TWIST TIPS

HOW TO DESIGN YOUR GENE

Twist Bioscience uses a silicon-based DNA synthesis platform to generate high-quality synthetic DNA. Designing and ordering genes from Twist Bioscience is simple when using our online ordering platform. There are two options available for synthetic DNA: Twist Gene Fragments (linear dsDNA) and Twist Clonal Genes (DNA cloned into a selection of either Twist Cloning and Expression Vectors, or a custom vector of your choice).

When a new gene sequence is uploaded to our online ordering platform, an algorithm instantly provides accurate feedback in real-time regarding the likelihood of successful synthesis. Prior to placing an order, any problematic regions within the uploaded sequence are highlighted, and suggestions for maximizing synthesis success are provided. The following guidelines offer additional tips for optimizing and customizing Twist Clonal Genes and Gene Fragments for various applications.

KEY CONSIDERATIONS FOR SUCCESSFUL SEQUENCE DESIGN

Keep your sequences above 300 bp

- If your gene sequence is shorter than the minimum length, a filler sequence can be added to the 3' or 5' end of your sequence. **Learn more about filler sequences here**.
- If you are adding a filler sequence to Clonal Genes, please work with the Twist Support Team or customersupport@twistbioscience.com who can help you with your design.

Follow sequence parameters

Sequence parameters, such as GC content, repeat length, and repeat density, can impact the success of synthesis. When multiple metrics fall just below their respective thresholds, their combined effect can decrease the probability of successful synthesis. To minimize failure rate and increase probability for success, follow the parameters below:

- Avoid repeats of ≥ 20 bp or Tm ≥ 60°C
- Global GC content between 25% and 65%
- Minimize the difference in GC content between the highest and lowest 50 bp window so that it does not exceed 50%
- Minimize homopolymers so that they do not exceed greater than 13 nt
- Try to avoid clusters of short repeats (8-9 bp) and avoid direct repeats longer than 12 - 16 bp
- For HIS tags use a combination of CAC and CAT codons i.e. CACCAT...

Include restriction enzyme sites as needed

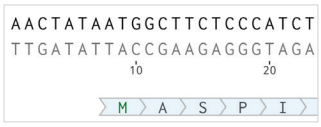
- Some Twist Vectors do not contain restriction sites flanking the custom gene sequence. If you require restriction sites, for example for downstream subcloning, make sure to add these to your insert sequence.

Double check the final sequence before ordering

- Gene Fragments with adaptors will contain Adaptor sequences outside of your designed sequence. To maximize success downstream, we recommend reviewing the following information here. If adapters are not wanted, please order your Gene Fragments without adapters.
- For Clonal Genes, use the Export Table button to download the full construct and review it carefully to confirm that the sequence is correctly inserted into the vector, and any open reading frames are accurate.

Keep proteins in frame

- If your synthetic DNA encodes for a protein, before placing your order, click on each gene in the listing to view its translation and confirm the resulting protein is in frame or download the draft or construct genbank maps via the download button to review your sequences offline.



Include a stop codon in protein sequences

- Check the annotations available in the construct view for each protein encoding gene to ensure each sequence includes a stop codon.

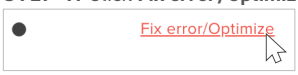
Notify Twist if an encoded protein inhibits the growth of *E. coli*

- If you are ordering cloned genes, the encoded proteins may affect cell viability and thus may impact the manufacture of your final construct.
- Please notify Twist if your protein contains elements known to be toxic (e.g. membrane proteins, DNA/RNA binding proteins, or bacterial toxins) to host *E. coli* cells by emailing the Twist Support Team or customersupport@twistbioscience.com.

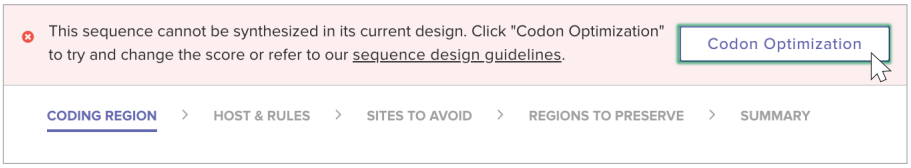
SCORING YOUR SEQUENCES

When a new gene sequence is uploaded into our online ordering platform, a proprietary algorithm measures the sequence length, GC content, repeats, homopolymers, and other potentially problematic regions, and automatically scores for feasibility of assembly. Sequences receiving a “Standard” score are not expected to experience any synthesis difficulty. “Complex” sequences may experience increased turnaround time or risk of manufacturing failure. “Not Accepted” sequences are likely to have a very low success rate. If a sequence is determined to be “Not Accepted”, we will not attempt to synthesize these sequences until the problematic regions are optimized or removed. Our codon optimization tool will suggest modifications that will help improve synthesis success rate, while keeping the translated sequence intact.

STEP 1: Click **Fix error/Optimize** link to access the option for codon optimization.



STEP 2: Click **Codon Optimization** to optimize sequences for DNA synthesis. Follow the recommendations to resolve all indicated problems with your order.



VECTOR DIVERSITY

When ordering Clonal Genes, choose from our catalog of cloning and expression vectors or onboard your own vector of choice. For more information on vector onboarding, click **here**.

