MacVector 13.0

for Mac OS X

Primer Design Tutorial

MacVector, Inc.

Software for Scientists

Copyright statement

Contents

Introduction

MacVector provides a number of tools to help you design primers for use in Polymerase Chain Reaction (PCR) experiments. MacVector 12.6 introduced a new "Quicktest Primer" interface that simplifies the design of primers with mismatches and/or tails. This can be used in conjunction with the existing **Primer3** "Primer Design" functionality to design pairs of primers with mismatches and/or tails and to easily generate the predicted product of the reaction, complete with tails and/or mismatches. The Quicktest functionality was enhanced in MacVector 13 to display Restriction Sites around the primer location, including "one-out" sites and the effect of mutations on those sites on overlapping open reading frames.

This tutorial shows you how you can design a primer with a mismatch that changes a protein coding region to create a restriction enzyme site, add a tail containing a restriction enzyme site, find a matching primer to amplify a specific segment from a gene, add a restriction enzyme site tail to that primer, then finally generate a new DNA sequence representing the predicted product of the amplification.

Sample Files

This tutorial uses a sample file that is installed with MacVector. You can find the file at this location;

```
/Applications/MacVector/Tutorial Files/Align To
Reference/Sequence Confirmation/SequenceSample
```

Tutorial

Opening SequenceSample

> Select **File I Open** and navigate to the `/MacVector/Tutorial Files/Align To Reference/Sequence Confirmation/` folder. Select `SequenceSample` and click **Open**. A window will open – click on the **Map** tab to view a graphical representation of the protein coding regions of the sequence.
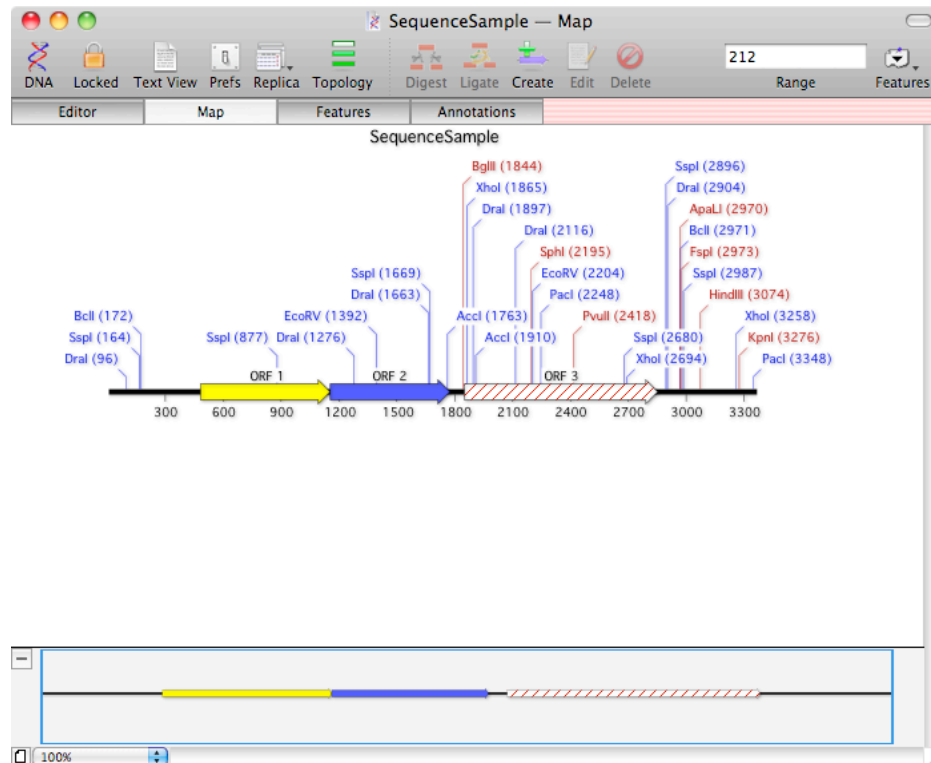
Note there are three ORFs annotated on the sequence. You can see the translations of the ORFs in the Editor tab as well.

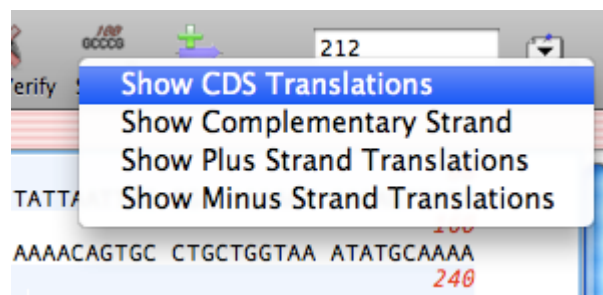Click on the **Editor** tab. Then click and hold on the **Display** button. A popup menu will appear. Select **Show CDS Translations.**

If you scroll through the sequence, you will see that the amino acid CDS translations are displayed over the sequence.

Invoking Quicktest Primer

Scroll through the sequence until you see the stop codon (*** or *) at the end of ORF 1 starting at residue 1152. Select approximately 20 residues across this codon.
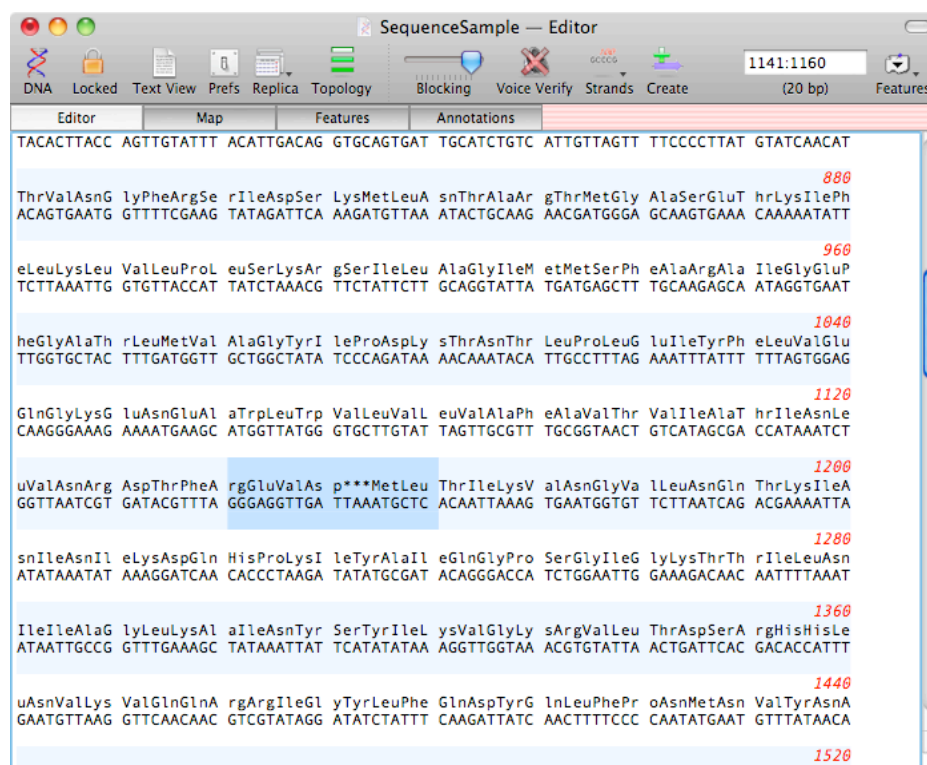
From the main menu, choose **Analyze | Primers | Quicktest Primer…**

The floating **Quicktest Primers** dialog will open;

When **Quicktest Primer** is invoked with a short (75 residues or less) selection in the parental sequence window, the selected sequence is automatically displayed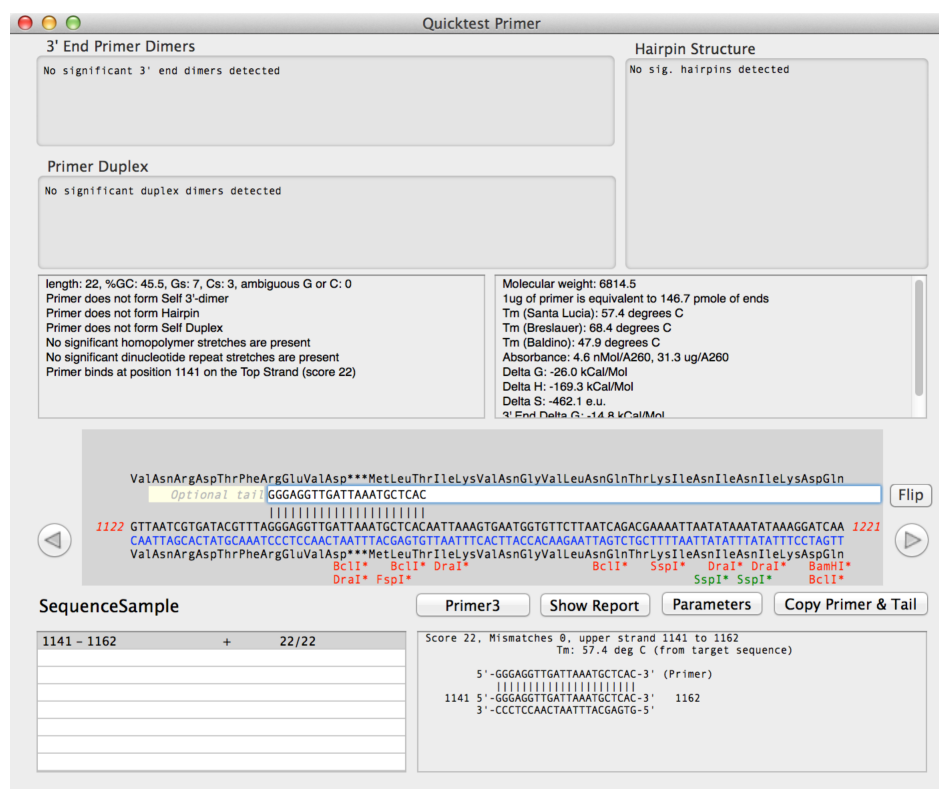 in the central primer edit box (with the white background and the focus ring). The dialog is sensitive to the current active nucleic acid window so that if the primer in the edit box binds to the active nucleic acid sequence, the matching sequence is shown aligned below the primer. In addition, the translations of any CDS features that overlap the binding region are shown below the sequence. Note that the plus and minus strands of the binding region are shown in the current colors (the defaults are black for the plus strand and blue for the minus strand). If the primer binds preferentially to the minus strand, the minus strand will be shown above the plus strand.

Finally, restriction enzyme sites are shown below the template sequence with color coding. Black sites are fully-matching restriction enzyme sites in the template sequence. The sites with an asterisk (e.g. BclI*) are "one-out" sites (e.g. a 5 out of 6 match to the recognition sequence) where a single residue change could be made to create a complete recognition site. Green sites are those where changing the template sequence would NOT change the amino acid sequence of any overlapping protein coding sequence. These will also include sites where no coding sequence is annotated on the sequence. Red sites are those where introducing the residue change to match the complete site WOULD cause a change to the amino acid sequence of a protein coding open reading frame. For this feature to work as expected, your source sequence must be annotated with CDS features representing coding regions. All sequences downloaded from the GenBank database will be annotated appropriately.

Resizing

There is a lot of information displayed in the **Quicktest Primer** dialog and sometimes there will be far more restriction enzymes displayed below the template sequence than can be viewed in the existing box. However, if you have a reasonably large screen, you can resize the dialog and the central primer/template editing pane will increase in size;
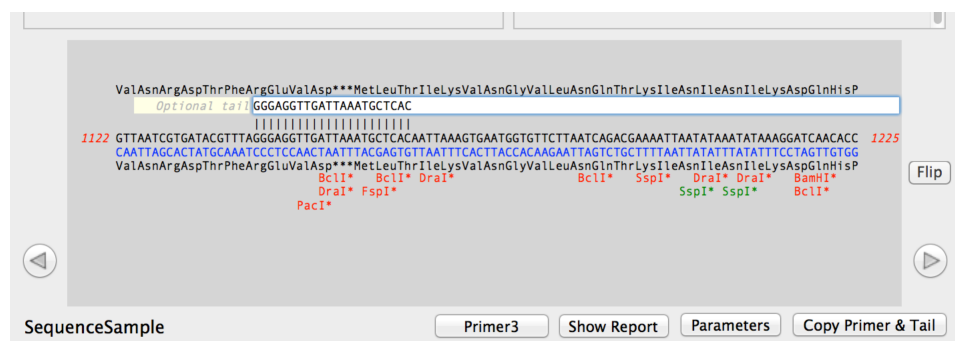
Changing The Amino Acid Display

The default setting is for amino acids to display using the single letter codes rather than triplet codes (e.g. " A " rather than "Ala"). This tutorial uses the three letter codes as it is more obvious to discuss glutamine as "Gln" rather than "Q". If you want to change you settings to match, choose the **MacVector I Preferences…** menu item and switch to the **Text View** tab;
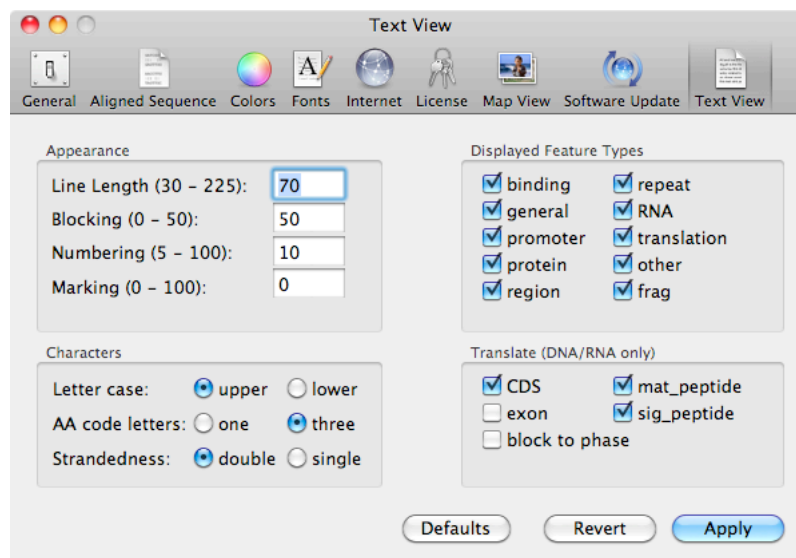
Make sure you have **AA code letters** set to **three**, and then click **Apply**.

> Open an unrelated nucleic acid sequence – e.g. `/MacVector/Sample Files/pBR322`.

When the unrelated sequence window opens, the **Quicktest Primer** dialog refreshes and the binding information is removed (because the primer does not bind).

If you click back on the `SequenceSample` window, the binding information reappears. You can use this behavior to quickly see if/where any primer binds to any open DNA sequence.

Nudging The Primer

When a primer binds to the target sequence, you can nudge it left and right to find the optimal location for the primer.

Make sure you have `SequenceSample` as the active window behind the **Quicktest Primer** dialog. Click on the round left/right buttons and watch what happens to the display.

Each time you click on a button, the primer gets "nudged" along the sequence. Each of the display panes dynamically responds to the new primer location, so when you nudge you can see any hairpin loops or primer dimers graphically displayed in the upper panes. The scrollable text boxes are also updated to report the binding and secondary structural information and the primer statistics (Tm, thermodynamic properties etc). If possible, you want to avoid primers that might form hairpin loops or primer dimers. That may not always be possible, but at least this display alerts you to potential problems.

Interacting With Restriction Sites

The restriction enzyme display is highly dynamic and interactive.

First, make sure the primer is nudged so that the *** termination codon is approximately centered in the primer. Move the mouse pointer over the BclI* site that is below and slightly to the left of the *** codon.

While the pointer is hovering over the BclI* site, the site highlights and the cut site and recognition sequence are temporarily displayed on the template sequence.
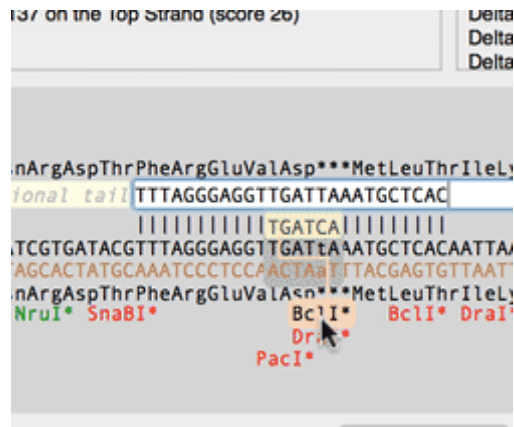
If you look carefully, you can see the mismatched residue is displayed in lower case. In this case, there is a "t" that would need to be changed to a "C" to create a functional BclI site.

Now click and hold the mouse button. This temporarily changes the sequence of the primer so that the "t" changes to a "C".
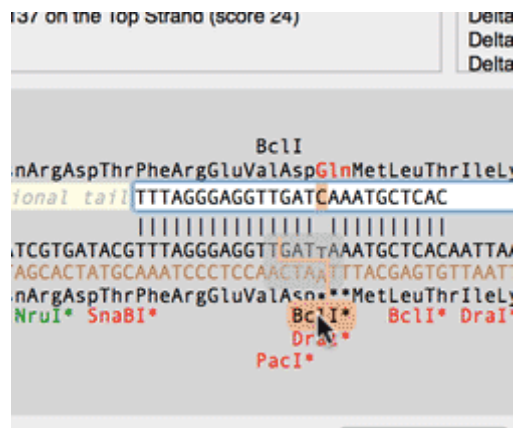
The primer sequence changes and there is now a mismatch between the primer and the template, represented by the gap in the vertical match characters between the sequences. Also, a BclI site has now appeared above the primer sequence. Restriction sites that would be created by changes in the primer are displayed in black text above the primer. Conversely, restriction sites that would be destroyed by mismatches in the primer are displayed in grey text. Any changes to the amino acid sequence of overlapping open reading frames are shown in red text. In this case the *** termination codon has been replaced by a red Gln codon.

To make the change permanent, simply double-click on the BclI* restriction site. In preparation for the next section of the tutorial, if you double-click, choose **Edit | Undo Edit** to revert the change.

Introducing a Mismatch Manually

> Nudge the primer so that the *** stop codon is approximately in the
> middle of the primer. Select the second "T" in the sequence
> ..GAT**T**AAA.. and type a "C".

The display updates to show that you have now introduced a mismatch and
also shows the change in the amino acid sequence in red;

The edit creates a new restriction enzyme site (*Bcl*I - TGAT*C*A) as we saw
above. You may also see that replacing the T with a C has introduced a
potentially significant internal primer duplex;
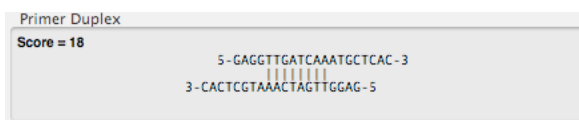
> Continue to nudge the primer. You'll see that as you nudge the primer to
> the left, the primer information is updated in real time. Unfortunately,
> because *Bcl*I has a typical 6 base pair dyad symmetrical recognition site,
> we can't do anything about the internal primer duplex.

Adding a Tail Containing a Restriction Enzyme Site

For the next step we will add a tail containing a *Bam*HI restriction site to the
primer.

> Click in the "Optional tail" box and type the sequence "CGCGGATCC"

Note that we add 3 additional residues (CGC) before the *Bam*HI recognition
sequence (GGATCC). Many restriction enzymes cannot cleave DNA if the
recognition sequence lies too closely to the end of a double stranded
molecule. Adding an additional 3 residues is usually sufficient to ensure that
the enzyme will be able to cut the final PCR product. As you type, the

restriction enzyme sites above the primer are automatically updated. You may find that adding the tail causes a potential hairpin loop to be created;

> Nudge the primer to the right. Eventually you will find a location where the hairpin loop is reduced to just 3 residues and it falls below the significance threshold. Do not let the mismatch fall off the end of the primer or you will have to re-enter the T -> C change.

Generating a Primer Report

Once we are happy with the primer and tail we have designed we may want to make a permanent record of the sequence of the primer and its characteristics. You can click on the **Copy Primer & Tail** button copy the full primer, ready to paste into other applications.

> Click on the **Show Report** button.

This opens a separate report window summarizing the information displayed in the **Quicktest Primer** dialog, including the primer properties, potential primer dimer and hairpin loops, the presence of any additional binding sites on the target sequence and the details of the primary binding site, restriction enzyme sites and CDS translations. The report is designed to be a simple one page document that you can print or save as a PDF file to create a permanent record of the primer.

Quicktest Primer Report

```
5' C GCG GAT CCG GGA GGT TGA TCA AAT GCT CAC AAT   3'
```

Primer Properties

Length	34	residues
GC	44.0	%
Gs: 7, Cs: 4, ambiguous G or C: 0		
Molecular weight	10492.0	
1ug of primer is equivalent to	129.4	pmole ends
Primer Tm (Santa Lucia)	61.8	degrees C
Primer+Tail Tm (Santa Lucia)	72.8	degrees C
Absorbance	3.0	nMol/A260
Absorbance	31.7	ug/A260
Delta G	-46.3	kCal/Mol
Delta H	-273.5	kCal/Mol
Delta S	-732.5	e.u.
3' End Delta G	-24.5	kCal/Mol

No significant homopolymer repeat stretches are present
No significant dinucleotide repeat stretches are present

Analysis Parameters

A-T score	2	
G-C score	3	
Mismatch score	-5	
3' end dimer score threshold	7	
Duplex score threshold	11	
Hairpin score threshold	8	
Hairpin minimum pairing	2	bases
Hairpin minimum loop size	3	bases
Hairpin maximum loop size	10	bases
Primer concentration	50.0	nM
Monovalent cation (Na+) conc.	50.0	mM
Divalent cation (Mg++) conc.	1.5	mM
dNTP concentration	0.6	mM

3' End Primer Dimers

```
No significant 3' end dimers detected
```

Primer Duplexes

Score = 22

```
        5-CGCGGATCCGGGAGGTTGATCAAATGCTCACAAT-3
                      |||||||||
3-TAACACTCGTAAACTAGTTGGAGGGCCTAGGCGC-5
```

Primer Hairpins

Score = 10

```
              GGG
             G    G
             G    T
             C    T
             C=G
             T=A
             A=T
             G=C
          5-CGCG    AAATGCTCACAAT-3
```

Primer Binding Details for SequenceSample

```
        BamHI         BclI
euValAsnArgAlaArgIleArgGluValAspGlnMetLeuThrIleLysValAsnGlyValLeuAsnGlnThrLysIleAsnIleAsnIleLysAspG
        CGCGGATCC GGGAGGTTGATCAAATGCTCACAAT
                  ||||||||||| ||||||||||||||
1120 TGGTTAATCGTGATACGTTTAGGGAGGTTGAT T AAATGCTCACAATTAAAGTGAATGGTGTTCTTAATCAGACGAAAATTAATATAAATATAAAGGATC  1219
     ACCAATTAGCACTATGCAAATCCCTCCAACTA A TTTACGAGTGTTAATTTCACTTACCACAAGAATTAGTCTGCTTTTAATTATATTTATATTTCCTAG
euValAsnArgAspThrPheArgGluValAsp * **MetLeuThrIleLysValAsnGlyValLeuAsnGlnThrLysIleAsnIleAsnIleLysAspG
                          BclI*   BclI* DraI*             BclI*  SspI*    DraI* DraI*   BamHI*
                          DraI* FspI*                                    SspI* SspI*     BclI*
                          PacI*
```

Alternate Binding Sites in SequenceSample

```
No alternate binding sites detected.
```

Finding a Suitable Matching Primer

Once the first primer has been designed, the next task is to find a suitable matching second primer for use in a PCR reaction. To do this we will make use of the integration between **Quicktest Primer** and the **Primer3** interface in MacVector.

Make sure you have the `SequenceSample` window as the front most window, then click on the **Primer3** button in the **Quicktest** dialog.
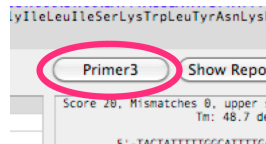
The Primer3 parameters sheet will drop down over the `SequenceSample` window and the **Quicktest Primer** dialog will close.

You can see that the main core of our primer is inserted into the 5' primer edit box and the tail is displayed over the top of that box. In addition, the **Use this Primer** radio button is selected for the 5' primer.

The trickiest part of using Primer3 to find a matching 3' primer is defining the region we want to be amplified. In this case, we know our designed primer overlaps the end of ORF1 and the start of ORF2. Lets set up the search to amplify from our primer position to somewhere beyond the end of ORF2.

> Make sure the popup menu is set to Amplify Feature/Region, then click
> on the region button ( ) and choose **1155 to 1775; CDS, ORF2** from
> the popup menu.

Because our primer overlaps the beginning of ORF2 MacVector will ensure that Primer3 accepts it as a valid primer. If the primer lay entirely within ORF2, we would need to adjust the **Region To Amplify** parameters so that the 3' end of the Region to Amplify would lie either within or to the right of the primer.

> Click on the **Advanced Options** button.

You will see the advanced options area displayed. Note how the length, percent G+C, Tm and GC clamp options have all been pre-filled to match the characteristics of the core of the input primer. Normally you should not

have to change these, as they will ensure that any 3' primers found will closely match your designed primer.

Click **OK** to run **Primer3**.

The Primer3 job will start and should complete within a few seconds. You should get a result sheet similar to this;

In particular, look at the **Pairs… Accepted** result. You should see that at least one pair has been accepted. If you don't see this then something went wrong – the most common problem is that you set up the region to amplify incorrectly, so that the left hand primer either did not bind within 200 residues of the region to amplify. If you click OK and examine the raw **Primer3 output** window it will usually have a message telling you what went wrong. If this is the case, simply choose **Analyze | Primers | Primer3…** and change the region to amplify settings.

Make sure the **Spreadsheet** and **Graphical map** checkboxes are selected, then click **OK**.

The `SequenceSample` **Primer3 Spreadsheet** tab displays a list of the pairs of primers found by Primer3.

Rank	L Start	L Len	L Tm	L GC%	L Sequence	R Start	R Len	R Tm	R GC%	R Sequence	P Size
1	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	2014	20	47.96	25.00	TCAACAATGGTTAACTTATT	867
2	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	1865	20	48.13	30.00	GTATCTTTCTTGCATTAGAT	718
3	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	1898	20	48.15	25.00	TAAACCAATTTCCTTAAACA	751
4	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	2037	20	47.84	25.00	TGCTAAATTCAATGTAATCT	890
5	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	1864	20	47.83	30.00	TATCTTTCTTGCATTAGATC	717
6	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	2007	20	47.82	30.00	TGGTTAACTTATTTATTCCC	860
7	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	1940	20	47.81	25.00	TACTATAAGCACATGTTTTT	793
8	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	1798	20	47.80	30.00	TAACAGATAAATAGCACTTC	651
9	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	2025	20	47.77	30.00	TGTAATCTCTATCAACAATG	878
10	1148	20	48.18	25.00	CGCGGATCCTGATCAAATGCTCACAATTA	2028	20	47.77	30.00	CAATGTAATCTCTATCAACA	881

You will see that all the left primers are the primer we designed using **Quicktest Primer**, complete with CGCGGATCC tail and the A->C mismatch. The Tm displayed (**L Tm**) is the Tm for the target site – i.e. the 20 residues where the tail-less core will bind. This does NOT take the mismatch into account. The right hand primers are all suitable matching primers with similar Tm's and length to our core starting primer. These are displayed as the "real" primer sequence i.e. the sequence you would send off for synthesis. They are the reverse complement of the plus strand of the target sequence. You can click in one of the primer cells and copy the primer to paste into an external application, or click elsewhere on a row to copy all of the data on the line so you have a record of the start, length and Tm data as well.

The `SequenceSample` **Primer3 Map** tab displays the results graphically (here shown zoomed in to the amplified region);
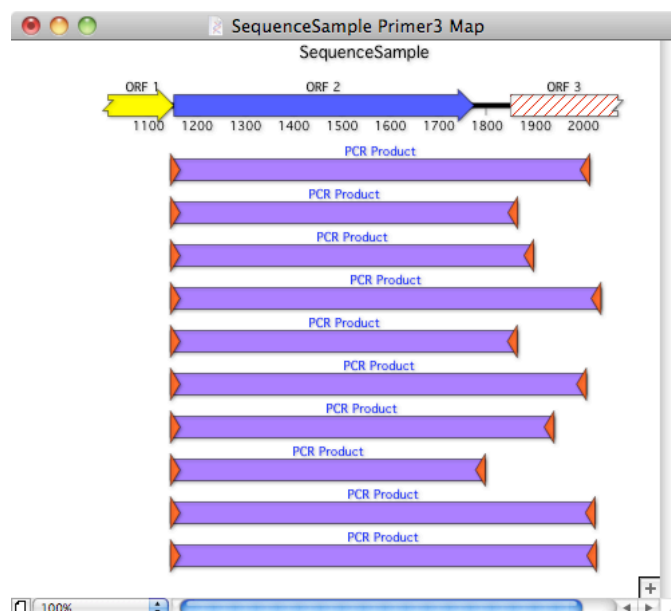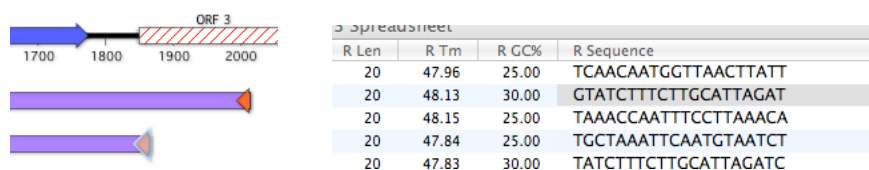

The **Map** is interactive. You can click on one of the small primer triangles, and then choose **Edit | Copy** and the sequence of the primer (including tail and/or mismatches) will be copied to the clipboard. Clicking on a purple

PCR Product object followed by **Edit I Copy** will copy the predicted PCR product, again including mismatches and or tails in each of the primers.

Adding a Tail to a Matching Primer

Once a suitable matching primer has been found, it is often useful to be able to add a tail to that primer as well. Again, we'd like to know if that might create significant secondary structure in the primer, so we need to use **Quicktest Primer** once more.

> In the `SequenceSample` **Primer3 Map** results tab, click on the right hand primer of the second pair. Note how the same primer highlights in the spreadsheet and also in the parental `SequenceSample` **Editor** window.

R Len	R Tm	R GC%	R Sequence
20	47.96	25.00	TCAACAATGGTTAACTTATT
20	48.13	30.00	GTATCTTTCTTGCATTAGAT
20	48.15	25.00	TAAACCAATTTCCTTAAACA
20	47.84	25.00	TGCTAAATTCAATGTAATCT
20	47.83	30.00	TATCTTTCTTGCATTAGATC

> Choose **Edit I Copy** to copy the primer sequence. Then choose **Analyze I Primers I Quicktest Primer** to open the **Quicktest Primer** dialog. If a primer is present in the main edit box, double-click in it to select the entire sequence, then choose **Edit I Paste** to replace it with the sequence you copied. Finally, if it is not already active, click on the original `SequenceSample` window to make sure **Quicktest** is reporting the binding results for the right sequence.

The **Quicktest Primer** dialog should appear and refresh with the new primer displayed in the primer edit box and with the template sequence "flipped" to show the correct binding between the new "minus strand" primer and the template. You should see that the upper strand of the template is now displayed in the "complementary strand" color ;
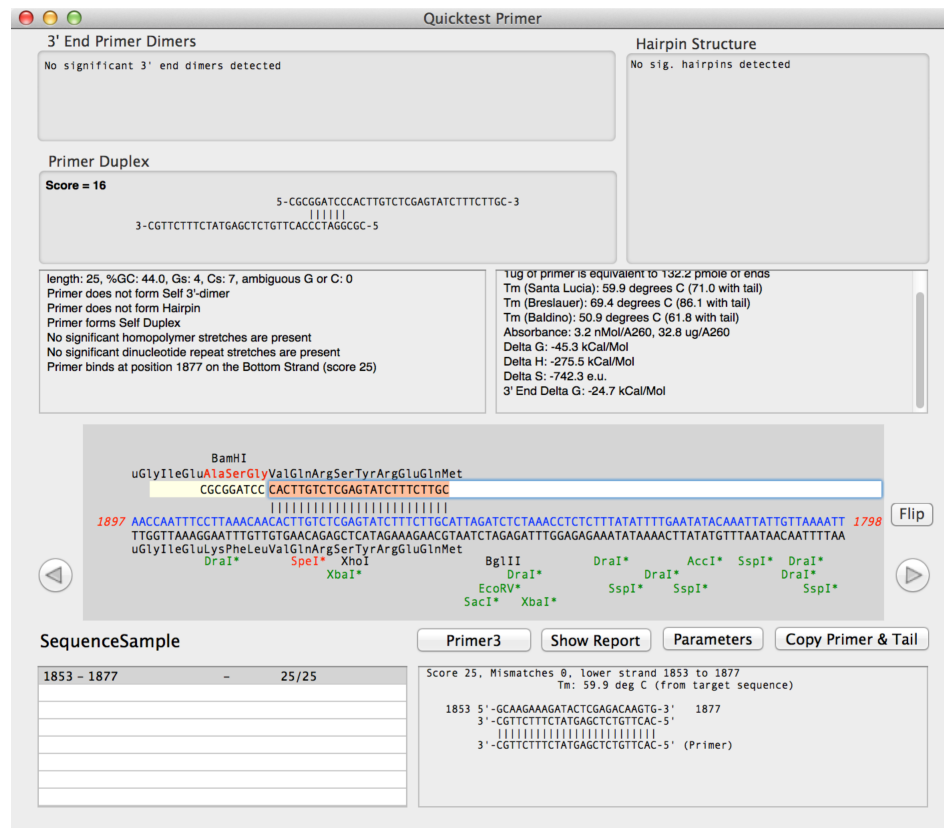
The "minus" strand is now shown above the black "plus" strand and the red numbering is larger at the left hand side than on the right hand side. This indicates that the primer is binding to the opposite strand, as you would expect for the second primer in a PCR pair. This is also indicated in the list box at the lower left corner that displays the binding sites in the sequence.

Usually, the tail sequence will be remembered between invocations of the **Quicktest Primer** dialog. But if it is not;

Click in the "Optional tail" edit box and type in the same *Bam*HI restriction site tail we used previously (CGCGGATCC).
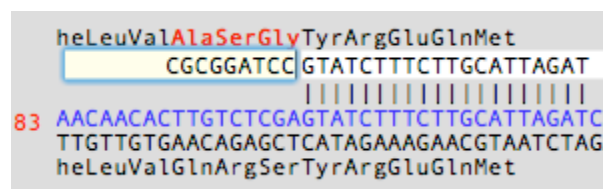
Finally, click on the **Primer3** button to reopen the **Primer3** sheet.

This time, because both primers have been defined (with tails), the Primer3 dialog will automatically switch into **Test Primer Pair** mode.

Note how the dialog remembers the last primer that we used for the 5' end, including the mismatch we introduced and the tail we added. In this case, the new primer is for the 3' end. MacVector always assumes that if you switch from **Quicktest** to **Primer3** with a primer that matches the minus strand (as we did here), then that should be treated as the 3' primer. In Test Primer Pair mode, the analysis bypasses Primer3 completely and simply tests the primers for binding and reports the primer properties and the predicted amplification products.

> Click **Test** to run the **Primer3** analysis

The analysis will proceed and you should be presented with a successful result dialog where just a single pair is considered and accepted.

> Click **OK** to display the result windows. Click on the single PCR Product graphical object in the **Map** tab and choose **Edit | Copy**. This will copy the predicted PCR product, including the mismatch and tail from the 5' primer and the tail we assigned to the 3' primer
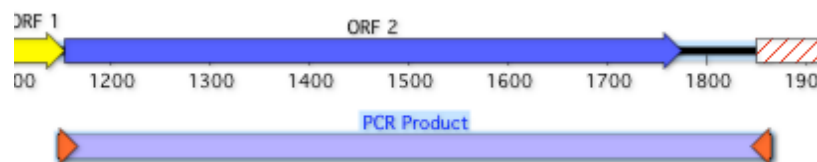
> Finally, choose **File | New From Clipboard** and switch to the **Map** tab of the newly created `Untitled` sequence.
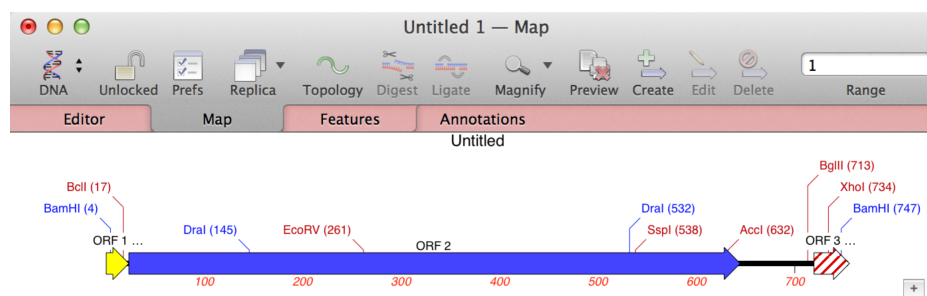
The new sequence retains all of the features from the source sequence i.e. ORF 2 and parts of ORF 1 and ORF 3. The sequence is now flanked by *Bam*HI sites from the tails we added to the primers. Finally, there is a *Bcl*I site near the left hand end – the mismatch we introduced into the 5' primer.

Cloning the Amplified Fragment into a Vector

In this example, after completing the PCR reaction you would typically digest the amplified product with *Bam*HI and then clone it into a suitable vector. This is simple to emulate in MacVector. For this example we will clone the fragment into the *Bam*HI site of classic early cloning vector pBR322.

> Click on the *Bam*HI site at the left side of the `Untitled 1` **Map**. Hold down the `<shift>` key and click on the *Bam*HI site at the right side of the **Map**. Click on the **Digest** toolbar button.

> Choose **File | Open** and open a suitable vector that contains a *Bam*HI cloning site. If you don't have a favorite vector, choose `/Applications/MacVector/Sample Files/pBR322.` Switch to the **Map** tab and click on the *Bam*HI site at position 375.
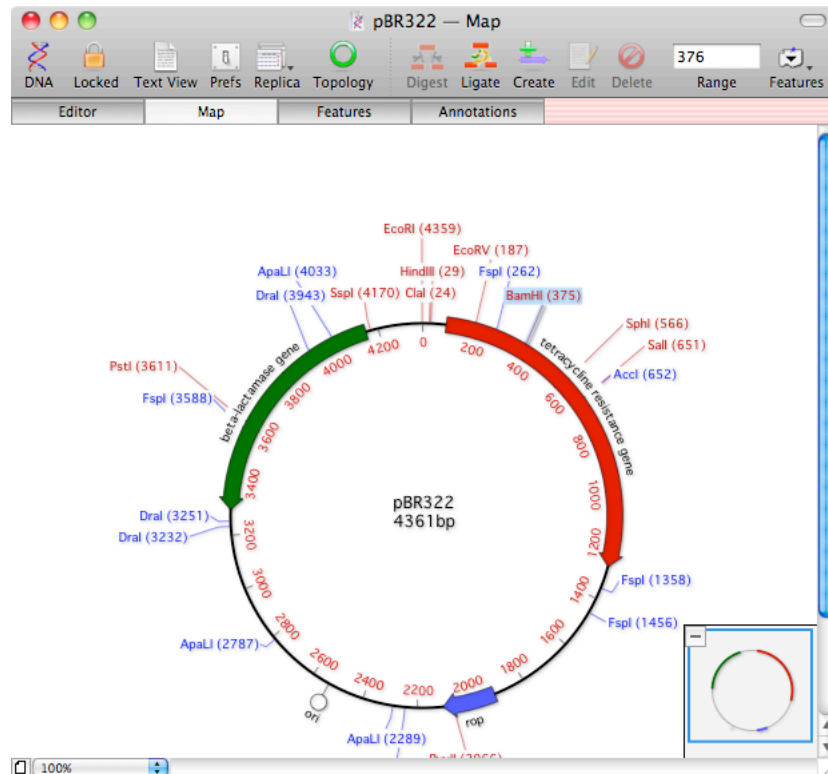
Click on the **Ligate** toolbar button.

After unlocking the document, the ligation sheet will appear;
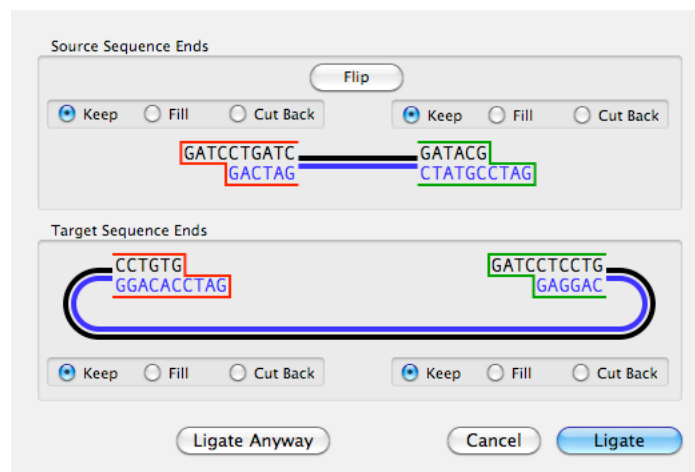
This gives you the opportunity to "flip" the incoming fragment so it is inserted in the opposite orientation, and also lets you fill or cut back sticky ends when emulating clonings that require manipulating overhangs. To learn more about "click cloning" with MacVector, check out the `/MacVector/Documentation/Click Cloning Tutorial.pdf` document.

Click on the **Ligate** button.

The PCR fragment with *Bam*HI ends gets inserted into the target pBR322 document, creating a new molecule complete with all the features of the original target DNA sequence and including the new *Bcl*I site we introduced with the mismatched primer.
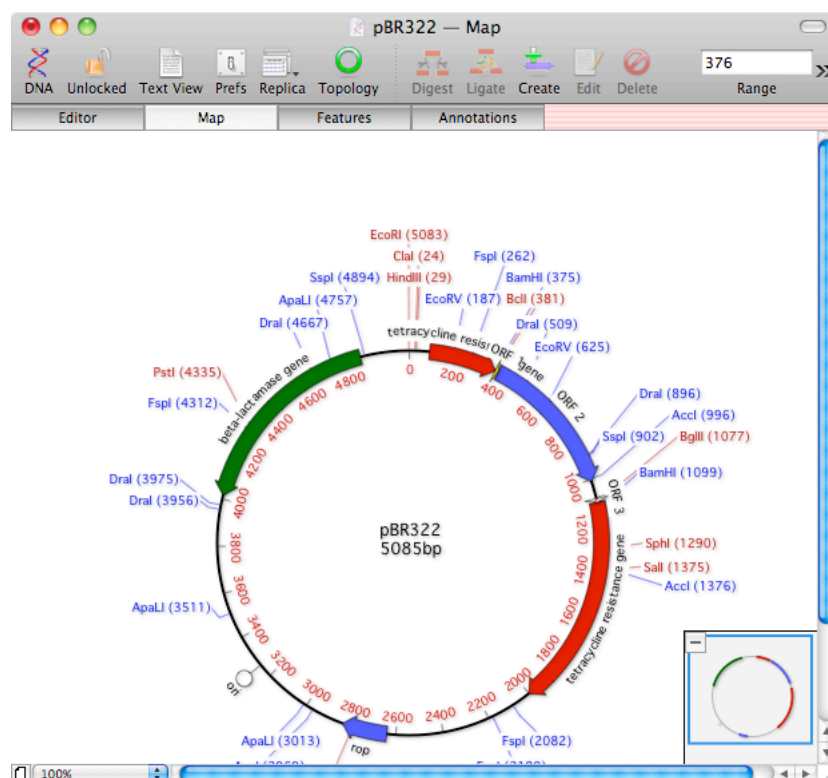
Additional Information

You can use the principles above to design any primer with or without tails or mismatches and document the final construct using MacVector's point and click interface. For more information on the "click cloning" interface, read the `Click Cloning Tutorial` in the `Documentation` folder. If you use T/A cloning to clone your PCR fragments, then you should read the `Gateway + TOPO Tutorial` document.