGCATATAAACAGAGAAACGGGCGAATGGACCTGTTCGTGCCGTATCGGAGAACAGCCTCGCGGAGGCATGTGCCATGCTGGCGCGCGGGGCACTCTGGTTACG
CATATGGTCCACAGGACACTCGTCGCTTCCGGGCTTGCCCTCTATGTGGC



Suggested modification: Split sequence into two fragments

Test Sequence 2 (subfragment 1):

TCTCTGGCACGTGGGTGGCCTAGAGGAATCACATCCAAGGCTGGCTAAGCACGAAGCAAGTCTTGGGTGTAAGATTGTTGTCCCCTGTATACGGGATGGGGGTACTAGATGA
CCGCAGGGACTCCGACGTCAAGTACACCGCCCCGTCGTGGGCGCCGTTCGGGATCACGTTACCGCCAAAGGATGGGAGCAAGGCCCCTTCTCCCCCGCGGCCACGCCGGTAG
TGATCACGCCTTTGACCCTCCCGAGAGCCGGGAGGCGGAAATCCGCCACGAATGGGGAAGCATTTCCCCGACGGTCAATACGGGACGCTCCTAAGTTTTTCCACTCGCTTGA
GCCGGCTAGGCCTCTCTGCCCGAAGTTTCGACGGACTGCTGCCAACAC



Test Sequence 2 (subfragment 2):

TCTCTGCCCGAAGTTTCGACGGACTGCTGCCAACACCCAGGCATAGTTTTAGGGGGGTCACTCGGGGGCACCCGCAGCCAACTTGTCGGGTCCTGCCCGGCTGGTCCTCGG
GCTAATGCAGCGAATTGCCGAGGGCCCGGCCCCGCGCAACGGAACGTCTTTAGCTCCGGCAGGCAATTATAAGCACGAAGCAAGTCTTGGGAGGACAACGCAAGCATGGCG
CATATAAACAGAGAAACGGGCGAATGGACCTGTTCGTGCCGTATCGGAGAACAGCCTCGCGGAGGCATGTGCCATGCTGGCGCGCGGGGCACTCTGGTTACGCATATGGTC
CACAGGACACTCGTCGCTTCCGGGCTTGCCCTCTATGTGGC

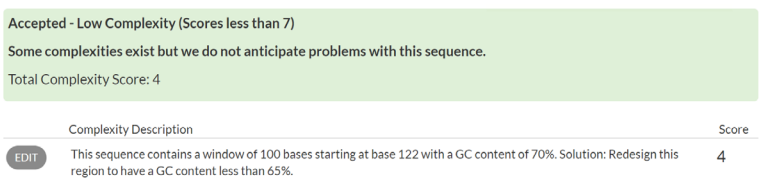


Figure 13. Example of reducing the complexity score of a gBlock Gene Fragment by splitting the sequence into sub-fragments. The original test sequence contains a sequence repeat which is difficult to manufacture. The complexity tool identified this complexity and denied the sequence. The sequence was split into two sub-fragments, which splits the repeating sequence. The sub-fragments were designed with homology to each other (in red) for reassembling. The resulting sub-fragments pass complexity screening and are acceptable for ordering.

Reduce complexity through codon optimization

When a DNA sequence is being used to create a protein product, high complexity DNA sequences can be modified using alternative codons. While the codon optimization approach reduces complexities within a DNA sequence, it also preserves the original amino acid sequence. At the same time, this approach can also be used to increase the chance of protein expression in a non-native host species.

The IDT Codon Optimization Tool can help to simplify the redesign. Users can input a DNA or protein sequence and select a host organism/cell type in which the gene(s) will be expressed. The tool provides codon usage tables as frequencies for a variety of species and optimizes the user's sequence for the chosen model. Both manual and automated optimization options are available, which provides flexibility in revising sequences to preserve experimental specifications. Manual optimization allows the selection of individual codons from a list, arranged in decreasing frequency of codon usage for that species. In comparison, automated optimization improves sequences based on optimal codon frequencies as well as overall structure and sequence complexities. Lastly, the tool identifies restriction enzyme sites present in the final sequence.

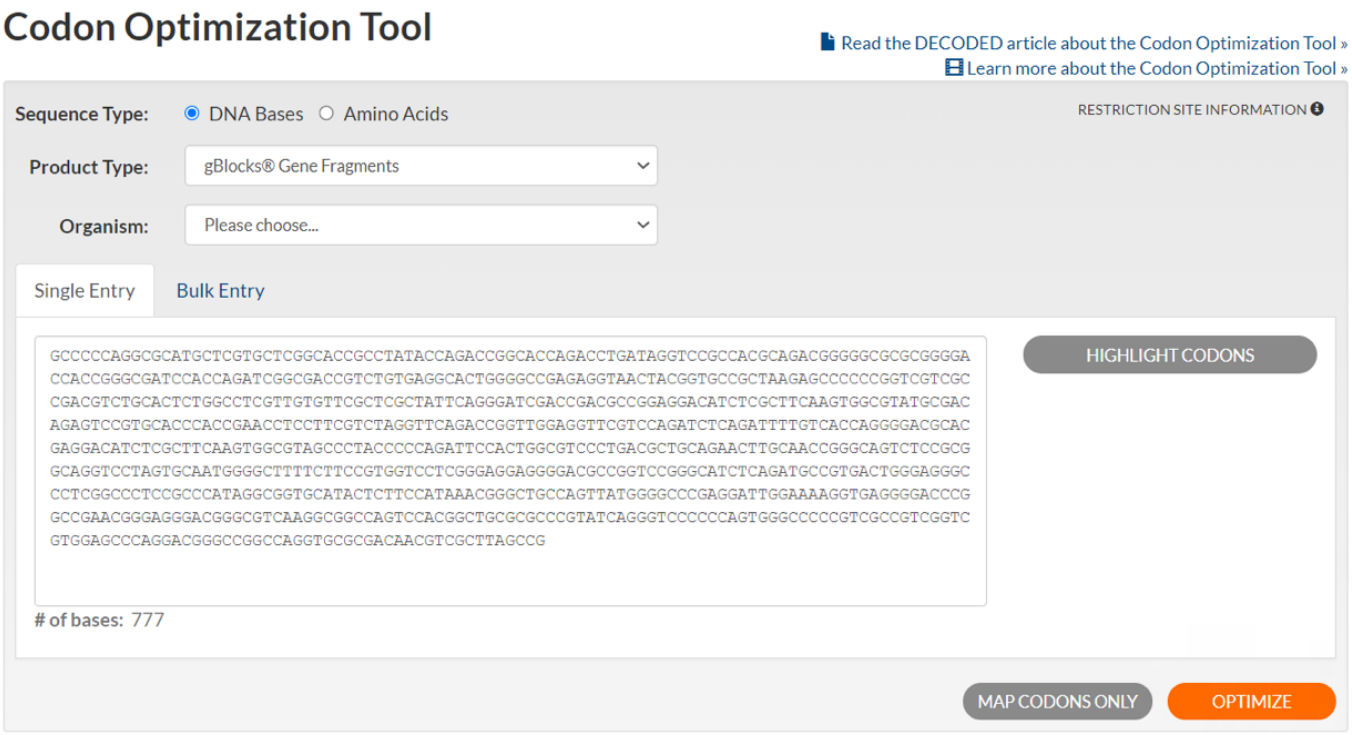


Figure 14. The IDT Codon Optimization Tool. This tool can be used to reduce complexity by using alternative codons for the amino acid. When using the automated optimization function the tool optimizes the sequence for reduced complexity while avoiding rare codon usage. For more information on how to use this tool, watch this video.

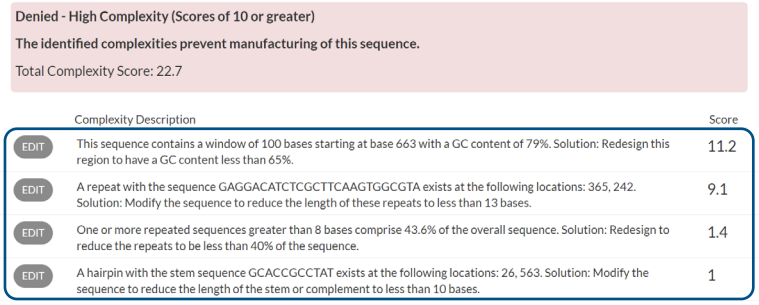
Codon optimization modifies the original DNA sequence; hence, the technique cannot be used when the base sequence needs to be preserved. For example, a promoter sequence with several repeats or high GC stretches cannot be codon optimized. Modifying a promoter sequence might disrupt the binding and assembly of proteins that could further impact their regulatory function. On the other hand, protein coding sequences that are not involved in DNA binding can be codon optimized.

Here is an example of how codon optimization can help reduce complexity scores:

Issue: High complexity sequence due to high GC content and repeat regions

Test Sequence 3:

GCCCCCAGGCGCATGCTCGTGCTCGGCACCGCCTATACCAGACCGGCACCAGACCTGATAGGTCCGCCACGCAGACGGGGGCGCGCGGGGACCACCGGGCGATCCACCAGAT
CGGCGACCGTCTGTGAGGCACTGGGGCCGAGAGGTAACTACGGTGCCGCTAAGAGCCCCCCGGTCGTCGCCGACGTCTGCACTCTGGCCTCGTTGTGTTCGCTCGCTATTCA
GGGATCGACCGACGCCGGAGGACATCTCGCTTCAAGTGGCGTATGCGACAGAGTCCGTGCACCCACCGAACCTCCTTCGTCTAGGTTCAGACCGGTTGGAGGTTCGTCCAGA
TCTCAGATTTTGTCACCAGGGGACGCACGAGGACATCTCGCTTCAAGTGGCGTAGCCCTACCCCCAGATTCCACTGGCGTCCCTGACGCTGCAGAACTTGCAACCGGGCAGT
CTCCGCGGCAGGTCCTAGTGCAATGGGGCTTTTCTTCCGTGGTCCTCGGGAGGAGGGGACGCCGGTCCGGGCATCTCAGATGCCGTGACTGGGAGGGCCCTCGGCCCTCCGC
CCATAGGCGGTGCATACTCTTCCATAAACGGGCTGCCAGTTATGGGGCCCGAGGATTGGAAAAGGTGAGGGGACCCGGCCGAACGGGAGGGACGGGCGTCAAGGCGGCCAGT
CCACGGCTGCGCGCCCGTATCAGGGTCCCCCCAGTGGGCCCCCGTCGCCGTCGGTCGTGGAGCCCAGGACGGGCCGGCCAGGTGCGCGACAACGTCGCTTAGCCG



Suggested modification: Codon optimize sequence using the organism that will produce the protein:

Test Sequence 3 (codon optimized):

GCCCCTCGCCGCATGCTTGTGTTGGGTACCGCCTATACCAGACCGGCACCAGACCTGATAGGTCCGCCACGCCGTCGGGGTCGCGCGGGGACCACCGGGCGTTCTACCAGAT
CGGCGACCGTCTGTGAGGCACTGGGTCCGAGAGGTAACTACGGTGCCGCTAAGAGCCCCCCGGTTGTCGCCGACGTCTGCACTCTGGCCTCGTTGTGTTCTCTCGCTATTCA
GGGATCGACCGACGCCGGAGGTCATCTCGCTTCAAGTGGCGTTTGCGACAGAGTCCGTGCACCCACCGAACCTCCTTCGTCTAGATTTAGACCGGTTGGAGGTTCATCCAGA
TCTCAGATTTTGTCACCAGGGGACGCAAGAGGACATCTTGCTTCAAGTGGCGTAGCCTTACCCCCAGATTCCACTGGCGTCCCTGACGCTGCAGAACTTGCAACCGGGCAGT
CTCCGCGGCAGGTCTTAGTGCAATGGGGCTTTTCTTCCGTGGTCCTCGGGCGCAGAGGACGCCGGTCCGGGCATCTCAGATGCCGTGACTGGGAGGGCCCTCGGCCCTCCGC
CCATCGTCGGTGCATACTCTTCCATAAACGGGCTGCCAGTTATGGGGCCCGTGGATTGGAAAAGGTGCGTGGACCTGGCCGTACAGGAGGGACTGGCGTCAAGGCAGCTTCT
CCACGTCTGCGTGCACGTATCCGTGTCCCCCCAGTGGGCCCCCGTCGCCGTCGGTCGTGGAGCCCAGGACGTGCAGGCCAGGTTCGCGATAATGTTGCATGACCG

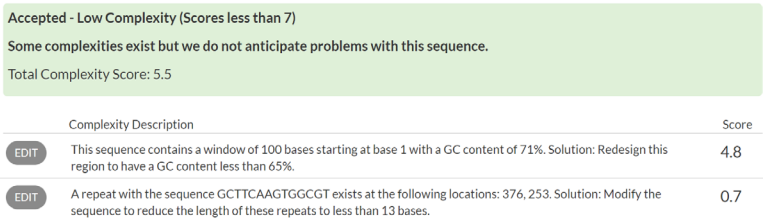


Figure 15. Example of reducing the complexity score of a gBlock Gene Fragment by codon optimization. The original test sequence contained multiple complexities throughout the sequence, including GC content extremes and repetitive sequences throughout. The complexity tool identified these complexities (highlighted by the blue box) and denied the sequence. Codon optimization of this sequence reduced the complexities so that it can be ordered while maintaining the same amino acid sequence.

Table 5. Summary of strategies to reduce complexities of DNA sequences.

Adding stabilizing sequences	Splitting into sub-fragments	Codon optimization
• Add extra bases to fragment terminal ends	• Create sub-fragments of original sequence	• Modify DNA sequence while preserving amino acid sequence
• Most effective when complexities are at terminal ends	• Most effective when internal complexities are spaced apart and can be separated	• Most effective for complexities spread across the entirety of the sequence
• Stabilizing bases may need to be removed	• Sub-fragments need to be reassembled	• Resulting optimized DNA sequence will be different from the original sequence

APPENDIX A: TROUBLESHOOTING

Issue	Possible cause	Solution
My sequence did not pass the complexity screening.	Sequence has one or more regions with complexities that interfere with successful synthesis: GC extremities, secondary structures, runs of identical bases (homopolymers), and repeats	To reduce complexities, use one of the following: a. Add stabilizing sequences b. Split up the sequence into shorter fragments c. Use codon optimization
My sequence passed complexity screening but failed synthesis.	Sequences with moderate and low complexity scores are accepted for synthesis, yet they can fail to synthesize. Since the screening tool estimates manufacturing success based on previous data, it cannot certainly evaluate whether sequence synthesis will fail.	To reduce complexities, use one of the following: a. Add stabilizing sequences b. Split up the sequence into shorter fragments c. Use codon optimization
I added restriction sites to my DNA fragment, but the enzymes did not cut correctly and generated unexpected band sizes.	a. The enzyme recognition sites might be placed at the terminal ends of the fragment.	a. Without terminal flanking ends, the restriction enzyme might not be able to efficiently bind the DNA fragment and cut it. Ensure that there are 4–6 flanking bases for the enzyme to bind.
	b. The restriction enzymes used are methylation sensitive.	b. If the DNA sequence is methylated, restriction enzymes that are methylation sensitive cannot cut it. Check that the sequence is not methylated. Consider using a dcm-/dam- competent *E. coli* cell line.
	c. The digest and/or ligation conditions are not optimal.	c. Check the protocol for optimal experimental conditions (e.g., incubation times, temperatures, buffers, etc.). If two different enzymes are used, try one enzyme at a time.
My DNA fragment did not insert into the vector resulting in colonies with empty vectors.	a. Vector re-ligation	a. To avoid vector re-ligation, use calf intestinal alkaline phosphatase to dephosphorylate the ends at the cloning site.
	b. Insert to vector ratio not optimal	b. Refer to the cloning kit for optimal insert to vector ratio. For larger constructs, try increasing the insert to vector ratio.
	c. Concentrations of the used DNA fragments are inaccurate.	c. Recheck the concentrations of the fragments to ensure that sufficient DNA was used.

Issue	Possible cause	Solution
I got very few (or no colonies) after transformation.	a. Low or inefficient transformation	a. Use a positive control to rule out any issues with competent cells and antibiotic selection.
	b. Gene product toxic to bacterial cells	b. If the gene encoded in the fragment is toxic to the host cells, try a different competent cell line or a low copy plasmid. A low copy plasmid will undergo fewer replication cycles, generating lesser protein and lower toxicity in the process.
I see truncations and/or deletions in my constructs.	Byproducts of fragment synthesis cloned	Preparations of DNA fragments may contain shorter truncated fragments that get preferably inserted over full-length constructs. Screen more colonies to find the correct full-length product. See Table 3 for an estimate of the number of colonies to screen.
I see point mutations in my DNA sequence.	a. eBlocks and gBlocks Gene Fragments are not quality-controlled using sequencing methods (gBlocks HiFi Gene Fragments however are verified by next generation sequencing). Thus, it is possible that a point mutation is present in the eBlocks or gBlocks Gene Fragment.	a. Because eBlocks and gBlocks Gene Fragments are not clonally derived and not QC'ed by sequencing, the preparation may contain DNA pieces with point mutations. Use Table 3 to make sure you have screened enough colonies. If you are still having troubles finding colonies with the correct clone, contact genes@idtdna.com.
	b. PCR amplification prior to cloning can introduce mutations.	b. Avoid amplifying the DNA fragment prior to cloning. Instead, design enzyme recognition sites or sequence overlaps directly into the fragment. If amplification is necessary, use a high-fidelity polymerase.
	c. If the gene product is toxic to bacterial cells, the sequence may spontaneously mutate to allow survival of the cells.	c. Check if the gene encoded in the DNA sequence is toxic to the competent cells. If so, use a different competent cell line.

APPENDIX B: SYNTHETIC DNA FRAGMENTS—FAQS

Design and optimization

1. **What can I do if my sequence is too complex to manufacture and I am unable to modify it?**

   If the DNA sequence contains a protein coding region, try to use the codon optimization approach to reduce the overall sequence complexities. Codon optimization uses alternative codons to retain the final amino acid sequence, while also generating a sequence that is feasible to manufacture.

   If the DNA sequence does not contain a protein coding region or if codon optimization is not an option, IDT might be able to synthesize the sequence as a custom gene instead. Unlike DNA fragments, custom genes are delivered cloned in a pUC-derived vector, which improves synthesis flexibility around high-complexity regions.

2. **Which cloning method should I use to generate a DNA construct?**

   This guide covers several cloning methods based on different types of DNA fragments. Each fragment can be designed according to the requirements of the cloning method. Review the experimental protocols and method requirements before selecting the optimal approach.

3. **Does codon optimization enhance protein expression?**

   The DNA sequence generated after codon optimization contains a minimum of rare codons (defined as those used less than 10% of the time) for that species. The DNA sequence generated after codon optimization excludes rare codons should also improve protein expression. Yet, it is impossible to predict how the expression level of the altered DNA sequence compares to the expression level of the original sequence. Expression mechanisms, especially in eukaryotic organisms, are not fully understood and cannot be accurately predicted *in silico*. If maximum protein expression is desired, design 2 or 3 different sequences and test them in the lab.

4. **Why do I get a different DNA sequence each time I optimize codons using the automated option?**

   The IDT Codon Optimization Tool redesigns sequences to minimize complexities and facilitate DNA synthesis. In this process, the tool uses alternative codons for amino acids based on usage frequencies specific to the host organism. The tool balances both usage frequencies and overall sequence complexities. Hence, it will not always select the codon with the highest frequency at every recurrence of an amino acid as that might result in greater number of repeats and secondary structures. As a result, there are many permutations to codon optimization. Every optimization attempt may generate a different sequence depending on surrounding sequence complexities.

5. **Can I order DNA fragments containing multiple regions with variable bases?**

   IDT gene fragments can be designed with a single region containing consecutive variable bases. These variable regions contain a population of multiple bases at each site as opposed a single base. More specifically, constructs with up to 18 consecutive N (A,C,G,T) or K (G,T) bases can be ordered, with a minimum of 125 base pairs flanking sequence on either side of the variable bases. The maximum length for the entire sequence is 500 base pairs including the variable bases.

6. **Why do you allow only 18 consecutive N or K bases in DNA fragments?**

   Incorporating 6 NNK codons corresponds to about 1 billion possible combinations, and 18 N mixed bases will create a pool with 68.7 billion different DNA fragments. As the number of variable bases in the fragment increases, the number of molecules representing a particular sequence decreases. By limiting variable regions to 18 mixed bases, the overall pool of library constructs will have representation of each variation. Since most functional screens cover 1000–100,000 recombinant colonies, allowing 18 mixed bases should be adequate for producing useful results.

Ordering and using synthetic DNA fragments

1. **How do I resuspend DNA fragments once I have received them?**

   Standard gBlocks Gene Fragments and gBlocks HiFi Gene Fragments are dried down and shipped in tubes (unless otherwise requested to ship in plates or ship wet). Upon receipt, spin down the tube in a microcentrifuge for 3–5 seconds to ensure that the DNA pellet is at the bottom. Since the pellet can become statically charged or lodge in the cap during shipping, this step prevents the pellet from flying out of the tube or remaining in the cap and losing yield. After centrifugation, resuspend the pellet in molecular-grade water or TE buffer (pH 7.5–8) at the required concentration. Verify the concentration of the resuspended gBlocks Gene Fragments to ensure the full product has gone into solution. Generate aliquots of the stock to avoid more than 2 or 3 freeze-thaw cycles.

   Plate products, including eBlocks and gBlocks Gene Fragments, are shipped in on dry ice and resuspended in 20 µL of nuclease-free water at a 10 ng/µL concentration.

2. **Do I need to amplify the DNA fragments before cloning or digesting with restriction enzymes?**

   Because DNA fragments are synthesized from several smaller overlapping fragments, the final sample contains smaller byproducts. During PCR, these byproducts might be preferentially amplified, thereby outnumbering the correct full-length products. The issue is prevalent when amplifying non-clonal gBlocks Gene Fragments greater than 1 kb in length. Therefore, amplifying gBlocks Gene Fragments greater than 1 kb is not recommended.

   Most applications such as restriction digestion, blunt-end cloning, and seamless cloning can be performed directly with the yield supplied by IDT, rendering PCR amplification unnecessary. If the amplification means to add flanking bases to complete a cloning reaction, we strongly recommend ordering the fragments with the required flanking bases.

   If amplification of the custom DNA fragment is essential, use the minimum amount of starting template required by the PCR kit protocol, limit the cycle number to 10–12, and use a high-fidelity polymerase.

3. **Why are there errors in my DNA fragment?**

   gBlocks and eBlocks Gene Fragments are not clonally derived and not verified by sequencing, the sample might contain a few species with mutations (indels and substitutions). Most species in the sample, however, should have the correct sequence. Thus, by screening more colonies, the correct clone without mutations should be found.

   If the DNA fragment is amplified prior to cloning, amplification can also introduce point mutations. Depending on how early in the amplification process the mutation was introduced, the dominant species in the sample may or may not contain the correct sequence. Avoid amplifying the DNA fragment, but if you must, use a high-fidelity polymerase to minimize errors.

   In rare cases, the gene product of the DNA sequence can be toxic to the bacterial host cells. Often spontaneous mutations are introduced to promote survival. Ensure that the gene is not toxic to the host organism.

4. **Why am I seeing large deletions in my sequences?**

   Sequences with secondary structure motifs forming hairpins or stem-loops bend away from the DNA. During transcription, this region might be skipped and thus get lost leading to internal, large deletions. To avoid this, ensure that secondary structures are minimal when submitting the sequence for synthesis.

5. **What are the quality control checks performed for DNA fragments?**

   IDT provides three types of DNA fragments (eBlocks Gene Fragments, gBlocks Gene Fragments, and gBlocks HiFi Gene Fragments), and the QC criteria are different for each type. The eBlocks Gene Fragments are assessed by capillary electrophoresis, while gBlocks Gene Fragments are assessed by either capillary electrophoresis or ESI mass spectrometry. The gBlocks HiFi Gene Fragments are sequence-verified by next generation sequencing. All three fragments are quantified by measuring their optical densities.

6. **How should I store my DNA fragments long-term?**

   DNA fragments that are shipped dry (standard gBlocks Gene Fragments or gBlocks HiFi Gene Fragments) can be stored dry at room temperature. Resuspended fragments, including eBlocks Gene Fragments, can be stored at 4°C for less than 1 month. For longer-term, aliquot and store at –20°C and minimize freeze-thaw cycles (no more than four).

   If storing the DNA fragments at a concentration less than 1 ng/µL, add a carrier (such as tRNA or poly A) to the resuspension and dilution buffers at 0.1–1 mg/mL. We have observed a decrease in DNA fragment concentrations in solutions with a starting concentration of less than 1 ng/µL, even when stored in low-bind tubes. This may be related to the very high purity of DNA fragments. Synthetic DNA lacks normal cellular debris found in DNA isolated from natural sources. In the absence of these natural carriers, synthetic DNA will bind irreversibly to plastic, resulting in DNA loss. Hence, small amounts of carrier can prevent this loss.

REFERENCES

1. Shuman S. **DNA ligases: progress and prospects.** J Biol Chem. 2009;284(26):17365-17369. doi:10.1074/jbc.R900017200
2. Meselson M, Yuan R. **DNA restriction enzyme from E. coli.** Nature. 1968;217(5134):1110-1114. doi:10.1038/2171110a0
3. Mandel M, Higa A. **Calcium-dependent bacteriophage DNA infection**. J Mol Biol. 1970;53(1):159-162. doi:10.1016/0022-2836(70)90051-3
4. Historical Profile Paul Berg. Science History Institute. [Accessed 21 January 2022] **https://www.sciencehistory.org/historical-profile/paul-berg**
5. Cohen SN, Chang AC, Boyer HW, Helling RB. **Construction of biologically functional bacterial plasmids in vitro.** Proc Natl Acad Sci U S A. 1973;70(11):3240-3244. doi:10.1073/pnas.70.11.3240
6. Chial Heidi. Restriction Enzymes. Scitable. [Accessed 21 January 2022] **https://www.nature.com/scitable/spotlight/restriction-enzymes-18458113/**
7. Saiki R, Scharf S, Faloona F, et al. **Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of Sickle Cell Anemia. Science.** 1985;230(4732): 1350-1354. DOI: 10.1126/science.2999980
8. Human Genome Project FAQ. National Human Genome Research Institute. [Accessed 21 January 2022]. **https://www.genome.gov/human-genome-project/Completion-FAQ**
9. Li MZ, Elledge SJ. **Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC.** Nat Methods. 2007;4(3):251-256. doi:10.1038/nmeth1010
10. Engler C, Kandzia R, Marillonnet S. **A one pot, one step, precision cloning method with high throughput capability.** PLoS One. 2008;3(11): e3647. doi: 10.1371/journal.pone.0003647
11. Gibson D, Young L, Chuang RY, et al. **Enzymatic assembly of DNA molecules up to several hundred kilobases.** Nat Methods 2009; 6(5):343–345. **https://doi.org/10.1038/nmeth.1318**
12. David P. Clark, Nanette J. Pazdernik, Michelle R. McGehee; **Chapter 7, Cloning Genes for Synthetic Biology**; Molecular Biology Third Edition, 2019. 199-239.
13. Thermo Fisher Scientific. (2003) **Gateway® Technology.** [Online] Waltham, MA, Thermo Fisher Scientific, Inc. [Accessed 21 January 2022]
14. Cloning learning center. Takara. [Accessed 21 January 2022]. **https://www.takarabio.com/learning-centers/cloning**

IDT PROTOCOLS

Blunt-end cloning protocol

TA cloning method protocol

DNA cloning guide

Technical support: genes@idtdna.com

Doc ID: RUO21-0183_001 05/22