PROTEIN TAGS

Twist Gene Fragments and Clonal Genes allowing you to design the exact sequences you need. Include the sequences below to immediately incorporate the following for downstream applications:

- Epitope tags for purification or detection of expressed proteins
- Tags for fluorescence-based detection of protein expression
- Linker sequences for adjusting flexibility/rigidity of proteins
- Signal peptides for targeting proteins to organelles
- Sequences for adjusting the solubility and stability of proteins

Commonly used tags for purification, visualization, or modification of expressed proteins

TAG	DESCRIPTION	SEQUENCE
EPITOPE TAGS FOR PURIFICATION OR DETECTION OF EXPRESSED PROTEINS		
6xHis-tag	Mixed CAC/CAT His-tag; easier to synthesize than the low-diversity 6xCAT His-tag	CACCATCATCACCACCAT
Myc tag	Derived from the c-myc protein	GAACAAAAACTCATCTCAGAAGAGGATCTG
HA tag	Derived from human influenza hemagglutinin	TACCCATACGATGTTCCAGATTACGCT
Flag tag	Artificial antigen, hydrophilic and does not tend to denature or inactive attached proteins	GACTACAAAGACGATGACGACAAG
V5	Derived from an epitope on the P and V proteins of the paramyxovirus of simian virus 5 (SV5)	GGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACG
VSV-G	Derived from the Vesicular Stomatitis viral glycoprotein	TATACAGACATAGAGATGAACCGACTTGGAAAG
Protein A	Derived from a surface antigen of *Staphylococcus aureus*	GTGGACAACAAATTCAACAAAGAACAACAAAACGCGTTCTATGAGATCTT ACATT TACCTAACTTAAACGAAGAACAACGAAACGCCTTCATCCAAAGTT TAAAAGATGACCCAAGCCAAAG
Cleavage tag TEV	Tobacco etch virus protease cleavage site, facilitates removal of epitope tags following purification	GAGGATCTGTACTTTCAGAGC

PROTEIN TAGS, continued

TAG	DESCRIPTION	SEQUENCE
TAGS FOR FLUORESCENCE-BASED DETECTION OF PROTEIN EXPRESSION		
EGFP	Enhanced green fluorescent protein	GTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGG ACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATG CCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGT GCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGC TACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCG CCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCA TCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAA CAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAAC TTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCA GCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGA GCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTG CTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG
mCherry	Red fluorescent protein	GTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAA GGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGC GAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGG CCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCC TACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGC TTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAG GACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAAC TTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCC GAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAA GCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAG AAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTC CCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACT CCACCGGCGGCATGGACGAGCTGTACAAG
LINKER SEQUENCES FOR ADJUSTING FLEXIBILITY / RIGIDITY OF PROTEINS		
$(GGGGS)_1$	Short, flexible linker	GGAGGCGGTGGATCT
$(GGGGS)_2$	Medium, flexible linker	GGAGGCGGTGGATCTGGCGGAGGTGGTAG
$(GGGS)_3$	Long, flexible linker	GGAGGCGGTGGATCTGGCGGAGGTGGTAGTGGGGGTGGTGGAAGC

PROTEIN TAGS, continued

TAG	DESCRIPTION	SEQUENCE
SIGNAL PEPTIDES FOR TARGETING PROTEINS TO ORGANELLES (in Saccharomyces cerevisiae)		
NLS	Nuclear localization sequence (targets proteins to nuclei in plant and animal systems)	CCAAAAAAGAAGAGAAAGGTAGAAGACCCC
HDEL	Targets proteins to the endoplasmic reticulum	CACGACGAATTG
SEQUENCES FOR ADJUSTING THE SOLUBILITY AND STABILITY OF PROTEINS		
MBP	Maltose-binding protein	ATGAAAATAAAAACAGGTGCACGCATCCTCGCATTATCCGCATTAACGACGATGATG TTTTCCGCCTCGGCTCTCGCCAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAAC GGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATAC CGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGG TTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGG TGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCA GGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATT GCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGC CGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGA AAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCG CTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATT AAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGAC CTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCC TTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACAT CGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAAC CATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAAC AAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAA GCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAG AGTTGGTGAAAGATCCGCGTATTGCCGCCACTATGGAAAACGCCCAGAAAGGTGAAA TCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGAT CAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT
GST	Glutathione-S-transferase	ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTT CTTTTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGT GATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCC TTATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATA GCTGACAAGCACAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAAT GCTTGAAGGAGCGGTTTTGGATATTAGATACGGTGTTTCGAGAATTGCATATAGT AAAGACTTTGAAACTCTCAAAGTTGATTTTCTTAGCAAGCTACCTGAAATGCTGA AAATGTTCGAAGATCGTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCC ATCCTGACTTCATGTTGTATGACGCTCTTGATGTTGTTTTATACATGGACCCAA TGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAAAAACGTATTGAAGCTA TCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCATGGCCTTTGCAG GGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATAA

For more help in optimizing the design of your gene sequences, please contact the Twist Support Team at customersupport@twistbioscience.com or through our online order system.

**Gateway® is a registered trademark of Thermo Fisher Scientific Inc.*

CONTACT US sales@twistbioscience.com or visit twistbioscience.com