Preface

PCR is probably the single most important methodological invention in molecular biology to date. Since its conception in the mid-1980s, it has rapidly become a routine procedure in every molecular biology laboratory for identifying and manipulating genetic material, from cloning, sequencing, mutagenesis, to diagnostic research and genetic analysis. What's astounding about this invention is that new and innovative applications of PCR have been generated with stunning regularity; its potential has shown no signs of leveling off. New applications for PCR are literally transforming molecular biology. In the post-genomic era, PCR has especially become the method of choice to clone existing genes and generate a wide array of new genes by mutagenesis and/or recombination within the genes of interest. The fast and easy availability of these genes is essential for the study of functional genomics, gene expression, protein structure–function relationships, protein–protein interactions, protein engineering, and molecular evolution.

PCR Cloning Protocols was prepared in response to the need to have an up-to-date compilation of proven protocols for PCR cloning and mutagenesis. It builds upon the best-selling first edition, *PCR Cloning Protocols: From Molecular Cloning to Genetic Engineering*, a book in the Methods in Molecular Biology™ series published in 1997. We divided the new edition into five parts. Part I. **Performing and Optimizing PCR**, contains basic PCR methodology, including PCR optimization and computer programs for PCR primer design and analysis, as well as novel variations for cloning genes of particular characteristics or origins, emphasizing long-distance PCR and GC-rich template amplification. Part II. **Cloning PCR Products**, presents both conventional and novel enzyme-free and restriction site-free procedures to clone PCR products into various vectors, either directionally or non-directionally. Part III. **Mutagenesis and Recombination**, addresses the use of PCR to facilitate DNA mutagenesis and recombination in various innovative approaches to generate a wide array of mutants. Part IV. **Cloning Unknown Neighboring DNA**, contains a comprehensive collection of protocols to fulfill the frequent and challenging task of cloning uncharacterized DNA flanking a known DNA fragment. Finally, Part V. **Library Construction and Screening**, addresses particular applications of PCR in library and sublibrary generation and screening. Each part also contains an overview, which summarizes the current methods available and their underlying

strategies, advantages, and disadvantages for that particular topic. These reviews are especially helpful to new researchers to orient themselves with the field and to guide them to choose a procedure that is most suitable for their experiments.

We hope that *PCR Cloning Protocols* will provide readily reproducible laboratory protocols that researchers in the field will follow closely and thereby increase their success rate in their experiments.

We are indebted to Mirah Riben for her superb help during the editing of the book. We also thank Prof. John M. Walker, the series editor, for his help, advice, and guidance.

Bing-Yuan Chen

Harry W. Janes

2

Computer Programs for PCR Primer Design and Analysis

Bing-Yuan Chen, Harry W. Janes, and Steve Chen

1. Introduction

1.1. Core Parameters in Primer Design

1.1.1. T_m, Primer Length, and GC Content (GC %)

Heat will separate or "melt" double-stranded DNA into single-stranded DNA by disrupting its hydrogen bonds. T_m (melting temperature) is the temperature at which half the DNA strands are single-stranded and half are double-stranded. T_m characterizes the stability of the DNA hybrid formed between an oligonucleotide and its complementary strand and therefore is a core parameter in primer design. It is affected by primer length , primer sequence, salt concentration, primer concentration, and the presence of denaturants (such as formamide or DMSO).

All other conditions set, T_m is characteristic of the primer composition. Primer with higher G+C content (GC %) has a higher T_m because of more hydrogen bonds (three hydrogen bonds between G and C, but two between A and T). The T_m of a primer also increases with its length. A simple formula for calculation of the T_m ***(1,2)*** (*see* **Note 1**) is

$$T_m = 2 \times AT + 4 \times CG$$

where *AT* is the sum of A and T nucleotides, and *CG* is the sum of C and G nucleotides in the primer.

1.1.2. Primer Specificity

Primer specificity is another important parameter in PCR primer design. To amplify only the intended fragment, the primers should bind to the target sequence only but not somewhere else. In other words, the target sequence should occur only once in the template. Primer length not only affects the T_m, as discussed earlier, but also the uniqueness (specificity) of the sequence in the template ***(3)***. Suppose the DNA sequence is entirely random (which may not be true), the chance of finding an A, G, C,

From: *Methods in Molecular Biology, Vol. 192: PCR Cloning Protocols, 2nd Edition*
Edited by: B.-Y. Chen and H. W. Janes © Humana Press Inc., Totowa, NJ

or T in any given DNA sequence is one quarter ($1/4^1$), so a 16 base primer will statistically occur only once in every 4^{16} bases, or about 4 billion bases, which is about the size of the human genome. Therefore, the binding of a 16 base or longer primer with its target sequence is an extremely sequence-specific process. Of course, to be absolutely sure that the target sequence occurs only once, you would need to check the entire sequence of the template DNA, which is not possible in most cases. However, it is often useful to search the current DNA sequence databases to check if the chosen primer has gross homology with repetitive sequences or with other loci elsewhere in the genome. For genomic DNA amplification 17-mer or longer primers are routinely used.

1.1.3. Primer Sequence and Hairpin (Self-Complementarity) and Self-Dimer (Dimer Formation)

The hardest part in PCR primer design is to avoid primer complementarity, especially at the 3' ends. When part of a primer is complementary to another part of itself, the primer may fold in half and form a so-called hairpin structure, which is stabilized by the complementary base pairing. The hairpin structure is a problem for PCR because the primer is interacting with itself and is not available for the desired reaction. Furthermore, the primer molecule could be extended by DNA polymerase so that its sequence is changed and it is no longer capable of binding to the target site.

Similar to the hairpin structure, if not carefully designed, one primer molecule may hybridize to another primer molecule and acts as template for each other, resulting in primer-dimers. Primer-dimer formation causes the same problems to PCR reaction as the hairpin structure. It may also act as a competitor to amplification of the target DNA ***(4)***. Usually it is very hard and time-consuming to catch the hairpin structure or primer-dimer formation manually by a naked eye. However, they can be easily detected by primer analysis programs.

1.2. General Rules for PCR Primer Design

According to Innis and Gelfand ***(5)*** the rules for primer design is as follows:

1. Primers should be 17–28 bases in length;
2. Base composition should be 50–60% (G+C);
3. Primers should end (3') in a G or C, or CG or GC: this prevents "breathing" of ends and increases efficiency of priming;
4. Tms between 55–80°C are preferred;
5. Avoid primers with 3' complementarity (results in primer-dimers). 3'-ends of primers should not be complementary (i.e., basepair), as otherwise primer dimers will be synthesised preferentially to any other product;
6. Primer self-complementarity (ability to form secondary structures such as hairpins) should be avoided;
7. Runs of three or more Cs or Gs at the 3'-ends of primers may promote mispriming at G or C-rich sequences (because of stability of annealing), and should be avoided.

Because two different primers are needed for PCR reaction, primer-dimer formation between the two primers should also be checked and avoided if possible. It is desirable that primer T_ms should be similar (within 8°C or so). If they are too different, a suitable annealing temperature may be hard to find. At high annealing temperature,

the primer with the lower T_m may not work, whereas at low annealing temperature, amplification will be less efficient because the primer with the higher T_m will misprime.

In reality, primer selection is often empirical. It varies greatly from researcher to researcher in regard to the criteria they use.

1.3. Computer Programs for PCR Primer Design and Analysis

1.3.1. Computer Programs for Nondegenerate PCR Primer Design

For primer design, most researchers used to visually inspect target DNA sequence to find primer(s) with the characteristics they prefer, which are usually similar to the guidelines we mentioned earlier. As computers are widely used in molecular biology, a large number of computer programs have been specifically developed for nondegenerate primer selection, which makes the PCR primer design more efficient and reliable. Most sequencing analysis packages, such as Vector NTI (InforMax Inc.), usually contain a primer design module. In this chapter, we focus on free online (web) primer design programs (*see* **Note 2**). Selected computer programs for nondegenerate PCR primer design and their features are listed in **Table 1**.

From a computational point of view the design of nondegenerate PCR primers is relatively simple: find short substrings from DNA nucleotide string that meet certain criteria. Although the criteria vary between programs, the core parameters, such as the primer length, T_m, GC content, and self-complementarity, are shared by these programs.

1.3.2. Computer Programs for Degenerate PCR Primer Design

In the experiments to amplify the novel members of gene families or cognate sequences from different organisms by PCR, the exact sequence of the target gene is not known. We usually align all known sequences for this gene and find the most conserved regions, then design corresponding "degenerate" primers, which are a set of primers with nucleotide diversity at several positions in the sequence. Degeneracies obviously increase the chances of amplifying the target sequence but reduce the specificity of the primer(s) at the same time.

Designing degenerate primers has been considered more of an art than a science. There are much less computer programs for degenerate primer design (*see* **Table 2**) than for nondegenerate primer design.

1.3.3. Computer Programs for Primer Analysis

Even if you prefer to design primers by yourself, not by a computer program, it is advised that your primers should be analyzed by a computer program to determine T_m, possible hairpin structure, primer-dimers, and other properties before you place the order for them. **Table 3** lists two computer programs for this purpose.

2. Materials

1. Computer: A computer (PC or Macintosh) with high-speed internet access.
2. Programs: Web Browser, Netscape (5.0 or above) or Internet Explorer (4.0 or higher).
3. Input files for primer design: DNA sequence file DNA.txt (*see* **Table 4**) and protein sequence file Protein.txt (*see* **Table 5**) (*see* **Note 4**).

Table 1
Selected Computer Programs for Nondegenerate PCR Primer Design

Program	Operating System	Features	URL (*see* **Note 3**)
Oligos	Windows 9X/NT	Free download The program includes several tools: make complement, reverse complement and inverted strand; search the sequence; extract from selected sites. (Reference Lowe T 1990)	http://www.biocenter.-helsinki.fi/bi/bare-1_html/oligos.htm
GCG Prime	Unix	Commercial Available within GCG This program selects primers according to a number of user-specified criteria including length, GC content, and annealing temperature. Potential primers can also be tested for self-complementarity and complementarity to each other to minimize the formation of primer dimers during the PCR.	http://www.gcg.com/products/wis-pkg-programs.html#Primer
Primer3	Internet Browser	Free Lots of user-configurable parameters Primer design for both PCR and hybridization Nice interface with useful help pages	http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi
Web Primer	Internet Browser	Free Best for designing primers to clone yeast genes. Can use a standard yeast gene name or systematic yeast name as DNA source input	http://genome-www2.stanford.edu/cgi-bin/SGD/web-primer
iOligo	Windows 9X/NT, Mac	Commercial Retrieval of Sequences from NCBI Sequence Editor Analysis of Oligonucleotide's Characteristics Submission of Oligo Orders by email	http://www.caesar software.com/pages/products/ioligo/ioligo.shtml
xprimer	Internet Browser	Free The user can select repeat database and genome model. Nice graphical display of suggested primers	http://alces.med.umn.edu/webprimers.html
PCR Help!	Windows 9X/NT	Commercial Free Demo Download User-friendly "PCR Wizard" allows you to design primers to any given DNA template sequence as well as to generate a Techne Genius Thermal Cycler program file, which can be sent from a PC directly to multiple Genius thermal cyclers (up to 32)	http://www.techneuk.co.uk/CatMol/pcrhelp.htm
Oligo	Windows 9X/NT, Mac	Commercial Free Demo Download Nice graphical interface for searching, selecting, and analyzing primers from known sequences Cross-compatible Multiplex PCR Primer Search Priming Efficiency Calculations	http://www.oligo.net/
The Primer Generator	Internet Browser	Free Designs Site Directed Mutagenesis primers The program analyzes the original nucleotide sequence and desired amino acid sequence and designs a primer that either has a new restriction enzyme site or is missing an old one. This allows for faster sorting out of mutated and nonmutated sequences.	http://www.med.jhu.edu/medcenter/primer/primer.cgi

Table 2
Selected Computer Programs for Degenerate PCR Primer Design

Program	Operating System	Features	URL
GeneFisher	Internet Browser	Free Processes aligned or unaligned sequences of DNA or protein	http://bibiserv.techfak.uni-bielefeld.de/genefisher/
CODEHOP	Internet Browser	Free Design degenerate PCR primers from protein multiple sequence alignments. The multiple-sequence alignments should be of amino acid sequences of the proteins and be in the Blocks Database format	http://www.blocks.fhcrc.org/codehop.html
Primer Premier 5	Mac and Windows 9X/NT	Commercial Free Demo Download Reverse translate a protein sequence and design primers in regions of low degeneracy	http://www.premierbiosoft http://www.premierbiosoft com/primerdesign/primerdesign.html

Table 3
Selected Computer Programs for PCR Primer Analysis

Program	Operating System	Features	URL
Oligo Analyzer	Internet Browser	Free Caculate Tm, find possible primer hairpin structure and primer dimer formation, Blast search databases for primer homologs	http://playground.idtdna.com/program/oligocalc/oligocalc.asp
NetPrimer	Internet Browser	Free Java Applet Analyze basic properties and second structures for an individual primer or primer pair. Also give a primer rating and a report of the analysis results	http://www.premierbiosoft.com/netprimer/netprimer.html

3. Methods

3.1. Designing Nondegenerate PCR Primers Using Primer3

Primer3 was developed at Whitehead Institute for Biomedical Research and Howard Hughes Medical Institute. It contains so many parameters that most people only need a subset of them to use as the criteria for primer selection.

3.1.1. Design Primers with the Default Settings

Primer3 provides default values for core parameters (*see* **Table 6** for a selected list. Go to Primer3 web page for a complete list and their meanings). If these default settings meet your needs, then use the following method to select your primers.

Table 4
Input File DNA.txt

```
   1 GGGGAAGTGC AATCACACTC TACCACACAC TCTCTATAGT ATCTATAGTT GAGAGCAAGC
  61 TTTGTTAACA ATGGCGGCTT CCATTGGAGC CTTAAAATCT TCACCTTCTT CCCACAATTG
 121 CATCAATGAG AGAAGAAATG ATTCTACACG TGCAATATCC AGCAGAAATC TCTCATTTTC
 181 GTCTTCTCAT CTCGCCGGAG ACAAGTTGAT GCCTGTATCG TCCTTACGTT CCCAAGGAGT
 241 ACGATTCAAT GTGAGAAGAA GTCCATTGAT TGTGTCTCCT AAGGCTGTTT CTGATTCGCA
 301 GAATTCACAG ACATGTCTGG ATCCAGATGC TAGCAGGAGT GTTTTGGGAA TTATTCTTGG
 361 AGGTGGAGCT GGGACCCGAC TTTATCCTCT AACTAAAAAA AGAGCAAAAC CTGCGGTTCC
 421 ACTTGGAGCA AATTATCGTC TGATTGACAT TCCCGTAAGC AATTGCTTGA ACAGTAACAT
 481 ATCCAAGATC TATGTTCTCA CACAATTCAA CTCTGCCTCT CTAAATCGCC ACCTTTCACG
 541 GGCATATGCT AGCAATATGG GAGAATACAA AAACGAGGGC TTTGTGGAAG TTCTTGCTGC
 601 TCAACAAAGT CCGGAGAACC CCGATTGGTT CCAGGGCACT GCGGACGCTG TCAGACAATA
 661 TCTGTGGTTG TTTGAGGAGC ATAATGTTCT TGAATACCTT ATACTTGCTG GAGATCATCT
 721 GTATCGAATG GATTATGAAA AGTTTATTCA AGCCCACAGG GAAACAGATG CTGATATTAC
 781 TGTTGCCGCA CTGCCAATGG ACGAGAAGCG TGCCACTGCA TTCGGTCTCA TGAAGATTGA
 841 CGAAGAAGGA CGCATTATTG AATTTGCAGA GAAACCGCAA GGAGAGCAAC TGCAAGCAAT
 901 GAAAGTGGAT ACTACCATTT TAGGTCTTGA TGACAAGAGA GCTAAAGAAA TGCCTTTTAT
 961 CGCCAGTATG GGTATATATG TCATTAGCAA AGACGTGATG TTAAACCTAC TTCGTGACAA
1021 GTTCCCTGGG GCCAATGATT TTGGTAGTGA AGTTATTCCT GGTGCAACTT CACTTGGGAT
1081 GAGAGTGCAA GCTTATTTAT ATGATGGGTA CTGGGAAGAT ATTGGTACCA TTGAAGCTTT
1141 CTACAATGCC AATTTGGGCA TTACAAAAAA GCCGGTGCCA GATTTTAGCT TTTACGACCG
1201 ATCAGCCCCA ATCTACACCC AACCTCGATA TTTGCCACCT TCAAAAATGC TTGATGCCGA
1261 TGTCACAGAT AGTGTCATTG GTGAAGGTTG TGTGATCAAG AACTGTAAGA TTCACCATTC
1321 CGTGGTTGGG CTCAGATCAT GCATATCAGA GGGAGCAATT ATAGAAGACT CACTTTTGAT
1381 GGGGGCAGAT TACTACGAGA CTGATGCTGA GAGGAAGCTG CTGGCTGCAA AGGGCAGTGT
1441 CCCAATTGGC ATCGGCAAGA ATTGTCTATA CAAAAGAGCC ATTATCGACA AGAATGCTCG
1501 TATAGGGGAC AATGTGAAGA TCATTAACAA AGACAATGTT CAAGAAGCGG CTAGGGAAAC
1561 AGATGGATAC TTCATCAAGA GTGGGATCGT CACTGTCATC AAGGATGCTT TGATTCCAAG
1621 TGGAATCGTC ATTTAAAGGA ACGCATTATA ACTTGGTTGC CCTCCAAGAT TTTGGCTAAA
1681 CAGCCATGAG GTACAAACGT GCCGAAGTTT TATTTTCCTA TGCTGTAGAA ATCTAGTGTA
1741 CATCTTGCTT TTATGATACT TCTCATTACC TGGTTGCTGT AAAAATTATT CGTCTAAAAT
1801 AAAAATAAAT CTACCATTAC ACCA
```

1. Start a web browser (Netscape or Internet Explorer).
2. Replace the default URL address with http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi and hit return. After connection Primer3 web page will appear in your browser.
3. Open the DNA sequence input file DNA.txt (*see* **Note 5**) using your favorite text editor, such as Notepad in Windows, then copy the sequence by going to **Edit/Select All**, **Edit/Copy** in the menubar. Close file DNA.txt.
4. In your browser click on the top sequence input box, then paste the above sequence by going to **Edit/Paste** in the menubar.
5. Click **Pick Primers** Button (there are six Pick Primers buttons on the page. Click any one of them will do the same). After a few s/min, Primer3 Output will be returned. The top part of the output is shown in **Table 7**.

The other parts of the output not shown are: whole input sequence and arrows, which nicely indicate the location of the primers above; additional four primer pairs; and statistics about the primer selection process.

Table 5
Input File Protein.txt

```
>Protein1
MKSTVHLGRVSTGGFNNGEKEIFGEKIRGSLNNNLRINQLSKSL
KLEKKIKPGVAYSVITTENDTETVFVDMPRLERRRANPKDVAAVILGGGEGTKLFPLT
SRTATPAVPVGGCYRLIDIPMSNCINSAINKIFVLTQYNSAALNRHIARTYFGNGVSF
GDGFVEVLAATQTPGEAGKKWFQGTADAVRKFIWVFEDAKNKNIENILVLSGDHLYRM
DYMELVQNHIDRNADITLSCAPAEDSRASDFGLVKIDSRGRVVQFAENQRFELKAMLV
DTSLVGLSPQDAKKSPYIASMGVYVFKTDVLLKLLKWSYPTSNDFGSEIIPAAIDDYN
VQAYIFKDYWEDIGTIKSFYNASLALTQEFPEFQFYDPKTPFYTSPRFLPPTKIDNCK
IKDAIISHGCFLRDCTVEHSIVGERSRLDCGVELKDTFMMGADYYQTESEIASLLAEG
KVPIGIGENTKIRKCIIDKNAKIGKNVSIINKDGVQEADRPEEGFYIRSGIIISEKA
TIRDGTVI
>Protein2
MDALCAGTAQSVAICNQESTFWGQKISGRRLINKGFGVRWCKSF
TTQQRGKNVTSAVLTRDINKEMLPFENSMFEEQPTAEPKAVASVILGGGVGTRLFPLT
SRRAKPAVPIGGCYRVIDVPMSNCINSGIRKIFILTQFNSFSLNRHLARTYNFGNGVG
FGDGFVEVLAATQTPGDAGKMWFQGTADAVRQFIWVFENQKNKNVEHIIILSGDHLYR
MNYMDFVQKHIDANADITVSCVPMDDGRASDFGLMKIDETGRIIQFVEKPKGPALKAM
QVDTSILGLSEQEASNFPYIASMGVYVFKTDVLLNLLKSAYPSCNDFGSEIIPSAVKD
HNVQAYLFNDYWEDIGTVKSFFDANLALTKQPPKFDFNDPKTPFYTSARFLPPTKVDK
SRIVDAIISHGCFLRECNIQHSIVGVRSRLDYGVEFKDTMMMGADYYQTESEIASLLA
EGKVPIGVGPNTKIQKCIIDKNAKIGKDVVILNKQGVEEADRSAEGFYIRSGITVIMK
NATIKDGTVI
```

Table 6
Selected Default Settings for Primer3

Parameter	Minimum	Optimum	Maximum
Primer size (base pairs)	18	20	27
Primer T_m (°C)	57	60	63
Max T_m Difference (°C)			100
Primer GC%	20		80
Product size (basepairs)	100	200	1000

Table 7
Primer3 Output

```
WARNING: Numbers in input sequence were deleted.

No mispriming library specified
Using 1-based sequence positions
OLIGO      start  len  tm     gc%    any    3' seq
LEFT PRIMER  890   20  59.99  45.00  4.00   0.00  ctgcaagcaatgaaagtgga
RIGHT PRIMER 1090  20  59.83  50.00  4.00   2.00  ttgcactctcatcccaagtg
SEQUENCE SIZE: 1824
INCLUDED REGION SIZE: 1824

PRODUCT SIZE: 201, PAIR ANY COMPL: 7.00, PAIR 3' COMPL: 3.00
```

3.1.2. Design Primers with User-Defined Settings

Often the default values need to be altered because they do not meet a researcher's needs or Primer3 did not find an appropriate PCR primer pair. The following are helpful guidelines for adjusting these parameters if Primer3 failed to select a primer:

a. Adjust location: pick a wider range to examine and allow for longer product size;
b. Change primer size: usually easier to find compatible primers if they are shorter;
c. Lower primer T_m.

Because there are so many configurable parameters in Primer3, it is impossible to explain their uses and try to change them here. Fortunately, the default values need not to be altered for most parameters. The readers should read the Primer3 help page and understand the uses of the parameters before trying to change them.

In the following method, we will try to design primers to clone the coding region in DNA.txt, which is from nucleotide 71 to 1636.

1. Start a web browser (Netscape or Internet Explorer).
2. Replace the default URL address with http://www-genome.wi.mit.edu/cgibin/primer/primer3_www.cgi and hit return. After connection, Primer3 web page will appear in your browser.
3. Open the DNA sequence input file DNA.txt using your favorite text editor, such as Notepad in Windows, then copy the sequence by going to **Edit/Select All**, **Edit/Copy** in the menubar. Close file DNA.txt.
4. In your browser, click on the top sequence input box, then paste the above sequence by going to **Edit/Paste** in the menubar.
5. Type 71,1565 in the **Targets** input box. Change **Product Size/Max** from 1000 to 1824, then click **Pick Primers** button. (There are six Pick Primers buttons on the page. Click any one of them will do the same). After a few s/min, Primer3 Output will be returned. Are there any primers returned?

3.2. Designing Degenerate PCR Primers Using GeneFisher

GeneFisher is an interactive degenerate primer design software. The current version, GeneFisher 1.22, processes aligned or unaligned sequences of DNA or protein. In the following method, we will use two unaligned protein sequences as the sequence input and design degenerate primers which could amplify the cDNAs encoding these two proteins and their related family members (if any).

1. Start web browser (*see* **Note 6**).
2. Replace the default URL address with http://bibiserv.techfak.unibielefeld.de/genefisher/ and hit return. After connection, **GeneFisher Interactive PCR Primer Design** home page will appear in your browser.
3. Click **Start** button on the page. After a few s/min, the Interactive **Primer Design** interface will be open.
4. In the **User Data** area of the page, type your **E-mail ID** and **Project** name. Click **this** button in the **Sequence Data** area to clear the sample sequence, then copy the two protein sequences from Protein.txt and paste to the **Sequence Data** area (*see* **Note 7**). Click **OK** button in the **Submit Query** area to accept your choice.

5. After a few s/min **GeneFisher Sequence Input** page will appear. Check that the protein lengths match with the input sequences. Click **OK** button to accept the two protein sequences.
6. After **GeneFisher Alignment Status** page returns, click **OK** button to use ClustalW as the alignment tool. ClustalW Multiple Sequence Alignment Setup page will appear. Click **OK** button to accept the default parameters.
7. Click **Progress** button on the **GeneFisher Clustal Alignment** page, which will open a new browser window. Click **Reload** button repeatedly in the new window to check the status of the alignment. If the last line on the page shows "GDE-Alignment file created" (The alignment time depends on your input. It takes several minutes for Protein.txt.) Then click **Alignment** button in the original window, which will show the alignment results in a new window.
8. We are satisfied with the alignment results, so go to the original window and click the **Consensus** button. **Sequence Consensus** page will return. Click **OK** button to accept the default consensus parameters, which will open the **GeneFisher Consensus** page.
9. Click Progress button to check the consensus calculation progress. If you are satisfied with the consensus calculation, click **Consensus** button on the **GeneFisher Consensus** page, which will show the alignment results in a new window. Click **Go!** button in the original window to generate primers.
10. After a few s/min, **Primer Design** page will appear. Click **OK** button to accept the default settings for primer design, which will open **GeneFisher Primer Calculation** page. Wait a few s/min, then click **Results** button, which will open the **Primer Calculation Results** page. Unfortunately, the results show that no primer pairs were generated. The rejection statistics underneath give some clues on why the primer selection fails. Click the **Redo** button to return to the **Primer Design** page.
11. Make the following changes to the primer parameters:
    a. Set primer length from 15 to 22 bp.
    b. Set GC content from 35 to 85%.
    c. Set melting temperature T_m from 42 to 65°C.
    d. Set product size from 100 to 1500 bp.
    e. Set primer degeneracy 512-fold.
    f. Set 3' GC content from 35 to 85%.

    Repeat the primer design step above **(step 10)**. This time, seven primer pairs were returned (*see* **Table 8**). If you click the primer sequence link (Forward Primer or Reverse Primer), the **GeneFisher Primers Profile - Data Sheet** about that primer pair will be returned in a new window. If you click the primer position link (FPPos. RPPos.), the **Textual Primer Pair Visualization** of that primer pair will be shown in a new window.

3.3. Analyze PCR Primers Using NetPrimer

1. Start a web browser (Netscape or Internet Explorer).
2. Replace the default URL address with http://www.premierbiosoft.com/netprimer/netprimer.html and hit return. After connection, NetPrimer web page will appear in your browser.
3. Click the **click here** link in the page to launch the NetPrimer applet.
4. After the applet is launched, type the following sequence in the **Oligo Sequence** input area: ctgcaagcaatgaaagtgga, then click the **Analyze** button. The analysis results of the primer, such as T_m, molecular weight, GC%, rating, and stability, will be shown in the

Table 8
GeneFisher Output (IUB Code for Sequence)

7 best Pairs (of max. 7)							
ID	Forward Primer	Reverse Primer	Qual.	Prod. Len.	T_mDiff.	FPPos.	RPPos.
1	NTAYMGNATGRAYTAYATGGA	GTyTGrTArTArTCnGCnCCCA	659	653	6	653	1306
2	NTAYMGNATGRAYTAYATGGA	TyTGrTArTArTCnGCnCCCAT	658	652	5	653	1305
3	NTAYMGNATGRAYTAYATGGA	AynGTnCCdATrTCyTCCCA	398	388	6	653	1041
4	NTTYAANGAYTAYTGGGA	GTyTGrTArTArTCnGCnCCCA	290	278	12	1028	1306
5	NTTYAANGAYTAYTGGGA	TyTGrTArTArTCnGCnCCCAT	289	277	11	1028	1305
6	NTAYMGNATGRAYTAYATGGA	GTyTTrAAnACrTAnACnCCCA	264	248	6	653	901
7	NTAYMGNATGRAYTAYATGGA	TyTTrAAnACrTAnACnCCCAT	262	247	5	653	900

Results area of the applet. You may also click the following buttons: **Hairpin**, **Dimer**, **Palindrome**, and **Repeat & Run**, to check the corresponding properties about the primer.

4. Notes

1. This formula only gives a very approximate T_m in the absence of denaturating agents such as formamide and DMSO, and it is only valid for primers < 20 nucleotides in length. For PCR purposes T_m-5°C is a good annealing temperature to start with. However, optimal annealing temperatures can only be determined experimentally for a certain primer/template combination and there is no formula currently available to accurately define their relationships.

   For longer primers, the nearest-neighbor method ***(6)*** offers a reliable estimation of the T_m and its formula is the following:

$$T_m = \Delta H / (A + \Delta S + R \times \ln[C/4]) - 273.15 + 16.6 \times \log[salt]$$

   where:

   ΔH (cal/mole) is the sum of the nearest-neighbor enthalpy changes for DNA helix formation (<0).

   A (cal/degree Celsius/mole) is a constant for helix initiation, which is equal to –10.8 cal /degree Celsius/mole for nonself-complementary sequences and = –12.4 for self-complementary sequences.

   ΔS (cal/degree Celsius/mole)is the sum of the nearest-neighbor entropy changes for helix formation (<0).

   R is the molar gas constant (1.987 cal/degree Celsius/mole).

   C is the primer concentration.

   [salt] is the salt concentration.

   However, primer design programs may use different formula to calculate T_m. For example, the Primer3 program uses the following formula:

$$T_m = 81.5 + 16.6(\log10([Na+])) + 0.41 \times (\%GC) - 600/length$$

   where [Na+] is the molar sodium concentration, (%GC) is the percent of Gs and Cs in the sequence, and length is the length of the sequence.

2. Keep in mind that internet is not a secure place to send your sequences. If you care about the privacy of your sequences or do not have internet access, then download a freeware or buy a commercial package and use them instead to design your primers.
3. URL stands for Universal Resource Locator, which is a unique address on the internet. However, URL is quite dynamic. Old web sites could be shut down and new sites could be set up, resulting in change of web address for a particular page or disappearance of a web page. Try to search for the new address of a web page by using a search engine, such as www.google.com.
4. You can also use your own in-house sequences for primer designs discussed in the **Methods** section. Of course, the results will vary.
5. For the input sequence of Primer3, numbers and blanks are ignored. Other letters are treated as N. FASTA format is acceptable. It is assumed that the strand direction is 5'–>3'.
6. Although GeneFisher system has been optimized for Netscape Navigator version 4.x and above, our testing showed that it works fine with Internet Explorer 5.0 and above when protein sequences are used as inputs. Netscape Navigator should be used when DNA sequences are used as inputs.
7. If you use your own input sequences, make sure that the sequences have significant homology. Otherwise the primer pair which meets your parameters will be very hard to find, if not impossible. Do not use the **Browse**... button to load your sequences, as it appears that there is a bug in reading the sequences from a file by GeneFisher.

References

1. Wallace, R. B., Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T., and Itakura, K. (1979) Hybridization of synthetic oligodeoxyribonucleotides to phi chi 174 DNA: the effect of single base pair mismatch. *Nucl. Acids Res.* **6,** 3543–3557.
2. Rychlik, W. and Rhoads, R. E. (1989) A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. *Nucl. Acids Res.* **17,** 8543–8551.
3. Wu, D. Y., Ugozzoli, L., Pal, B. K., Qian, J., and Wallace, R. B. (1991) The effect of temperature and oligonucleotide primer length on the specificity and efficiency of amplification by the polymerase chain reaction. *DNA Cell Biol.* **10,** 233–238.
4. Rychlik, W. (1993) Selection of primers for polymerase chain reaction, in *PCR Protocols. Current Methods and Applications* (White, B. A., ed.), Humana, Totowa, NJ, pp. 31–40.
5. