

Gene Designer: Gene Design Automation Tool

DNA2.0 Inc.

Introduction

Gene Designer from DNA2.0 Inc. is tool for molecular biologist and synthetic biologist. This software enables the user to:

- Design large and small DNA fragments.
- Optimize expression in desired hosts using codon optimization.
- Build and Manipulate DNA from building blocks such as promoters and ORFs.

and much more..-

Support

We provide a detailed use case manual with the software and are happy to help you with any design. For support call us at +1 650 853 8347 or email us at support@dna20.com

Gene Synthesis

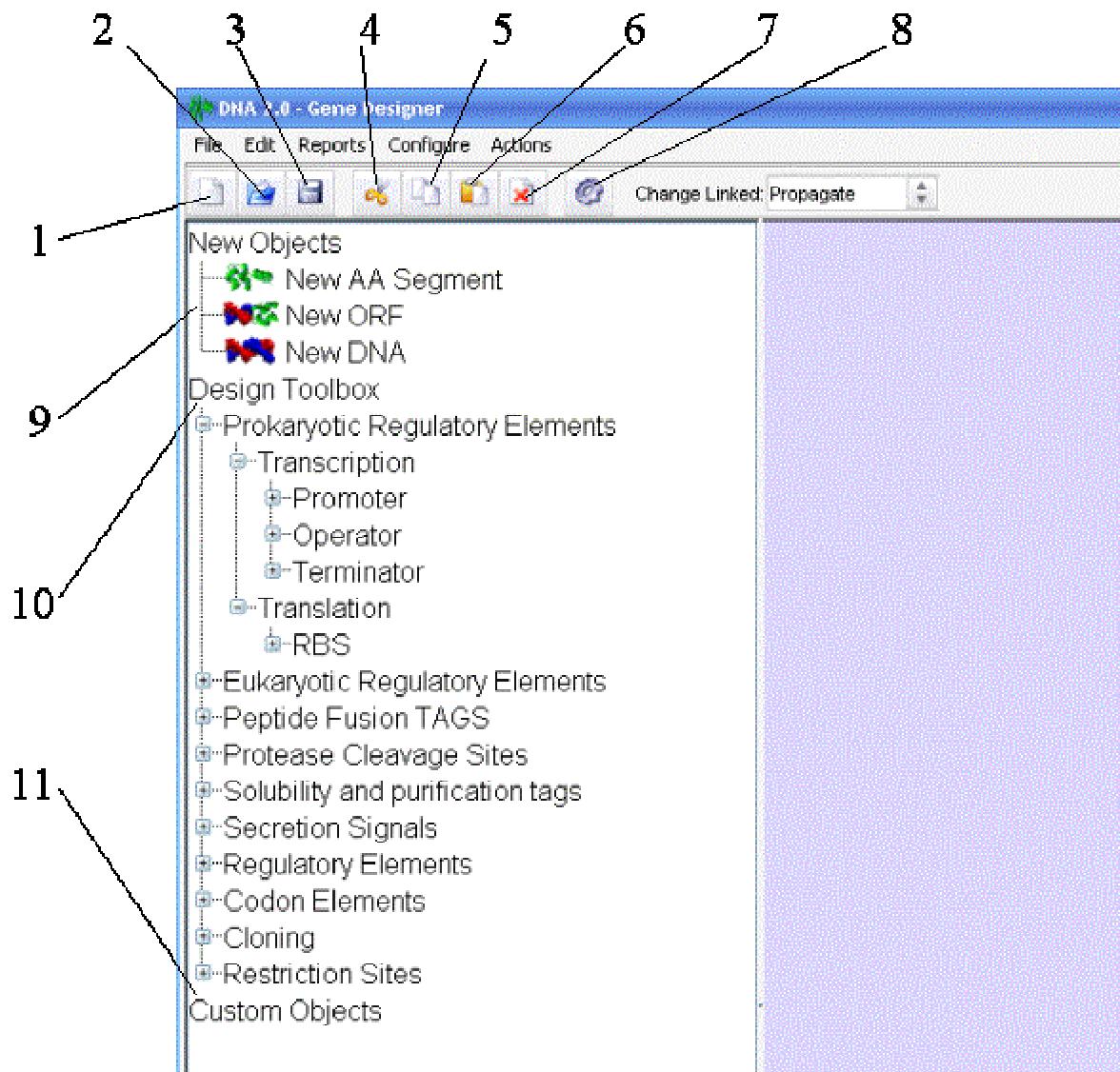
- DNA2.0 Custom Gene Synthesis: You can order your genes directly from Gene Designer.
- For synthesis of your designed genes DNA2.0 offers great rates and rapid turnaround services. For details visit <http://www.dnatwopointo.com/>

Plane Gene

- Gene Elements like expression ready fluorescent proteins from our Planet Gene Catalog are available in our design toolbox in Gene Designer.
- DNA2.0 also offers Planet Gene : The Online resource for codon optimized genes where you can obtain Human Genes for expression in your host, Fluorescent protein of your choice and Human Genes optimized to escape RNAi induced silencing for controls. Visit <http://www.planetgene.com> for details.

Getting Started

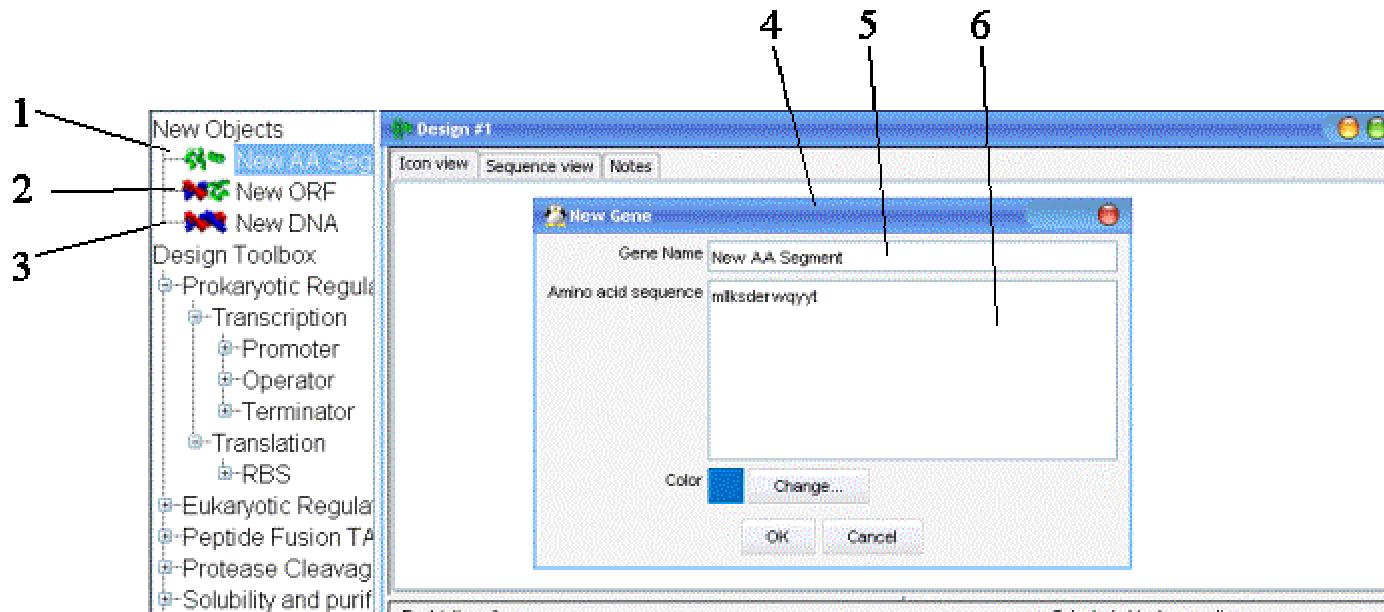
1. [Create a new project](#)
2. Open an existing project
3. Save project
4. Cut selected sequence
5. Copy selected sequence
6. Paste selected sequence
7. Delete selected sequence
8. [Backtranslate](#) amino acid sequence to corresponding DNA sequence
9. New sequence objects: [amino acids only](#), [amino acids plus DNA \(ORF\)](#), [DNA only](#)
10. Design toolbox of [pre-designed sequence elements](#)
11. Design toolbox of [custom designed sequence elements](#)



Creating Sequence Elements

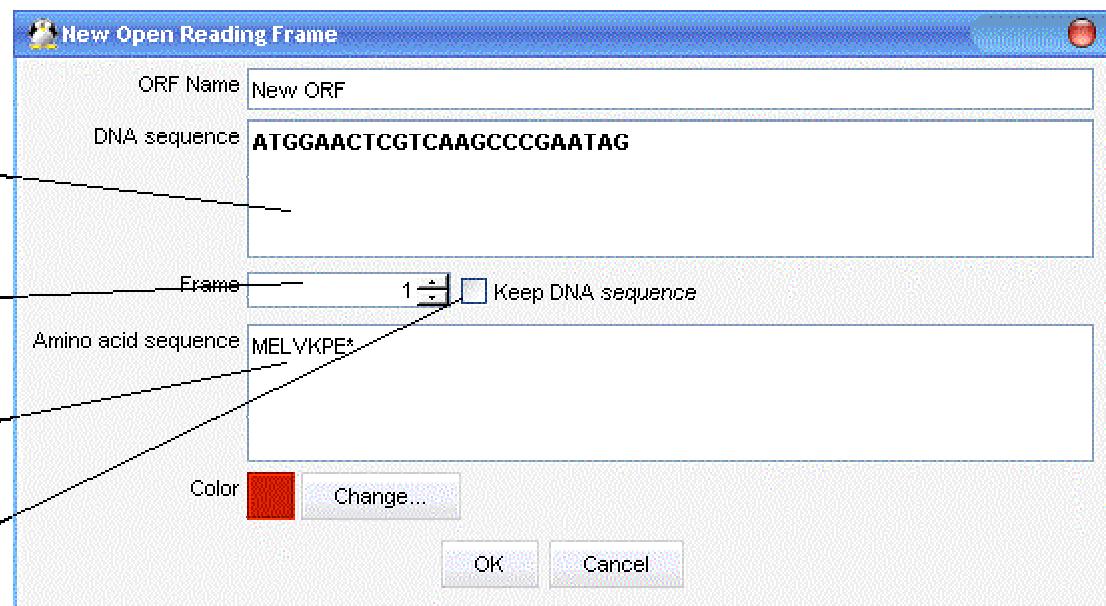
Amino Acid and DNA Sequence Elements

- Create or open a new project ([see overview](#)).
- Select a new object by clicking with the mouse and dragging into the open project window
- This creates a new sequence element window (4) with space for a name (5) and sequence (6). Name and sequence can be entered by typing or pasting.
- Amino acid sequence elements (1) accept one letter code for 20 amino acids. Non-standard characters are automatically filtered out.
- DNA sequence elements (3) accept 1 letter code for 4 bases. Non-standard characters are automatically filtered out.
- [see overviewORF \(open reading frame\) elements](#) (2) accept 1 letter code for 4 bases. Non-standard characters are automatically filtered out.



ORF Elements

ORF sequence elements are entered as DNA sequence (1). The frame can be selected (2) to define whether the first codon starts at the first, second or third base. The amino acid sequence is translated as the DNA sequence is entered (3). The DNA sequence entered can be retained (4) or optimized by subsequent back-translation.



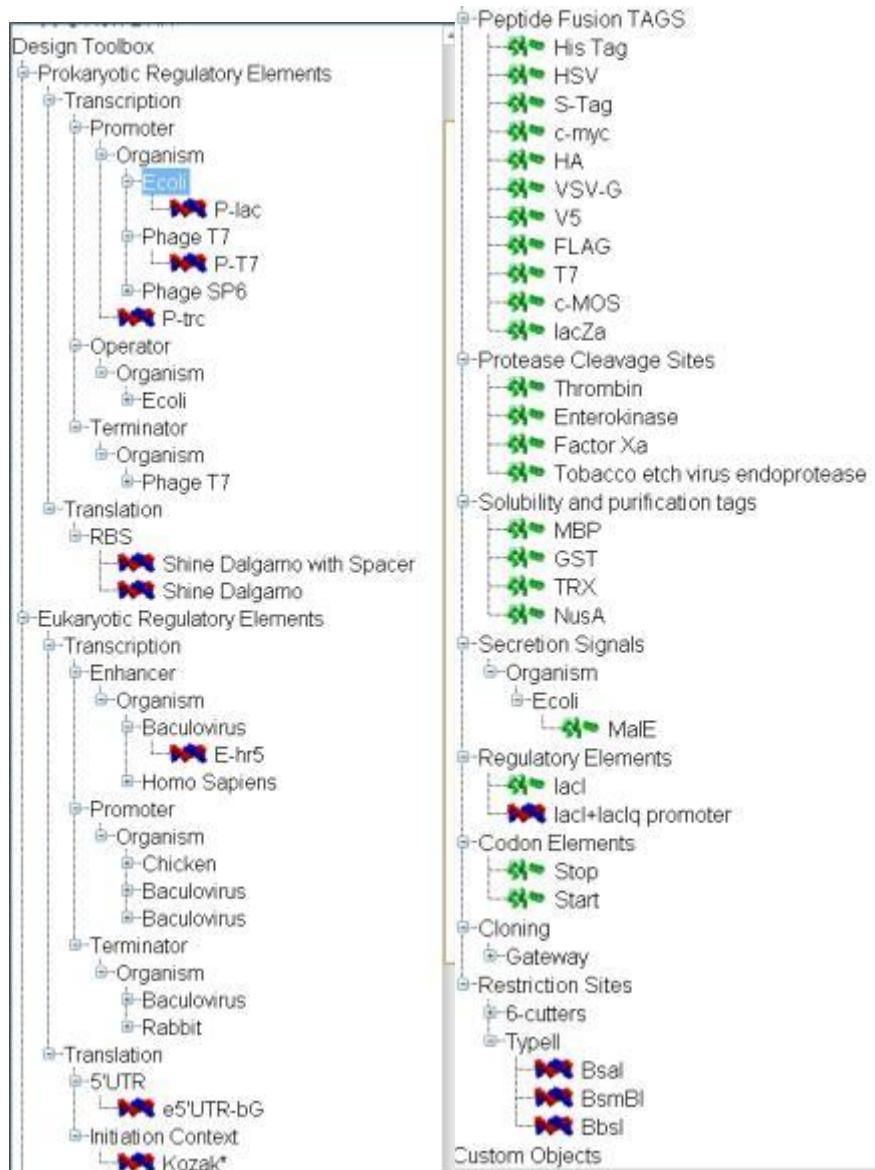
Library Sequence Elements

Pre-designed sequence elements can be organized in a hierarchical tree. For example

- Elements can be divided into regulatory, expressed and cloning elements.
 - Regulatory elements can be divided, for example between prokaryotic and eukaryotic elements.
 - Prokaryotic elements can be further subdivided, for example into transcriptional and translational elements.
 - Transcriptional elements can be further divided, for example into enhancers, promoters, operators, terminators, polyadenylation signals. These elements can be further subdivided according to organism of origin.
 - Translational elements can be further divided, for example into 5' and 3' untranslated regions, ribosome binding sites, initiation AUG contexts, termination codons.
 - Expressed elements can be further subdivided, for example into peptide fusion tags, protease cleavage sites, solubility and fusion tags, secretion signals. These elements can be further subdivided according to organism of origin.
 - Cloning elements can be further subdivided, for example into recombinase recognitions sequences, restriction enzyme recognition sequences.

It is also possible to create custom objects by dragging the icon for a created element onto the library tree into the Custom Objects folder.

It is also possible to rearrange the hierarchy by dragging objects from one position to another.



Custom Sequence Elements

Drag and drop any sequence element created by the user into Custom Toolbox in the Tree. this will save the object for use in future projects.

Differences Between Sequence Elements

- Amino acid elements are subject to subsequent back-translation.
- DNA elements are fixed and will not be altered by backtranslation.
- ORF sequences are entered as DNA sequences. They are automatically translated. The DNA sequence may be fixed at the time of entry, in which case it is not affected by subsequent back-translation. If the DNA sequence is not fixed, the amino acid sequence will be backtranslated at the backtranslation step.

Positioning Sequence Elements: Icon View

In the icon view each sequence element can be represented by an icon. The order of the icons can be altered by clicking on an element with the mouse and dragging it to a new position. Each sequence element can be identified by a name and by a color selected by the user.



Sequence View

Viewing and Modifying Sequence Elements

In the sequence view, the DNA (1, 2) and / or amino acid sequence (3, 4) of each sequence element is shown. Each element is named above the sequence, together with the positions of its start and finish within the entire sequence. The DNA sequence of the entire project is shown above the individual elements (5). For amino acid elements that have not yet been back-translated this sequence is blank. The positions of the start and end of each sequence element is indicated, together with a marker every 20 bases (6).

For DNA elements, the amino acid sequence resulting from translation of that sequence is shown in all 6 reading frames (7). This allows rapid assessment of reading frames when combining amino acid and DNA elements.

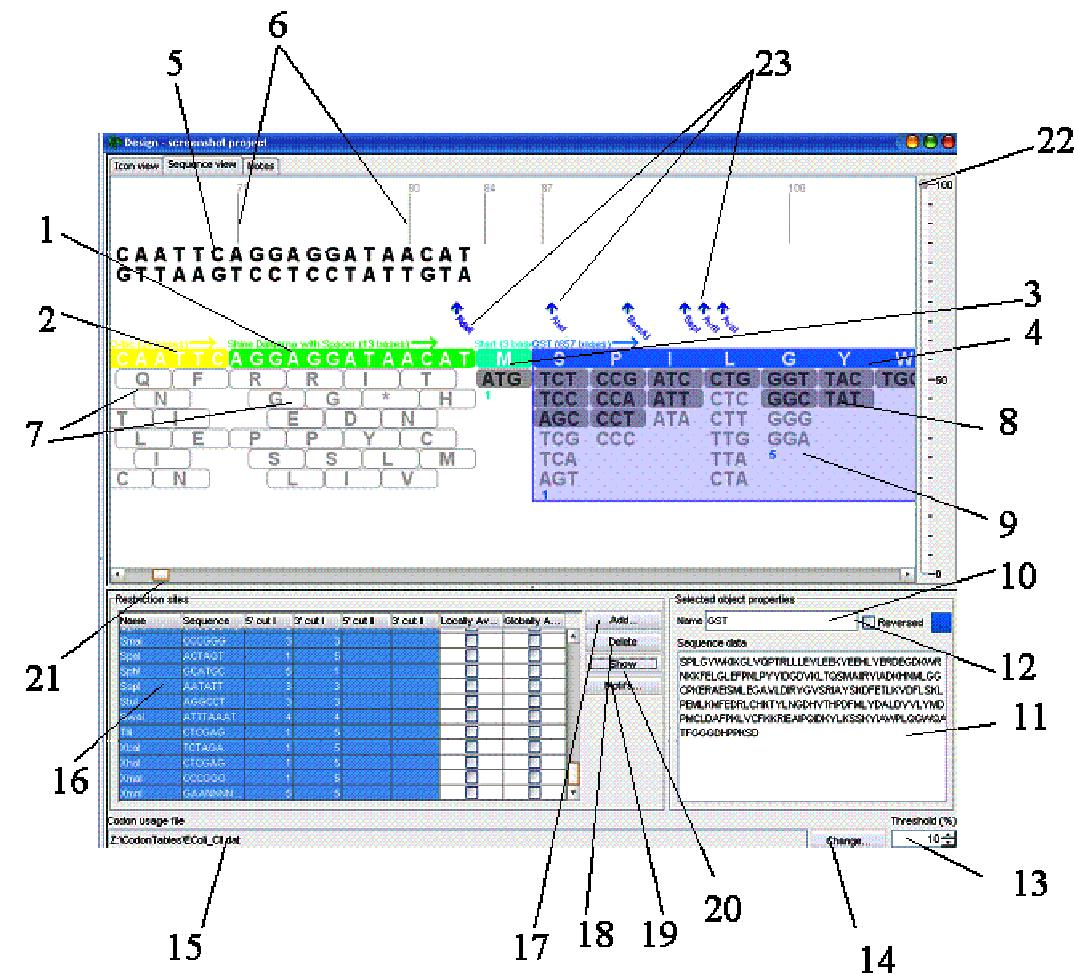
Advanced sequence element manipulations including merging, splitting and reverse complementing of elements are described [here](#).

Codon Bias Table

For amino acid elements, the possible codons encoding each amino acid are shown (8, 9). The codons for each amino acid are ranked according to use in the expression organism selected.

The codon bias table used is indicated (15) and can be changed (14) by choosing from list. New codon tables can also be imported in GCG format available from [KEGG](#) using the menu Conigure->Configure Codon Maps

The threshold (13) can be set to exclude codons that are used below a certain frequency in the selected organism. The codons for each amino acid are also color-coded, with those used at a frequency in the selected organism above the selected threshold shown in one color (8), and those that occur at a frequency in the selected organism below the selected threshold in another (9).

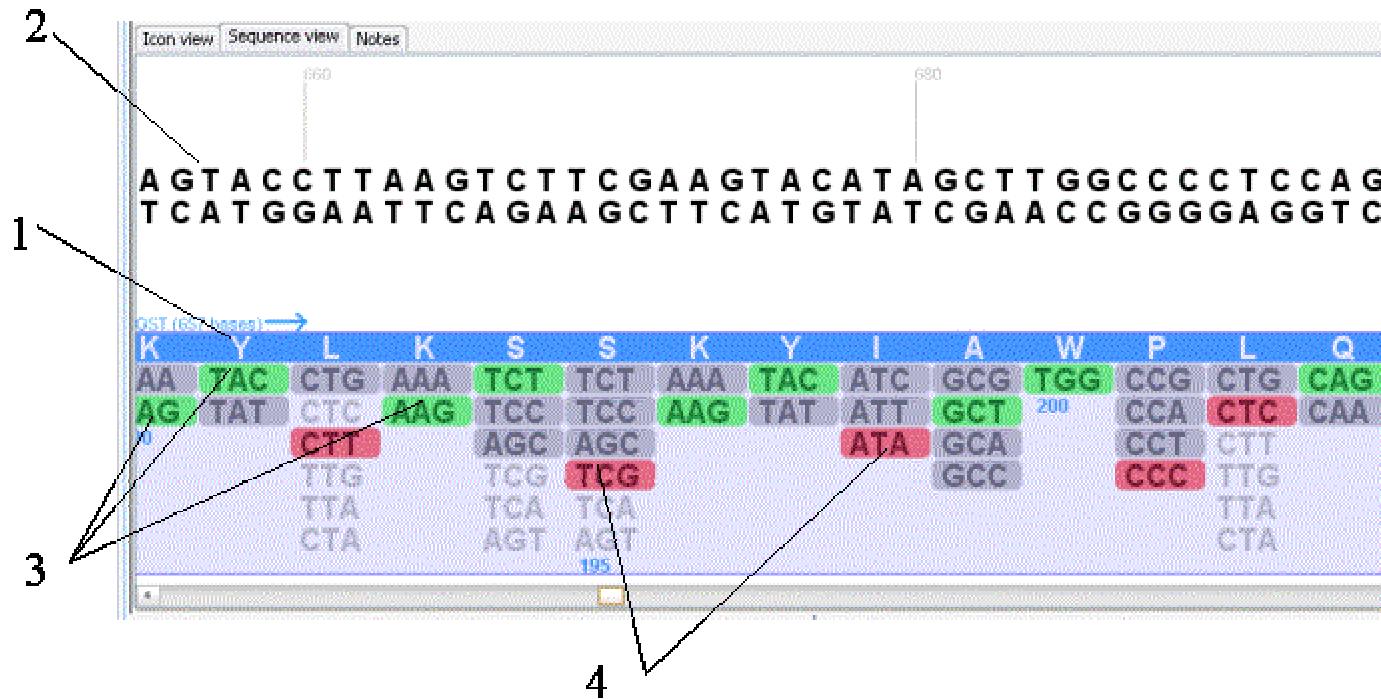


A sequence element can be selected (4) which displays its characteristics in an editing box (10-12). This contains the element name (10), sequence (11) and an option that allows the reversal of the element direction (12). A restriction site analysis box displays restriction enzyme recognition sequences (16). The list may be modified by additions (17) or deletions (18), and other motifs may be added (19). Selected restriction sites may be shown within the sequence (20). Places within amino acid or ORF elements where restriction sites could occur without

altering the amino acid sequence are indicated (23). The view of the sequences may be altered by scrolling through the sequence (21), or by altering the magnification (22).

Analyzing Codon Choices

A sequence for which both amino acid (1) and DNA (2) sequence are provided (for example an ORF) will display the codons used for each amino acid below the sequence (3 and 4). The codons for each amino acid are color-coded, with those used at a frequency in the selected organism above the selected threshold shown in one color (8), and those that occur at a frequency in the selected organism below the selected threshold in another (9). This allows a rapid assessment of the suitability of a DNA sequence for expression in the selected host organism.



Positioning Sequence Elements: Icon View

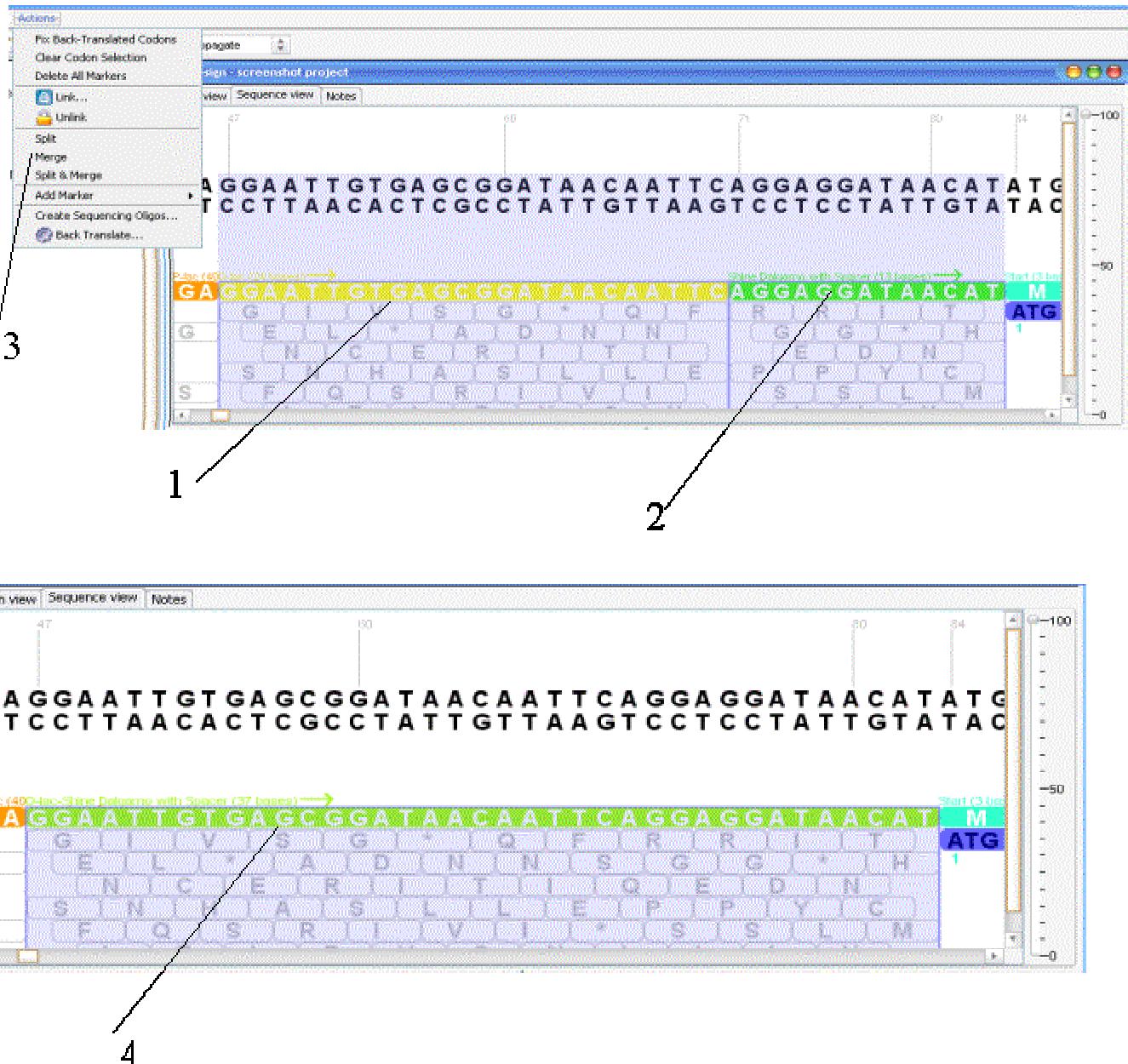
In the icon view each sequence element can be represented by an icon. The order of the icons can be altered by clicking on an element with the mouse and dragging it to a new position. Each sequence element can be identified by a name and by a color selected by the user.

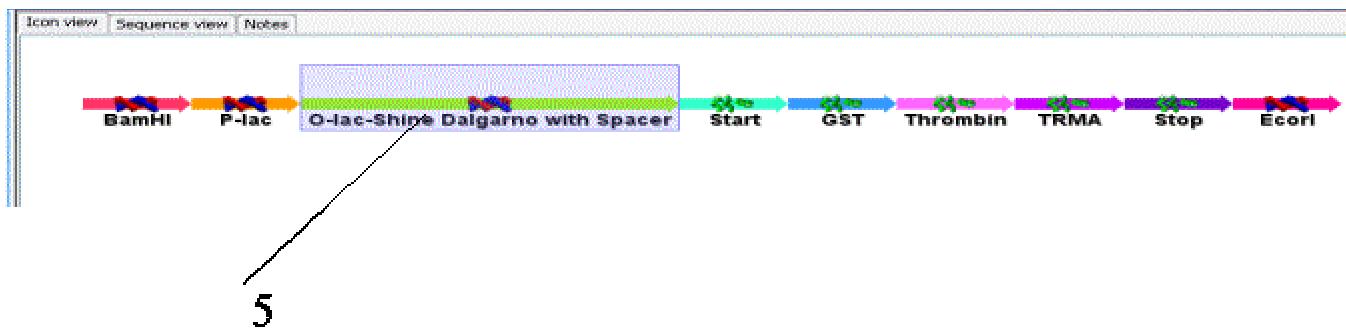


Advanced Sequence Element Manipulations

Merging Sequence Elements

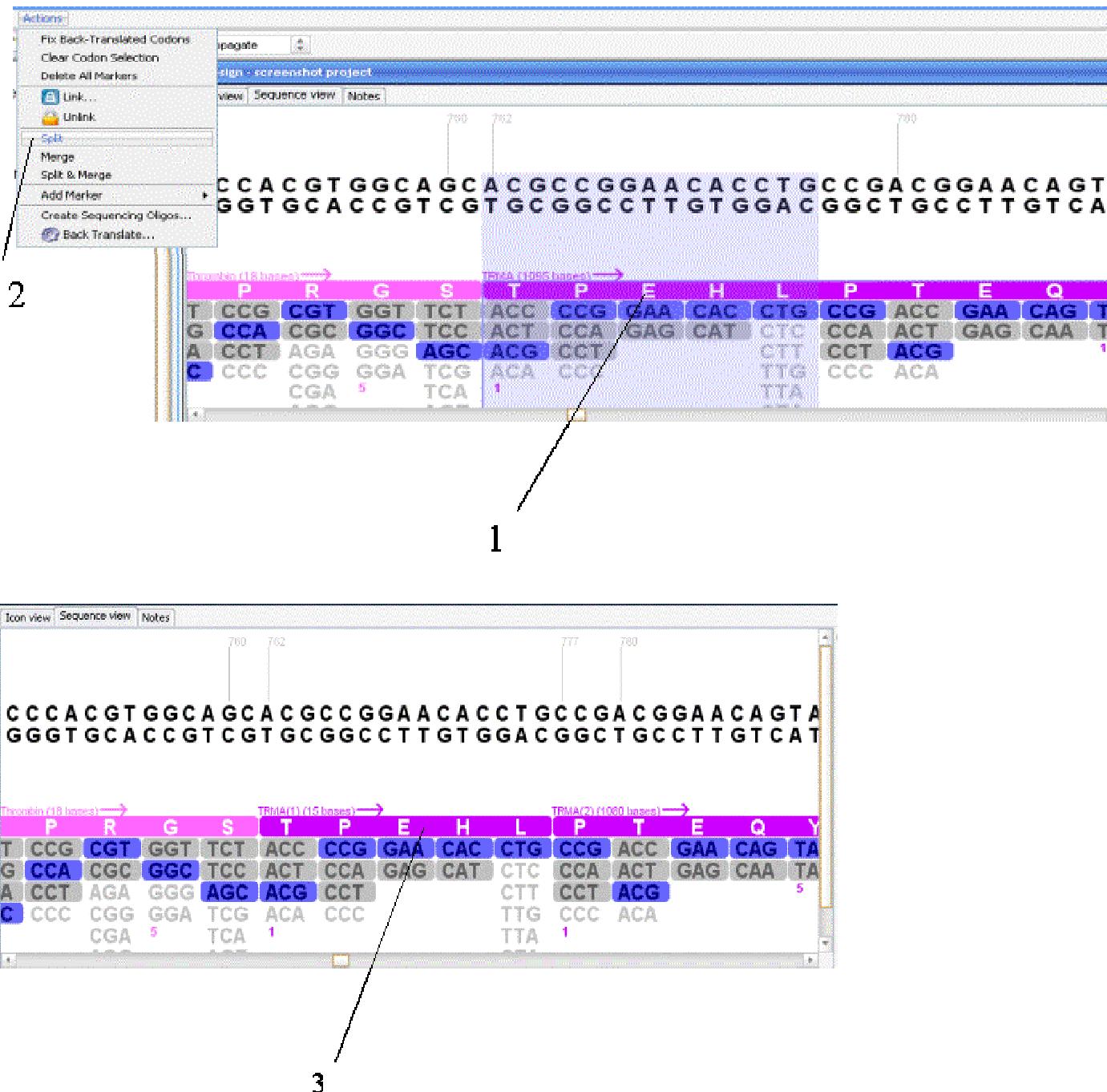
Two or more sequence elements (1,2) may be selected in sequence view. By selecting the Merge function (3) the sequence elements are combined into a single object in sequence (4) and icon (5) views.

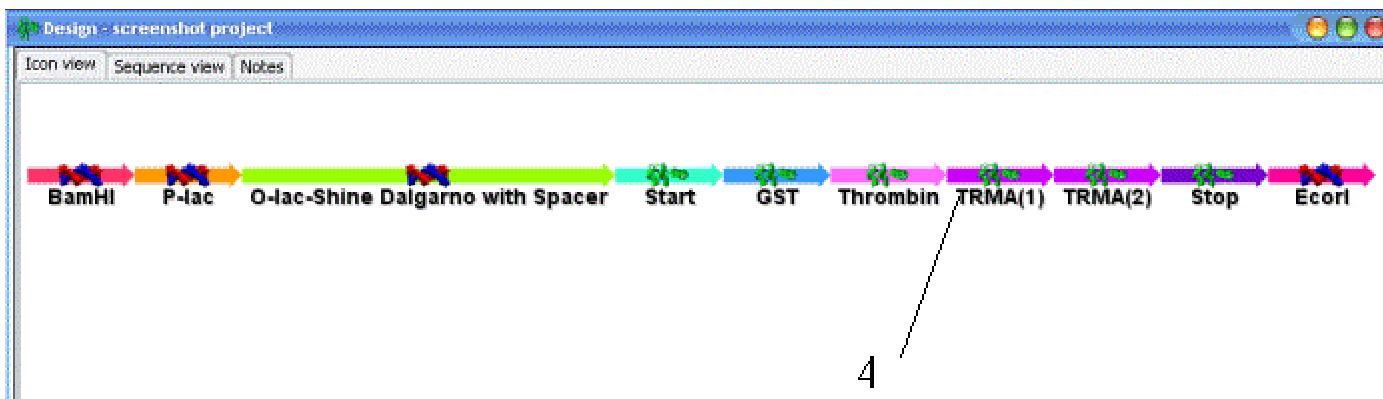




Splitting Sequence Elements

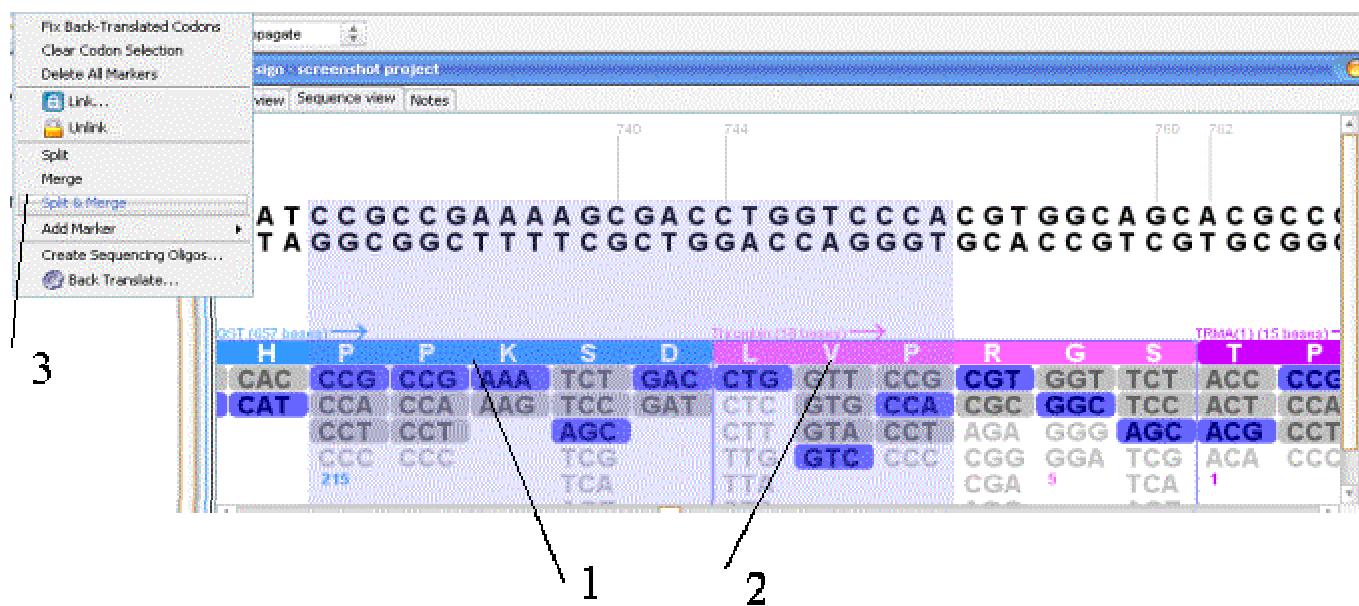
A section of a sequence element (1) may be selected in sequence view. By selecting the Split function (2) the selected element is separated into a new object in sequence (3) and icon (4) views.





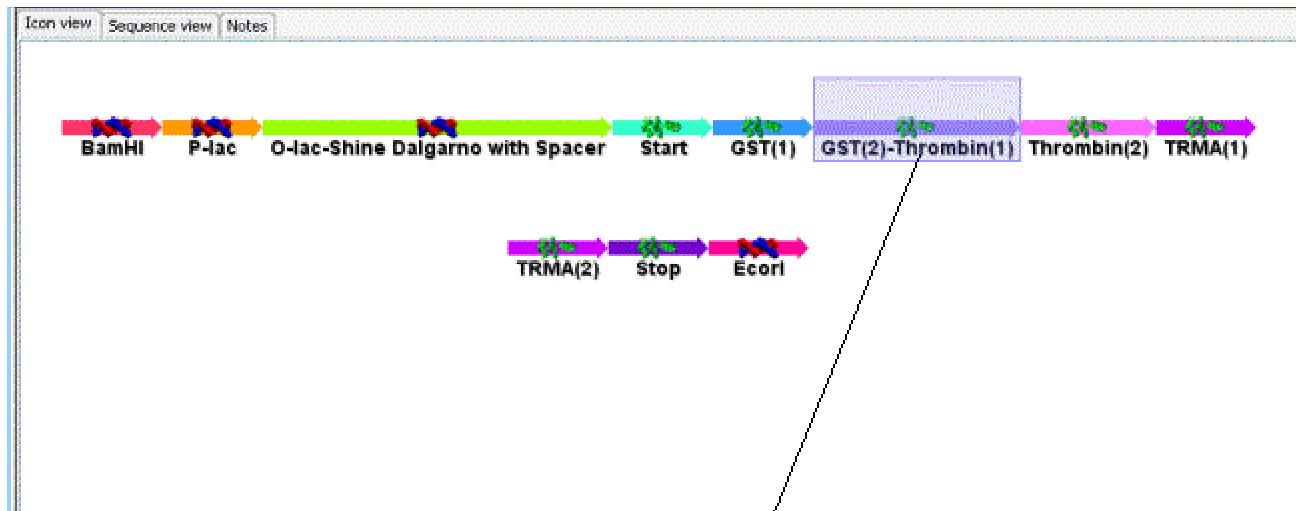
Splitting & Merging Sequence Elements

Two or more sequence elements that include one or more partial element (1,2) may be selected in sequence view. By selecting the Split & Merge function (3) the sequence elements are separated from the elements that they were part of and are combined into a single object in sequence (4) and icon (5) views.





4



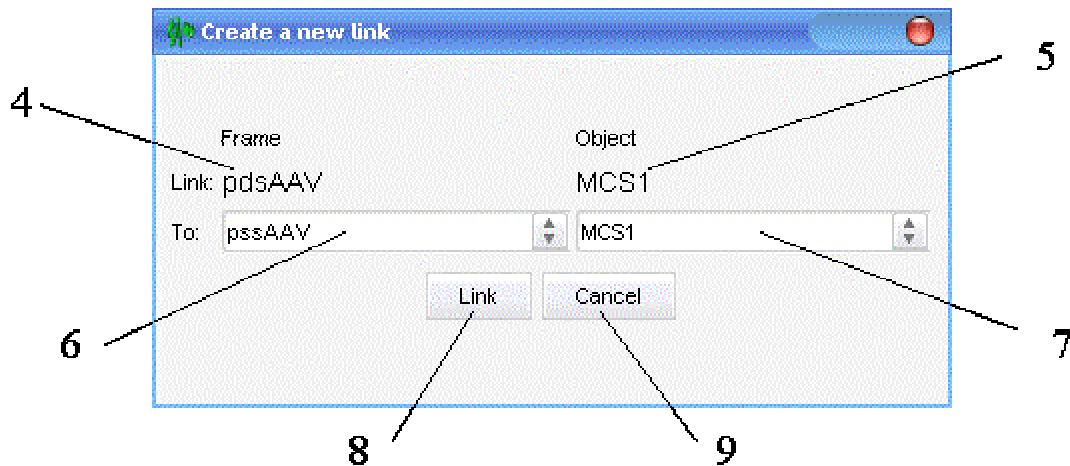
5

Linking Sequence Elements

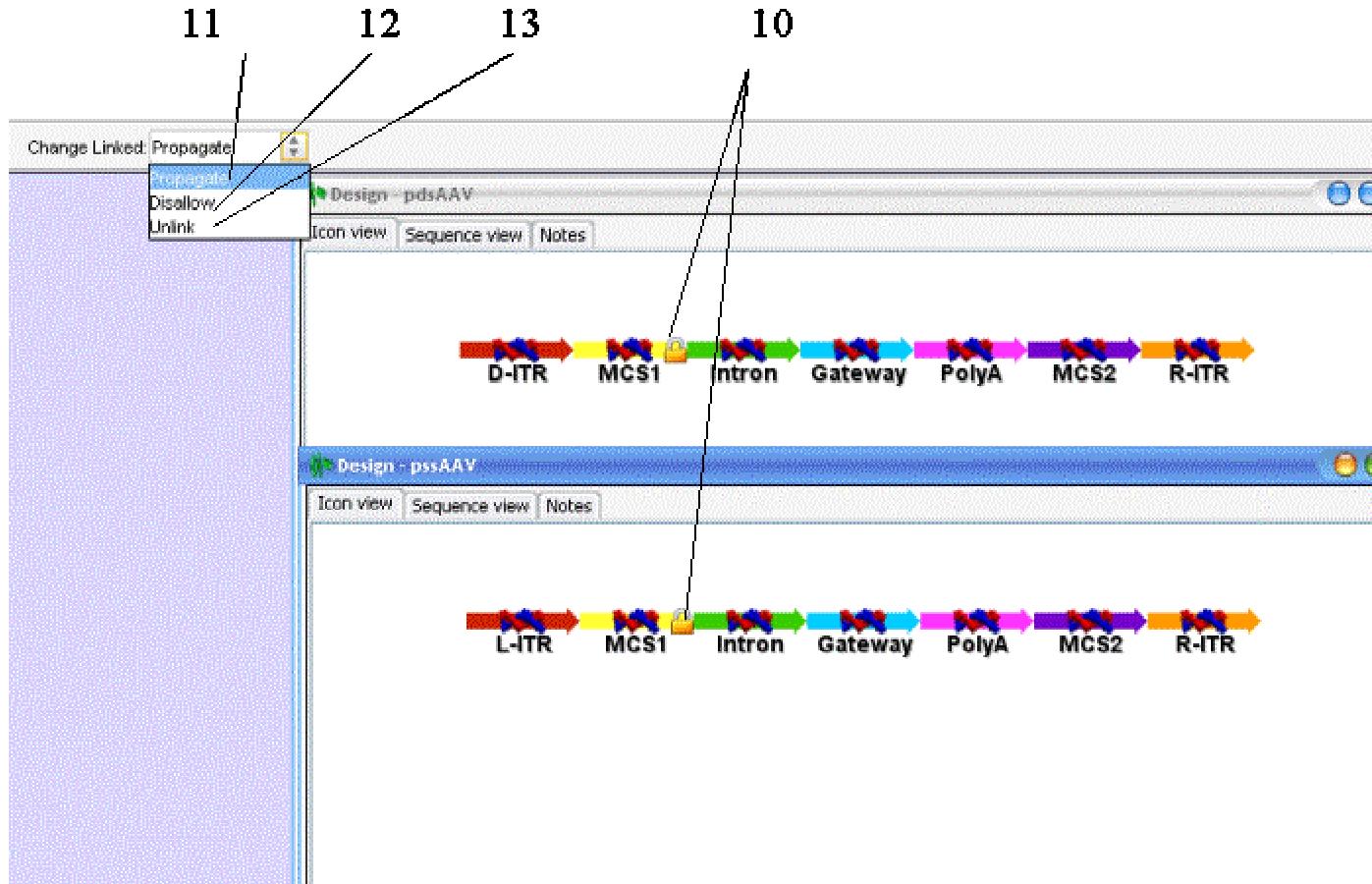
Sequence elements can be linked across projects. A change made to a sequence element in one project is propagated through all linked objects in other open projects.



A sequence element in one project is selected (2) and the Link action chosen (1). This creates a link dialog box which specifies the sequence element chosen (2 and 5) and the project (or frame) that it occurs in (4). This element can then be linked to another element (7) selected from another open project (6). Any open project can be chosen (6) and any sequence element within that project (7). Once an element has been chosen a link may be created (8) or the action cancelled (9).



Once sequence elements are linked, this is indicated in the icon view (10). The user can select options for changes to linked objects: they may be propagated through the linked elements (11), they may be disallowed (12) or the objects may be unlinked (13).



Reversing Sequence Elements

A sequence element may be reversed by selecting the element (1) and then selecting the reverse option (2). The arrow representing direction is then reversed in the icon view (3).

1

Icon view | Sequence view | Notes

Sequence view showing a DNA sequence from position 740 to 762. The sequence is:
 740: GCC GAAA AGC GAC CTGGT CCC AC GT GG CAG CAC GCC GGA A C A C T G C
 753: CGG CTT TC GCT GG ACC AGGT GC ACC GT CGT GC GGC TT GT GG ACC
 762: CGG CTT TC GCT GG ACC AGGT GC ACC GT CGT GC GGC TT GT GG ACC

Restriction sites and motifs shown in the sequence view:

- GST(2)-Thrombin(1) (24 bases) at positions 740-763
- Thrombin(2) (9 bases) at position 753
- TRMA(1) (15 bases) at position 762

2

Selected object properties:

Name: Throwbin(2)	<input type="checkbox"/> Reversed
Sequence data	
RGS	

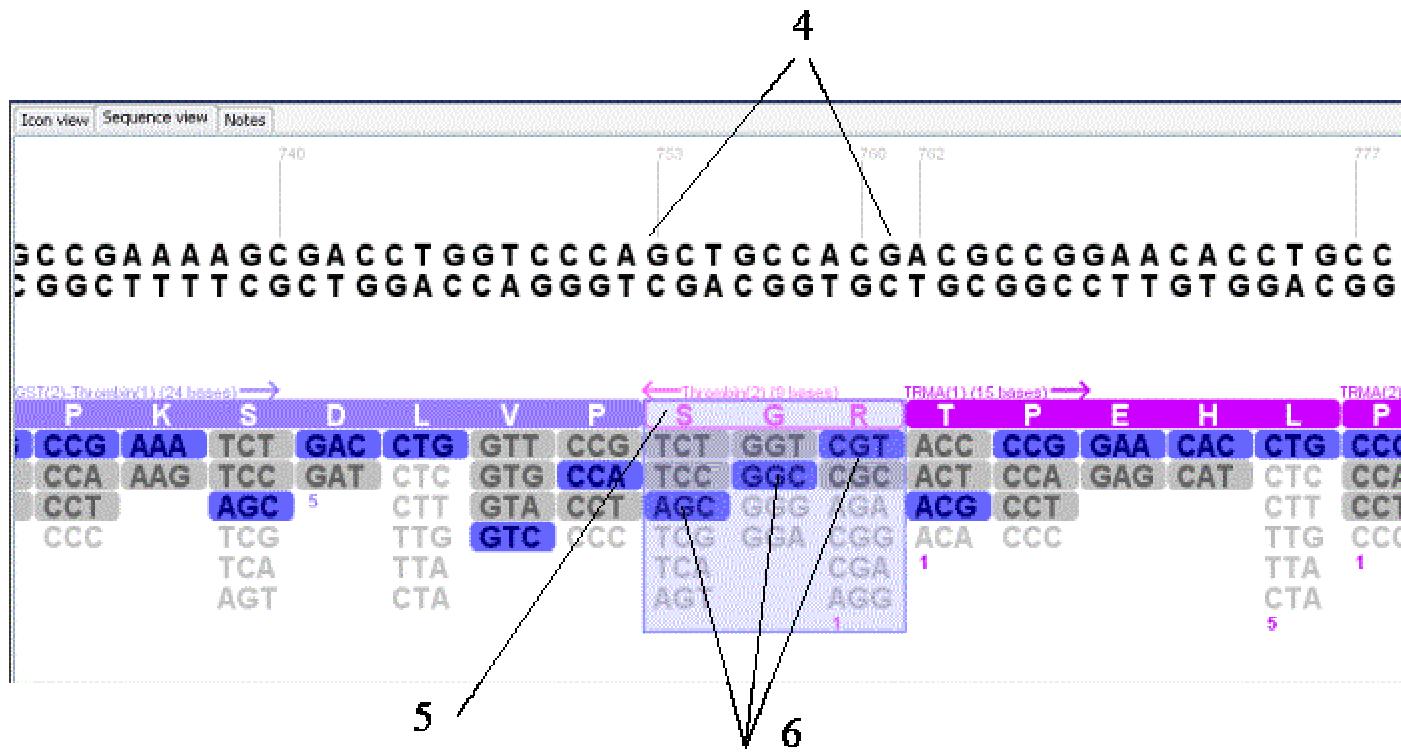
3

Icon view | Sequence view | Notes

Icon view showing the sequence structure with arrows indicating direction. The Thrombin(2) motif is shown with its direction reversed.

In the sequence view a reversed amino acid element is shown with the order of amino acids reversed (compare 1 and 5). However the codons for each amino acid are shown as the actual codons in the forward direction (6) for

ease of manipulation. The DNA sequence of each codon is reversed to show the actual DNA sequence (on the reverse strand) (4).

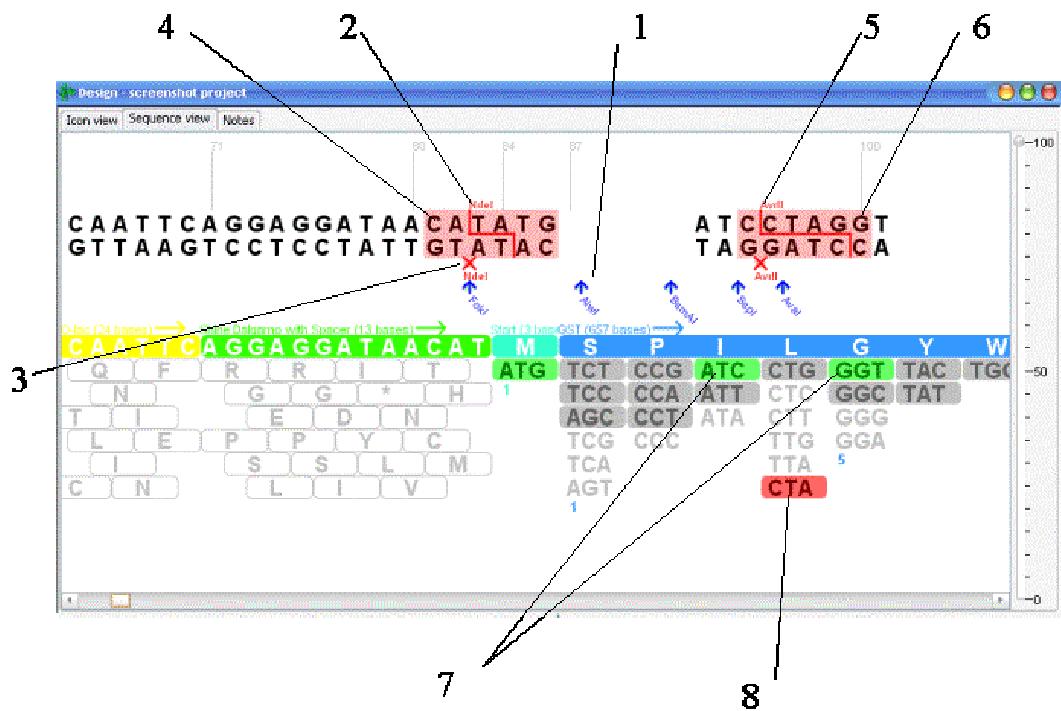


Protein 2 DNA

This feature allows the user to create a DNA sequence for an amino acid segment. User can choose any codon table provided by the software or import a new codon table from [KEGG](#) in GCG format. User can also customize the DNA by incorporating restriction sites and also [eliminate restriction sites](#)

Incorporating restriction sites

One or more restriction enzyme recognition sites can be selected from a list (see section 3.1, 16). By selecting the "show" button (see section 3.1, 20), the locations at which these restriction sites could occur in the amino acid sequence are indicated (1). Clicking on the arrow inserts the site and highlights the recognition sequence in the upper sequence panel (4 and 6), where the enzyme name (3) and overhangs generated by cleavage by that enzyme (2 and 5) are also indicated. The enzyme can be deselected by clicking the mouse on the name (3). Selection of a restriction recognition site fixes part of the DNA sequence, which then appears above the corresponding amino acid sequence. Below the amino acid sequence, the codons that are selected to incorporate the restriction site are indicated (7 and 8). These codons are also colored differently: those used at a frequency in the selected organism above the selected threshold shown in one color (7), and those that occur at a frequency in the selected organism below the selected threshold in another (8).



Eliminating restriction sites

A restriction site panel allows selection of one or more sites from a list by name (1) or recognition sequence (2). The panel also contains information about the location of cuts made by the enzyme (3). One or more site may be

avoided in a selected sequence element (4), or within the entire molecule under design (5). Avoided restriction sites will only be eliminated in back-translated regions. Restriction sites present in DNA regions, or ORFs whose DNA sequence has been fixed will not be removed. [See also 3 Viewing and Modifying Sequence Elements](#)

Name	Sequence	5' cut I	3' cut I	5' cut II	3' cut II	Locally Av.	Globally Av.	Motif...
AseI	ATTAAT	2	4			<input type="checkbox"/>	<input type="checkbox"/>	
Aval	CYCGRG	1	5			<input type="checkbox"/>	<input type="checkbox"/>	
Avall	GGWWCC	1	4			<input type="checkbox"/>	<input type="checkbox"/>	
AvrII	OCTAGG	1	5			<input checked="" type="checkbox"/>	<input type="checkbox"/>	
BamHI	GGATCC	1	5			<input type="checkbox"/>	<input type="checkbox"/>	
BbaI	GAAGAC	8	12			<input type="checkbox"/>	<input type="checkbox"/>	
BbvI	GCAGC	13	17			<input type="checkbox"/>	<input type="checkbox"/>	
BclI	TGATCA	1	5			<input type="checkbox"/>	<input type="checkbox"/>	
BglI	GCNNNN...	7	4			<input type="checkbox"/>	<input type="checkbox"/>	
BpuI	AGATCT	1	5			<input type="checkbox"/>	<input checked="" type="checkbox"/>	
BpiI	GCTNAGC	2	5			<input type="checkbox"/>	<input type="checkbox"/>	
BsaI	GGTCTC	7	11			<input type="checkbox"/>	<input type="checkbox"/>	
BsmAI	GTCTC	6	10			<input type="checkbox"/>	<input type="checkbox"/>	
BsmBI	CGTCTC	7	11			<input type="checkbox"/>	<input type="checkbox"/>	
BstEII	GGTNACC	1	6			<input type="checkbox"/>	<input type="checkbox"/>	
BstXI	CCANNNN...	8	4			<input type="checkbox"/>	<input type="checkbox"/>	
ClaI	ATCGAT	2	4			<input type="checkbox"/>	<input type="checkbox"/>	
DraIII	CACNNNG...	6	3			<input type="checkbox"/>	<input type="checkbox"/>	
EagI	CGGCGG	1	5			<input type="checkbox"/>	<input type="checkbox"/>	
EcoI	CTCTTC	7	10			<input type="checkbox"/>	<input type="checkbox"/>	
EcoRI	GAATTTC	1	5			<input type="checkbox"/>	<input type="checkbox"/>	
EcoRV	GATATC	3	3			<input type="checkbox"/>	<input type="checkbox"/>	
FokI	GGATG	14	18			<input type="checkbox"/>	<input type="checkbox"/>	
FseI	GGCGGGCC	6	2			<input type="checkbox"/>	<input type="checkbox"/>	
HindIII	AAGCTT	1	5			<input type="checkbox"/>	<input type="checkbox"/>	
KasI	GGGCC	1	5			<input type="checkbox"/>	<input type="checkbox"/>	
KpnI	GGTACC	5	1			<input type="checkbox"/>	<input type="checkbox"/>	
...	<input type="checkbox"/>	<input type="checkbox"/>	

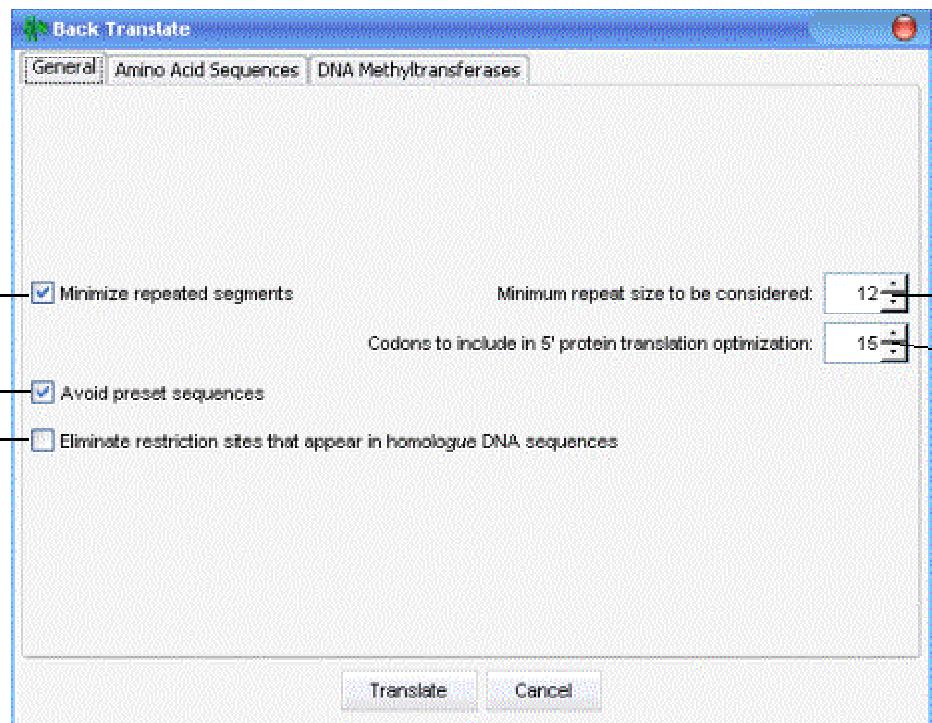
BackTranslation

This process converts all amino acid sequences that are not fixed into a DNA sequence that will encode them. The codon bias table is selected as described [here](#)

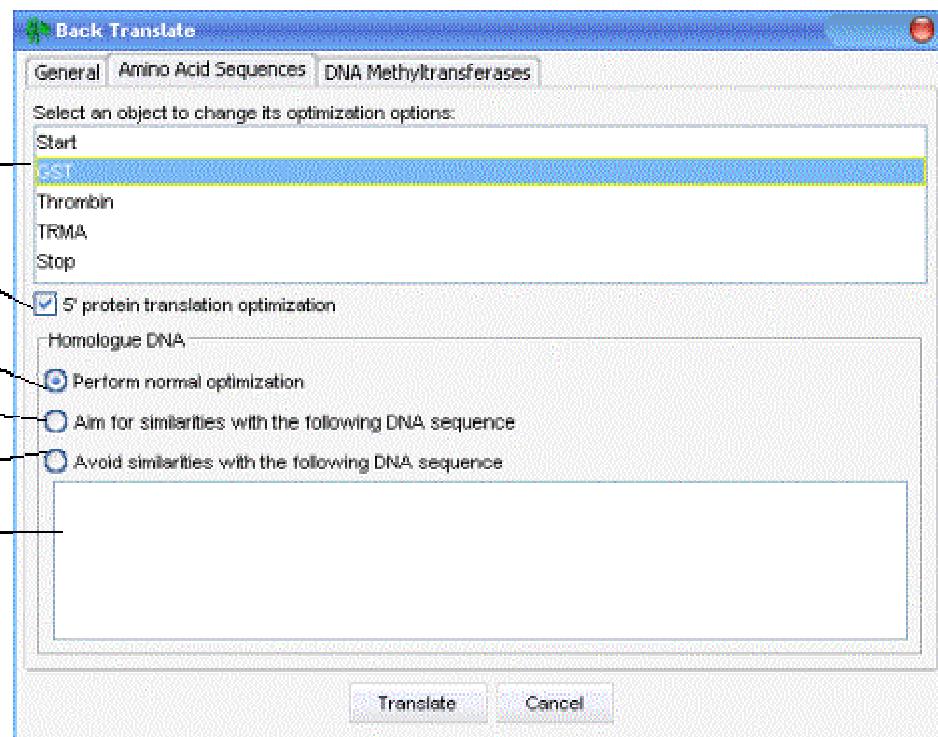
Optional settings for backtranslation include

- Minimizing repeat segments (1) including repeat size selection (4)
- Avoidance of pre-set sequences used as standards (2)

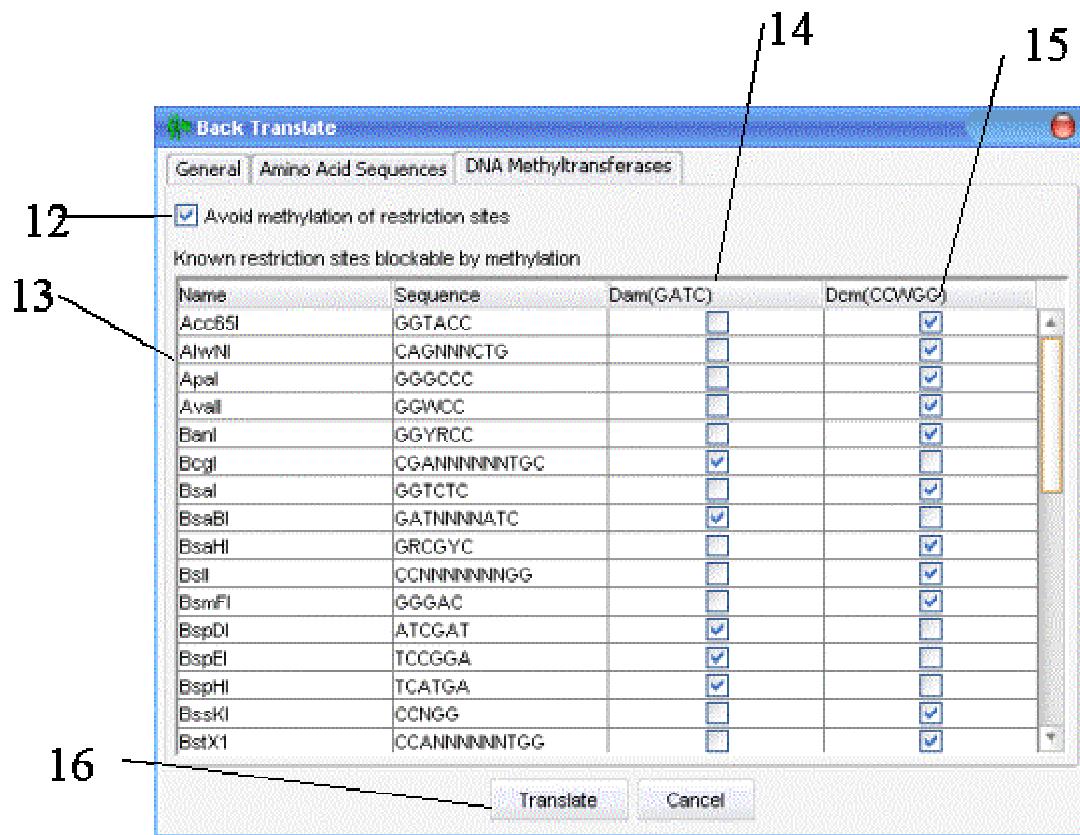
Each amino acid segment or open reading frame can also be individually selected (6) for independent optimization of the beginning of open reading frames for translation initiation (7), which involves minimizing secondary structures that could interfere with the initiating ribosome, and allows selection of the length of sequence so optimized (5).



Each amino acid segment or open reading frame can also be individually selected (6) and compared with a homologous reference sequence that can be pasted in (11). Sequence identity between the selected amino acid segment or open reading frame and the entered reference sequence can be maximized (9) or minimized (10), or not taken into account (8).



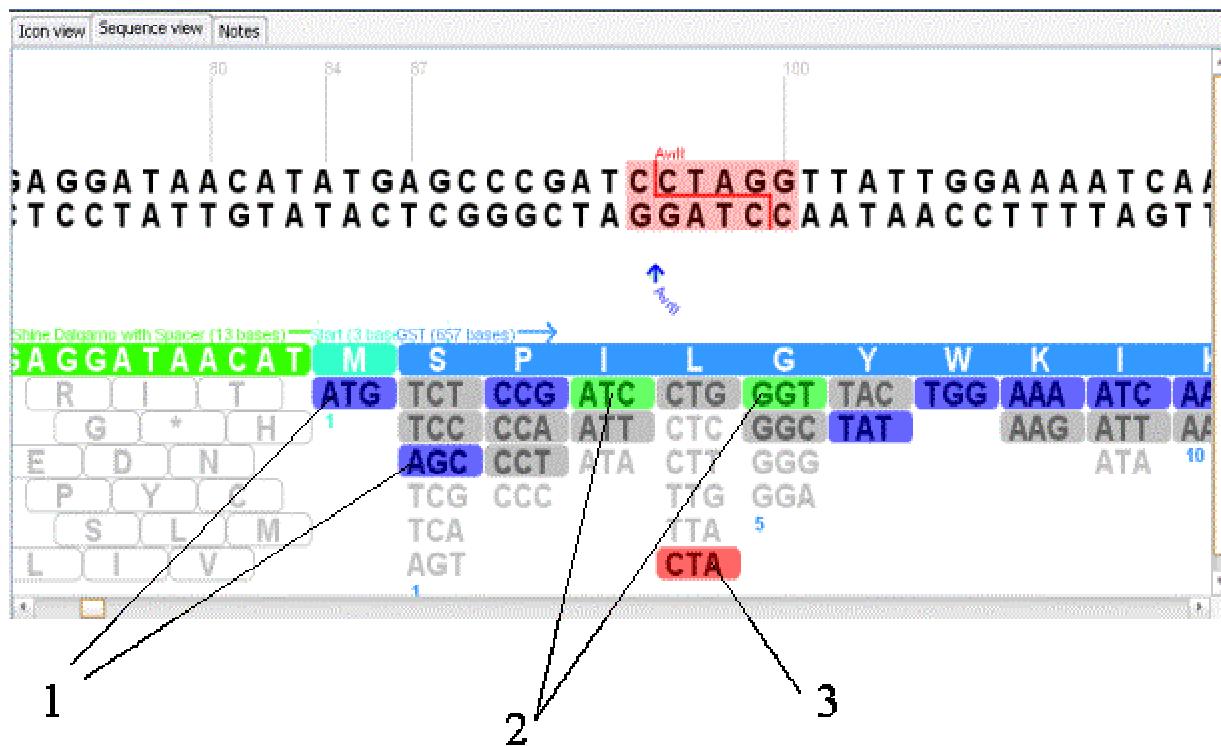
Restriction sites (13) known to be blocked by overlapping methylation, for example Dam (14) or Dcm (15) can be provided as a list. An option may be selected to eliminate overlapping inhibitory methylations (12).



Once all selections have been made the selected sequences may be backtranslated. The program uses the [codon table](#) selected.

BackTranslation View

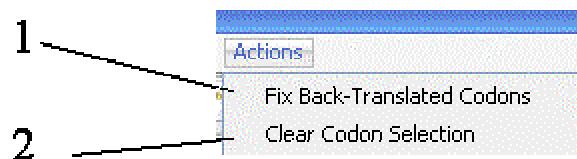
After back-translation, codons can be differentiated in the sequence view between codons assigned by the backtranslation (1), codons selected by the user (for example for inclusion of a restriction site) that are used at a frequency in the selected organism above the selected threshold shown in one color (2), and codons selected by the user that occur at a frequency in the selected organism below the selected threshold in another (3).



Backtranslation fixing / clearing

After backtranslation, codons for one or more elements can be fixed so that they are not altered by subsequent backtranslation steps (1).

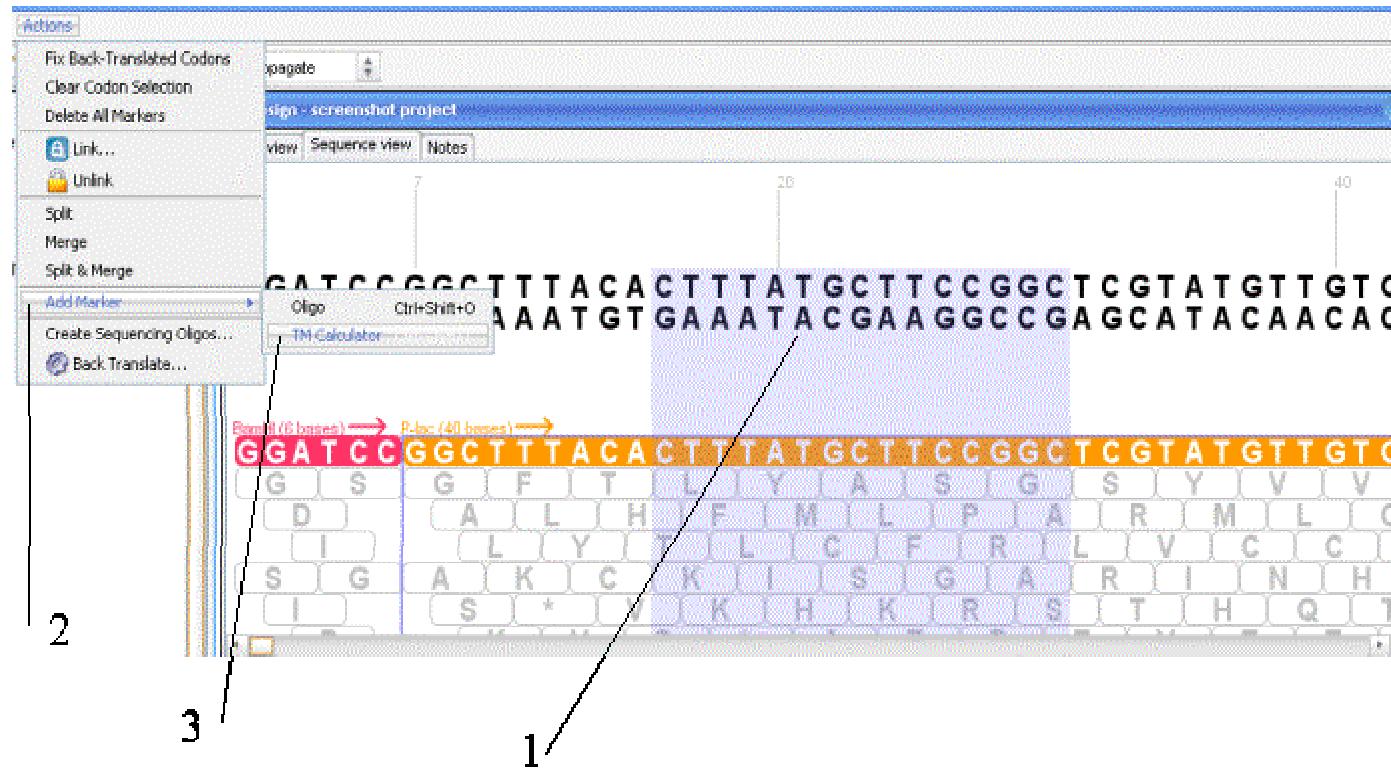
After backtranslation, codons for one or more elements can be cleared so that they may be re-selected, for example by subsequent backtranslation steps (2).



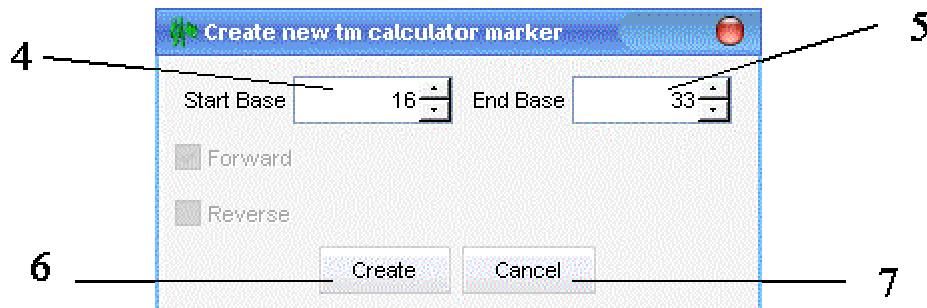
Markers and Oligos

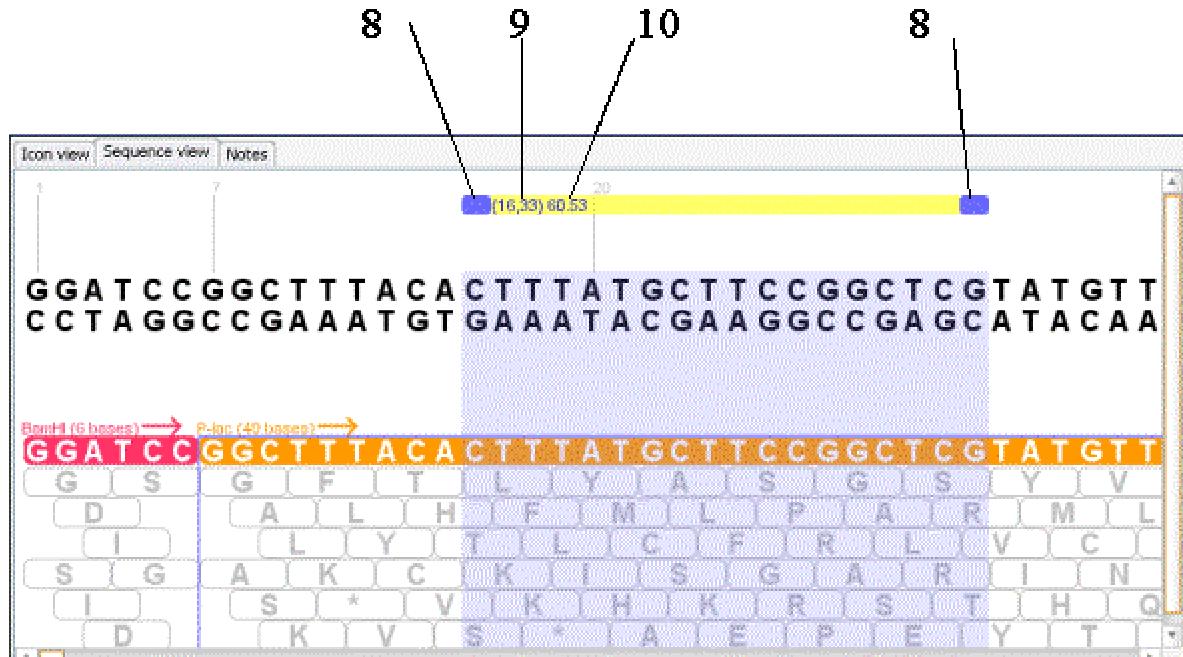
Creating a Tm Calculator

A sequence region may be selected (1) and a marker added (2). Selecting addition of a Tm calculator marker (3) allows modification of the start (4) and end (5) positions of the marker. The marker can then be created (6) or cancelled (7).



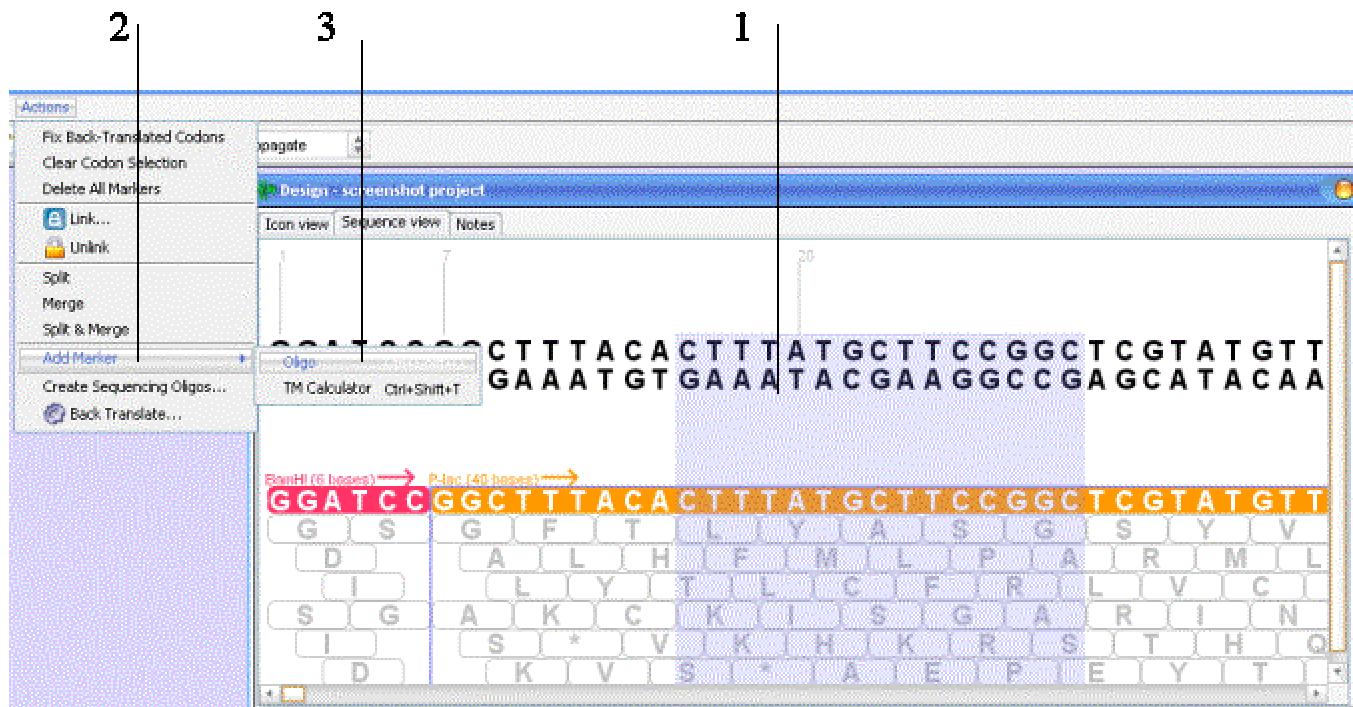
The ends of the Tm calculator are marked (8) and it contains start and stop positions (9) and the calculated Tm of the selected region of DNA (10). Either end (8) can be moved by clicking with the mouse and dragging. The Tm changes as this is done.



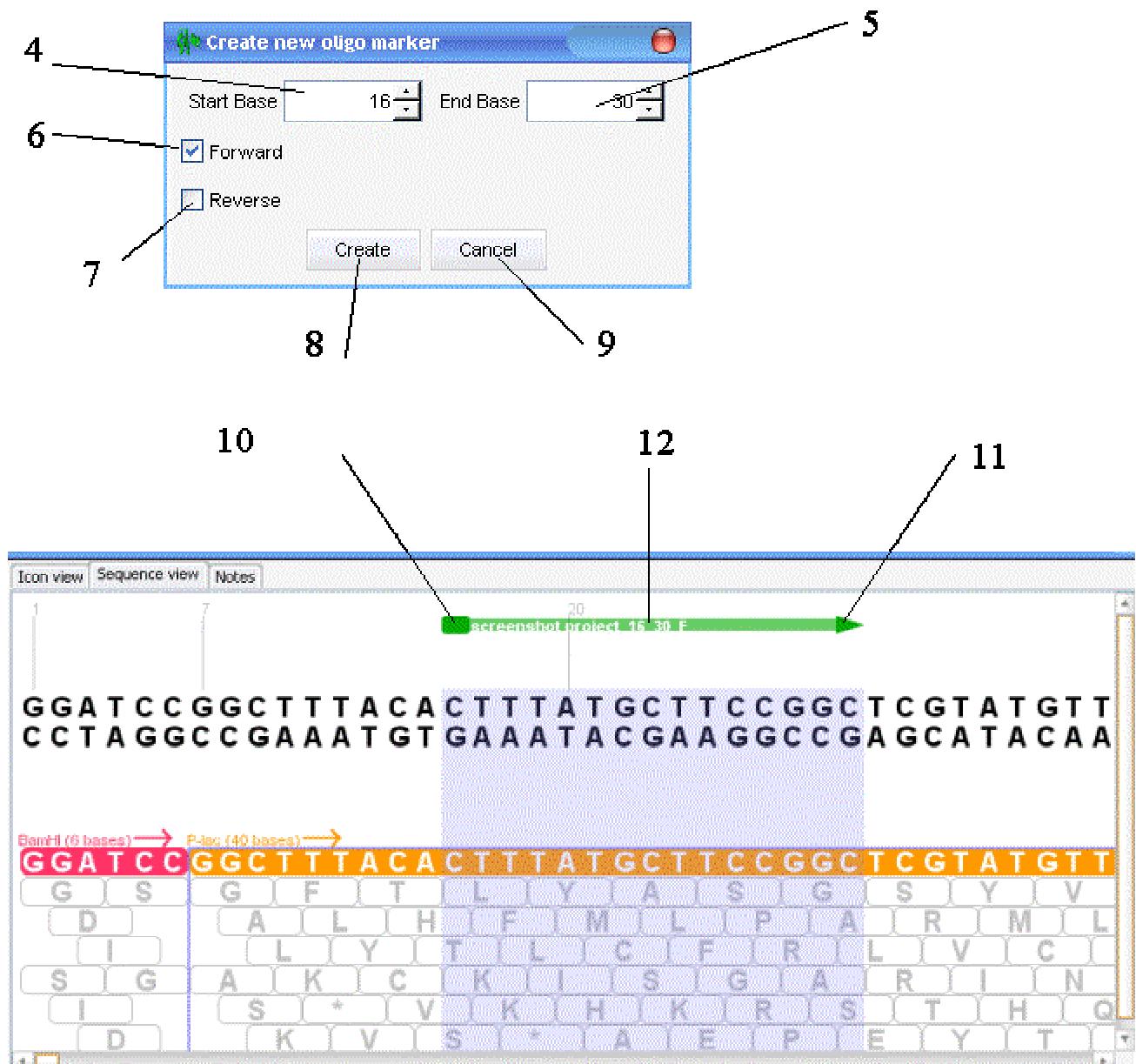


Creating an Oligo marker

A sequence region may be selected (1) and a marker added (2). Selecting addition of an oligo marker (3) allows modification of the start (4) and end (5) positions of the marker. Oligos may be generated in the forward (6) or reverse (7) directions. Once selected and oligo marker can be created (8) or cancelled (9).



The oligo marker has a 3' end (10) and a 5' end (11), both of which can be moved by clicking with the mouse and dragging. The marker also carries information describing its ends (12).



Creating Sequencing Oligos

A sequence region may be selected to obtain an optimal sequencing primer using the "create sequencing oligo" option under "Actions" menu. This will launch parameters dialog for the user.

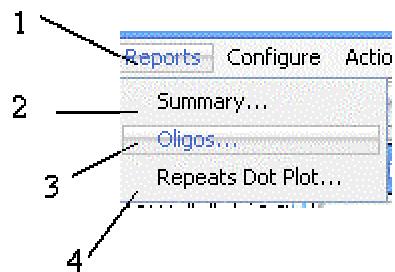
Sequence primers can also be created for the entire gene in forward and reverse directions spaced uniformly to aid sequence assembly. This option will allow the user to create primers optimally for sequencing large DNA fragments.

Sequencing oligos thus created can be exported and/or viewed in the icon view.

Reports

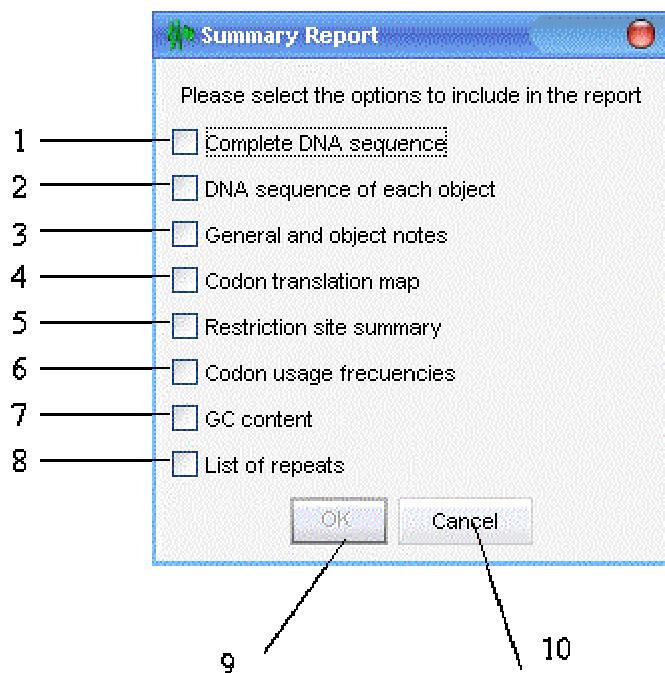
Creating Reports

Reports may be generated from the Report Menu (1). These can be a summary of information from the project (2), a report describing the oligos in the project (2) or a report visualizing the repeats present in the sequence (3).



Summary Report

A summary report can be generated to provide the complete DNA sequence (1), the DNA sequence of each sequence element (2), general notes for each sequence element (3), a codon translation map for each sequence element (4), a restriction site summary for the entire DNA sequence (5), the codon use frequencies for the entire sequence (6), the GC content of the entire sequence (7) and a list of repeats present in the entire sequence (8). Once the desired options have been selected a report may be generated (9) or cancelled (10). A sample report is shown below.



1. Complete DNA Sequence

GGATCCGGCTTACACTTATGCTTCGGCTCGTATGTTGTGGAGGAATTGTGAGCGGATAACAATTCAAGGAGGATAACA
TATGAGCCCCATCCTAGTTATTGAAAATCAAAGGCCCTGGTTCAGCCGACCGCTGCTGCTGGAATACCTGGAAGAAAAA
TACGAAGAACACCTGTACGAACCGATGAAGGTATAATGGCGAACAAAAAGTTGAACTGGGTCTGGAATTTCGAACC
TGCGTACTATATTGATGGTATGTAACCTGACCCAAATCATGGCCATCATCCGTTACATTGCCATAAACATAACATGCT
GGGTGGTTGTCTAAAGAACAGTGCGAACATTAGCATGCTGGAGGGTGCAGTCCTGGATATCGTTATGGTGTAGCCGCATT
GCTTACTCCAAGAACAGTCGAAACCCCTGAAGGTCGATTCCGTCCAAACTGCCGAAATGCTGAAATGTTGAGGACGTC
TGTGCCACAAAACGTACCTGAATGGCACCACGTAACCTCATCCGACTTCATGCTGTATGACGCGCTGGACGTAGTTCTGTA
CATGGACCCGATGTGCCTGGACGCATTCCCACCGTGTGTTCAAAAGCGTATTGAAGCCATCCCGCAGATCGATAAA
TACCTGAAATCCAGCAAATACATTGCATGGCGCTGCAGGGCTGGCAGGCAACCTCGCGGTGGCAGTCATCCGCCAAAA
GCGACCTGGTCCCACGTGGCAGCAGCCGAAACACCTGCCGACGGAACAGTACGAGGGCAGCTGGCTGAAAAGTTGTACG
TCTGCAATCTATGATGGCCCTTTCTGACCTGGTACCGAAGTCTCGTTCTCGGTGTCCTACTATCGTATGCGTGC
GAATTCCGTATCTGGCACGGTGACGACCTGTACCGACATTCTCGATCAGCAGACGAAATCTCGTATCCCGTGTGACT
CTTCCCAGCTGCGAGCGAACTGATCAACCAGCTGATGACTGCAATGATCGCAGGTGTACGCAACAACCCAGTGCCTGCTCA
CAAGCTGTTCAAATTGATTATCTGACTACTCTGAGCAACCAGGCTGTTGATCTCTGCTGTACCAAGAAACTGGACGAC
GAATGGCGTAGGAAGCGGAAGCAGCTGCGTACGCACACTGCGCAGACAACCTGAACGTGCACCTGATTGGCGTGTACGA
AAACCAAAATCGAACTGGATCAGGATTATATCGACGAACGCTGCCGGTGCAGGAAAGAAATGATCTACCGTCAGGTGGA
GAATTCTTCACCCAGCGAACGCAATGAAACATCCAGATGCTGGAATGGCGCTGGACGTTACCAAAGGTTCTAAAGGC
GACCTGCTGGAACTGACTGCGAACCGTAACTTAGCCTGGCTCTGCACGTAACCTCGACCGCTTCTGCCACCGAAA
TCGCAAAGCCCTCGTTGGCAGCCAAATATAACATTGCGCAAACACATCGATAACGTGCAGATCTCGCATGGCGC
AGAAGAAATTACCCAGCGATGAACGGCGTGAATTTAACCGTCTGCGAGGGCATCGATCTGAAATCTTACCAAGTGCAG
ACTATTTCTGTTGATCCGCCGCTTCCGGTCTGGACTCCGAAACCGAAAAGATGGTCAGGGTACCCCTGTATTCTGTATA
TCAGTTCCCGTACACTCACCATAATGGAATGTGGTGTACTGCTGACCGCGAAGTAAGAATT
TCAGTTCCCGTACACTCACCATAATGGAATGTGGTGTACTGCTGACCGCGAAGTAAGAATT

DNA sequence of each sequence element

>BamHI

GGATCC

>P-lac

GGCTTACACTTATGCTTCGGCTCGTATGTTGTGGAA

>O-lac-Shine Dalgarno with Spacer

GGAATTGTGAGCGGATAACAATTCAAGGAGGATAACAT

>Start

ATG

>GST(1)

AGCCCGATCCTAGGTTATTGAAAATCAAAGGCCTGGTCAGCCACCGCTCTGCTGGAATACCTGGAAGAAAAATACG
AAGAACACCTGTACGAACCGATGAAGGTGATAATGGCGAACAAAAAGTTGAACCTGGGTCTGGAATTCCGAAACCTGCC
GTACTATATTGATGGTGATGTAAGACTGACCCAATCCATGGCCATCATCCGTTACATTGCCGATAAACATAACATGCTGGGT
GGTGTCTAAAGAACGTGCCAAATTAGCATGCTGGAGGGTGCAGTCCTGGATATCCGTTATGGTGTAGCCGATTGCT
ACTCCAAAGACTTCGAAACCTGAAGGTCGATTCCTGTCACACTGCCGACTTCATGCTGTATGACGCCCTGGACGTAGTTCTGTACATG
CCACAAAACGTACCTGAATGGCACCACGTAACACTCCGACTTCATGCTGTATGACGCCCTGGACGTAGTTCTGTACATG
GACCCGATGTGCCCTGGACGCATTCCCAGAAACTGGTGTGTTCAAAAAGCGTATTGAAGCCATCCCCAGATCGATAAAATACC
TGAAATCCAGCAAATACATTGATGGCGCTGCAGGGCTGGCAGGCAACCTCGCCGGTGGCGATCAT

>GST(2)-Thrombin(1)

CCGCCGAAAGCGACCTGGTCCC

>Thrombin(2)

CGTGGCAGC

>TRMA(1)

ACGCCGGAACACCTG

>TRMA(2)

CCGACGGAACAGTACGAGGCGCAGCTGGCTAAAAAGTTGTACGTCTGCAATCTATGATGGCCCTTTCTGACCTGGTAC
CGGAAGTCTTCCGTTCTCCGGTGTCCACTATCGTATGCGTGAGAATTCCGTATCTGCACGACGGTACGACCTGTACCA
CATTATCTTCGATCAGCAGCGAAATCTGTATCCGCTTGACTCTTCCAGCTGCGAGCGAACTGATCAACCAGCTGATG
ACTGCAATGATCGCAGGTGTACGCAACAACCCAGTGTGCTGCGTCACAAGCTGTTCAAATTGATTATCTGACTACTCTGAGCA
ACCAGGCTGTGGTATCTGCTGTACCGACAAGAAACTGGACGACGAATGGCGTCAGGAAGCGGAAGCAGTGTGACGCACT
GGCGCACAGAACCTGAACGTGCCACTGATTGGCGTGTACGAAAACCAAATCGAAGCTGGATCAGGATTATATCGACGAA
CGTCTGCCGGTTGCAGGCAAAGAAATGATCTACCGTCAGGTGGAGAATTCTTCACCCAGCCGAACCGAGCAATGAACATCC
AGATGCTGGAATGGCGCTGGACGTTACCAAGGTTCTAAAGGCGACCTGCTGGAACGTACTGCGGCAACGGTAACCTTAG
CCTGGCTCTGGCACGTAACCTGACCGCTCTGCCACCGAAATCGCAAAGCCTCCGTTGCCAGCCAAATATAACATT
GGCGAACACATCGATAACCGTCAGATCATTGCGATGGCGAGAAGAATTCAACCCAGGCGATGAACGGCGTGCCTGAAT
TTAACCGTCTGCAGGGCATCGATCTGAAATCCTACAGTGCAGACTATTTCTGATCCGCCGTTCCGGTCTGGACTC
CGAAACCGAAAAGATGGTTCAAGCGTACCCCTGTATTCTGTATATCAGCTGCAACCCGAAACTCTGTGCAAAACCTGGAA
ACCCGTGAGCCAACCCATAAAGTCGAGCGTCTGGCTCTGTTGATCAGTTCCCGTACACTCACCATAATGGAATGTGGTGTAC
TGCTGACCGCGAAG

>Stop

TA

>EcorI

GAATT

General notes for each sequence element

Notes for BamHI

null

Notes for P-lac

Transcriptional promoter from the E coli lac operon

Notes for O-lac-Shine Dalgarno with Spacer

Transcription operator from the E coli lac operon-Consensus ribosome binding site plus 7 base spacer
that places an NdeI site at the initiation AUG Notes for Start

Start

Notes for GST(1)

Glutathione S-transferase

Notes for GST(2)-Thrombin(1)

Glutathione S-transferase-Cleaves between the arginine and glycine Notes for Thrombin(2)

Cleaves between the arginine and glycine

Notes for TRMA(1)

null

Notes for TRMA(2)

null

Notes for Stop

STOP

Notes for EcoRI

null

Translation Map for each sequence element

Start

1 ATG

1 M

GST(1)

1 AGCCCGATCCTAGGTTATTGGAAAATCAAAGGCCTGGTTAGCCGACCGTCTGCTGCTG

1 S P I L G Y W K I K G L V Q P T R L L L

61 GAATAACCTGGAAGAAAAATACGAAGAACACCTGTACGAACCGATGAAGGTGATAATGG

21 E Y L E E K Y E E H L Y E R D E G D K W

121 CGCAACAAAAAGTTGAACCTGGGTCTGGAATTCCGAACCTGCCGTACTATATTGATGGT

41 R N K K F E L G L E F P N L P Y Y I D G

181 GATGTAAACTGACCCAATCCATGCCATCATCCGTTACATTGCCGATAAACATAACATG

61 D V K L T Q S M A I I R Y I A D K H N M

241 CTGGGTGGTTGTCTAAAGAACGTGCCGAAATTAGCATGCTGGAGGGTGCAGTCCTGGAT

81 L G G C P K E R A E I S M L E G A V L D

301 ATCCGTTATGGTGTCAAGCCGCATTGCTTACTCCAAAGACTTCGAAACCCCTGAAGGTCGAT
101 I R Y G V S R I A Y S K D F E T L K V D
361 TTCCCTGTCCAAACTGCCGAAATGCTGAAAATGTTGAGGACCGTCTGTGCCACAAAACG
121 F L S K L P E M L K M F E D R L C H K T
421 TACCTGAATGGCGACCACGTAACCATCCGGACTTCATGCTGTATGACGCGCTGGACGTA
141 Y L N G D H V T H P D F M L Y D A L D V
481 GTTCTGTACATGGACCCGATGTGCCTGGACGCATTCCGAAACTGGTGTGTTCAAAAAG
161 V L Y M D P M C L D A F P K L V C F K K
541 CGTATTGAAGCCATCCCGCAGATCGATAAATACCTGAAATCCAGCAAATACATTGCATGG
181 R I E A I P Q I D K Y L K S S K Y I A W
601 CCGCTGCAGGGCTGGCAGGCAACCTTCGGCGGTGGCGATCAT
201 P L Q G W Q A T F G G G D H

GST(2)-Thrombin(1)

1 CCGCCGAAAAGCGACCTGGTCCCA
1 P P K S D L V P

Thrombin(2)

1 CGTGGCAGC
1 R G S

TRMA(1)

1 ACGCCGGAACACCTG
1 T P E H L

TRMA(2)

1 CCGACGGAACAGTACGAGGCGCAGCTGGCTAAAAAGTTGTACGTCTGCAATCTATGATG
1 P T E Q Y E A Q L A E K V V R L Q S M M
61 GCCCCTTTCTGACCTGGTACCGGAAGTCTTCCGTTCTCCGGTGTCCCCTATCGTATG
21 A P F S D L V P E V F R S P V S H Y R M
121 CGTGCAGAATTCCGTATCTGGCACGACGGTGACGACCTGTACCACATTATCTCGATCAG
41 R A E F R I W H D G D D L Y H I I F D Q
181 CAGACGAAATCTCGTATCCGCGTTGACTCTTCCAGCTGCGAGCGAACTGATCAACCAG
61 Q T K S R I R V D S F P A A S E L I N Q
241 CTGATGACTGCAATGATCGCAGGTGTACGCAACAACCCAGTGTGCGTCACAAGCTGTC
81 L M T A M I A G V R N N P V L R H K L F
301 CAAATTGATTATCTGACTACTCTGAGCAACCAGGCTGGTATCTCTGCTGTACCAACAG
101 Q I D Y L T T L S N Q A V V S L L Y H K
361 AAACTGGACGACGAATGGCGTCAGGAAGCGAACGACTGGTGACCGACTGCGCCACAG
121 K L D D E W R Q E A E A L R D A L R A Q
421 AACCTGAACGTGCACCTGATTGGCGTGCTACGAAAACCAAAATCGAACTGGATCAGGAT
141 N L N V H L I G R A T K T K I E L D Q D
481 TATATCGACGAACGTCTGCCGGTGCAGGCAAAGAAATGATCTACCGTCAGGTGGAGAAT
161 Y I D E R L P V A G K E M I Y R Q V E N
541 TCTTCACCCAGCCGAACGCAGCAATGAACATCCAGATGCTGGAATGGCGCTGGACGTT
181 S F T Q P N A A M N I Q M L E W A L D V
601 ACCAAAGGTTCTAAAGGCGACCTGCTGGAACGTGACTGCGGCAACGGTAACTTAGCCTG

201 T K G S K G D L L E L Y C G N G N F S L
 661 GCTCTGGCACGTAACCTCGACCGCGTTCTGCCACCGAAATCGCAAAGCCTCCGTTGCG
 221 A L A R N F D R V L A T E I A K P S V A
 721 GCAGCCCAATATAACATTGCCGAAACCACATCGATAACGTGCAGATCATTGCATGGCG
 241 A A Q Y N I A A N H I D N V Q I I R M A
 781 GCAGAAGAATTCACCCAGGCGATGAACGGCGTGCCTGAATTAAACGTCTGCAGGGCATC
 261 A E E F T Q A M N G V R E F N R L Q G I
 841 GATCTGAAATCCTACCAAGTGCAGACTATTTCGTTGATCCGCCGCGTTCCGGTCTGGAC
 281 D L K S Y Q C E T I F V D P P R S G L D
 901 TCCGAAACCGAAAAGATGGTTCAGGCACCCCTGAGCCAAACCCATAAAGTCGAGCGTCTGGCT
 301 S E T E K M V Q A Y P R I L Y I S C N P
 961 GAAACTCTGTGCAAAACCTGGAAACCCCTGAGCCAAACCCATAAAGTCGAGCGTCTGGCT
 321 E T L C K N L E T L S Q T H K V E R L A
 1021 CTGTTGATCAGTCCCGTACACTCACCATATGGAATGTGGTGTACTGCTGACCGCGAAG
 341 L F D Q F P Y T H H M E C G V L L T A K
 Stop
 1 TAA
 1 *

Restriction Sites for entire sequence

Name	Seq.	Locations
AatI	AGGCCT	115
AccI	GTMKAC	none
AflII	CTTAAG	none
AgeI	ACCGGT	none
AlwI	GGATC	0, 1246, 1(c), 91(c), 1652(c)
AlwNI	CAGNNNNCTG	1193, 1349
Apal	GGGCC	none
ApalII	GTGCAC	1205
AscI	GGCGCGCC	none
AseI	ATTAAT	none
AvaI	CYCGRG	none
AvaII	GGWCC	484, 577, 745
AvrII	CCTAGG	94
BamHI	GGATCC	0
BbsI	GAAGAC	863(c)
BbvI	GCAGC	756, 796, 1334, 1496, 139(c), 688(c), 992(c), 1057(c), 1725(c)
BclI	TGATCA	1005, 1801
BglI	GCCNNNNNGC	none
BglII	AGATCT	none
BlpI	GCTNAGC	none

BsaI	GGTCTC	none
BsmAI	GTCTC	1637(c)
BsmBI	CGTCTC	none
BstEII	GGTNACC	none
BstXI	CCANNNNNNTGG	none
ClaI	ATCGAT	647, 1526, 1613
DraIII	CACNNNGTG	1170
EagI	CGGCCG	none
EarI	CTCTTC	none
EcoRI	GAATTC	902, 1312, 1562, 1859
EcoRV	GATATC	383
FokI	GGATG	295(c), 530(c), 637(c), 725(c), 1345(c)
FseI	GGCCGGCC	none
HindIII	AAGCTT	none
KasI	GGCGCC	none
KpnI	GGTACC	853
MluI	ACGCGT	131
NarI	GGCGCC	none
NcoI	CCATGG	285
NdeI	CATATG	80, 1823
NheI	GCTAGC	none
NotI	GC GGCCGC	none
NsiI	ATGCAT	none
PacI	TTAATTAA	none
PciI	ACATGT	none
PmeI	GTTTAAAC	none
PstI	CTGCAG	689, 1604
PvuI	CGATCG	none
PvuII	CAGCTG	797, 990, 1013, 1723
SacI	GAGCTC	none
SacII	CCGCGG	none
SalI	GTCGAC	none
SapI	GCTCTTC	none
SfiI	GGCCNNNNNGGCC	none
SgrAI	CRCCGGYG	none
SmaI	CCCGGG	none
SpeI	ACTAGT	none
SphI	GCATGC	360
SspI	AATATT	none
StuI	AGGCCT	115
SwaI	ATTTAAAT	none
TliI	CTCGAG	none
XbaI	TCTAGA	none

XbaI	CTCGAG	none
XmaI	CCCGGG	none
XmnI	GAANNNNTTC	1312

Codon Usage Table

AmAcid		Codon	Number	/1000	Fraction
END	TAA	1	1.68	1.0	
END	TGA	0	0.0	0.0	
END	TAG	0	0.0	0.0	
ALA	GCG	11	18.58	0.25	
ALA	GCT	7	11.82	0.16	
ALA	GCA	18	30.40	0.41	
ALA	GCC	7	11.82	0.16	
CYS	TGC	6	10.13	0.66	
CYS	TGT	3	5.06	0.33	
ASP	GAC	22	37.16	0.56	
ASP	GAT	17	28.71	0.43	
GLU	GAA	37	62.5	0.86	
GLU	GAG	6	10.13	0.13	
PHE	TTC	16	27.02	0.69	
PHE	TTT	7	11.82	0.30	
GLY	GGT	15	25.33	0.55	
GLY	GGC	12	20.27	0.44	
GLY	GGG	0	0.0	0.0	
GLY	GGA	0	0.0	0.0	
HIS	CAC	12	20.27	0.70	
HIS	CAT	5	8.44	0.29	
ILE	ATC	21	35.47	0.61	
ILE	ATT	13	21.95	0.38	
ILE	ATA	0	0.0	0.0	
LYS	AAA	29	48.98	0.78	

LYS	AAG	8	13.51	0.21
LEU	CTG	66	111.48	0.98
LEU	CTC	0	0.0	0.0
LEU	CTT	0	0.0	0.0
LEU	TTG	0	0.0	0.0
LEU	TTA	0	0.0	0.0
LEU	CTA	1	1.68	0.01
MET	ATG	21	35.47	1.0
ASN	AAC	21	35.47	0.91
ASN	AAT	2	3.37	0.08
PRO	CCG	21	35.47	0.72
PRO	CCA	3	5.06	0.10
PRO	CCT	5	8.44	0.17
PRO	CCC	0	0.0	0.0
GLN	CAG	22	37.16	0.81
GLN	CAA	5	8.44	0.18
ARG	CGT	25	42.22	0.75
ARG	CGC	8	13.51	0.24
ARG	AGA	0	0.0	0.0
ARG	CGG	0	0.0	0.0
ARG	CGA	0	0.0	0.0
ARG	AGG	0	0.0	0.0
SER	TCT	8	13.51	0.28
SER	TCC	9	15.20	0.32
SER	AGC	11	18.58	0.39
SER	TCG	0	0.0	0.0
SER	TCA	0	0.0	0.0
SER	AGT	0	0.0	0.0
THR	ACC	12	20.27	0.48
THR	ACT	7	11.82	0.28
THR	ACG	6	10.13	0.24
THR	ACA	0	0.0	0.0
VAL	GTT	10	16.89	0.31
VAL	GTG	8	13.51	0.25

VAL	GTA	8	13.51	0.25
VAL	GTC	6	10.13	0.18
TRP	TGG	7	11.82	1.0
TYR	TAC	18	30.40	0.66
TYR	TAT	9	15.20	0.33

GC content

GC Percentage: 51.206434316353885%

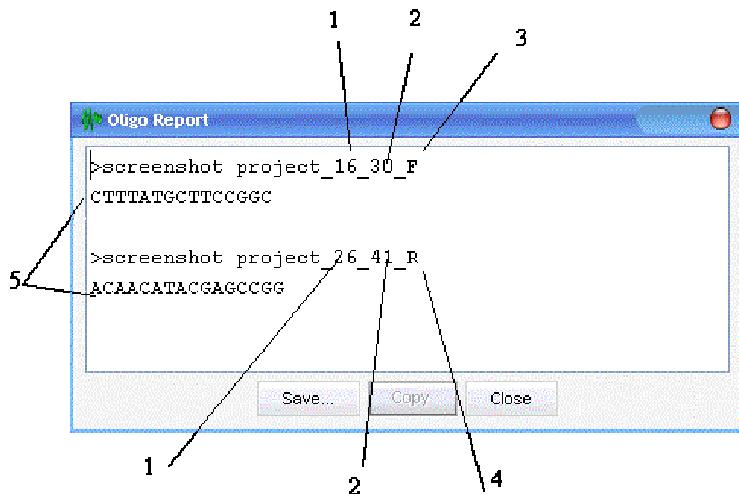
Repeat Analysis

Repeats greater than or equal to 12, in screenshot project

None

Oligo Report

The oligo report displays the sequences of all oligos present in the project (5). For each oligo the start (1) and end (2) positions are displayed, together with the orientation: either forward (3) or reverse (4) relative to the direction of the sequence.



Custom Gene Synthesis from DNA2.0

(<http://www.dnatwopointo.com>)

- [Why make synthetic genes?](#)
- [Our prices and delivery guarantees](#)
- [Our past performance](#)
- [Product specifications](#)
- [How to order](#)
- [Request a quote](#)

Why make synthetic genes?

Our mission is to simplify the steps between thinking of an experiment and doing it. Synthetic DNA provides an immediate and easy path from sequence databases and DNA manipulation software to physical reality. Many of the limitations of natural DNA can be overcome by using DNA-2-Go, our custom gene synthesis service:

- Any sequence is available. Finding a sequence in a computer database is much easier than locating a source of the DNA, even just to use as a template for PCR amplification. Suitable DNA may be impossible to find, costly or time-consuming to obtain. Synthesizing the gene eliminates availability problems and provides several additional degrees of freedom:
- Heterologous expression can be greatly enhanced. Different organisms prefer different codons to encode the same amino acid. While there is still considerable debate over the reasons for this, it is extremely clear that genes containing codons that are infrequently used in the desired expression host can suffer from poor protein expression. Because a synthetic gene is designed from the ground up, the codons used to encode different amino acids can be chosen to suit the expression host. We have seen numerous examples where codon optimization drastically increases protein expression.
 - See our review "Codon bias and heterologous protein expression" published in the July 04 issue of Trends in Biotechnology. E-mail info@dna20.com for a copy.
- > Modify the sequence. Quickly add N- or C-terminal affinity tags, insert linkers, create fusion proteins or fusion transcripts, add upstream ribosome binding sites to give optimal bacterial expression or surround a eukaryotic initiating AUG codon with an optimal Kozak consensus sequence. Introns can be removed and sites for protein modification such as phosphorylation, glycosylation or PEGylation can be added or eliminated as desired. Gene synthesis allows you to define exactly what specifications the gene should conform to, instead of consuming your time and effort in working around the limitations of the natural gene.
- Create variants. We use our DNA-2-Go gene synthesis process to create variants for our DeNovo Genes protein engineering technology. For this we developed a process where we can create many variants of a gene at a significantly reduced per base pair cost. This service is also available to our DNA-2-Go customers. For a variant-set quote, e-mail us at info@dna20.com or call us toll-free at 1-877-DNA-TOGO (1-877-362-8646).

- Too much freedom? Everyone at DNA 2.0 has considerable first-hand experience in gene design and synthesis. We pride ourselves on our scientist-to-scientist support, and will be very happy to walk you through the entire gene design process, making sure that your gene includes all of the features that you will need.

Delivery Time and Turnaround Guarantee

- 50 bp to 1675 bp under 10 days Guaranteed within 15 days
- 1676bp to 3kb under 15 days Guaranteed within 20 days
- 3kb onwards under 25 days

Special Rates: Contact us with details while requesting quotes. We offer great rates and discounts for Academic labs. Bulk orders. Gene variant sets.

Our past performance

(over the most recent 1000 genes made)

Size	Average(Days)	Median(Days)
0-500bp	9.2	9
500-1kb	9.8	9
1kb-1.5kb	10.5	10
1.5kb-2kb	14.7	13
2kb-3kb	15.6	13
3kb-10kb	26.1	23

Product Specifications

- **Cloning Vehicles :** Genes are cloned into our own pJ2 vector or the Invitrogen Gateway® pDONR221 vector at no additional charge. Let us know if you have a preference. Our default cloning vector is pJ2
- Cloning into any customer-specified alternative vector is also available for an additional charge. Please inquire for quote.

- Transfection ready DNA, Purified Protein and/or Antibodies You need milligrams of superclean plasmid DNA? And you want us to express and purify the protein as well? And generate antibodies to it? We can probably help you, either directly or through our integrated biotech service alliance. Please inquire for quote and specifics. Gene Design
- Assistance with gene design including codon optimization, feature choices and other bioinformatic assistance is available free of charge. Email info@dna20.com or call 1-877-DNA-TOGO. Delivery Specifications

You will receive:

- 2µg lyophilized plasmid containing your gene insert.
 - An E.coli culture stab transformed with a plasmid containing your gene.
 - Sequence chromatograms covering both strands of your gene (electronic).
 - The full sequence of the plasmid containing your gene insert (electronic).
- Confidentiality. All of your information is kept completely confidential. Non-disclosure agreements are available upon request.

How to obtain a Quote

- You can receive a quote for your newly designed gene via Gene Designer software. Just go to the menu item "actions" and select "quote gene".
- or E-mail us at info@dna20.com or call us toll-free at 1-877-DNA-TOGO (1-877-362-8646)

How To Order:

- You can order your newly designed gene via Gene Designer software. Just go to the menu item "actions" and select "order gene". You need a new account with DNA2.0 prior to your first order.
- E-mail us at info@dna20.com or call us toll-free at 1-877-DNA-TOGO (1-877-362-8646)
- What do you need?
 - The DNA sequence
 - A purchase order (PO) number or credit card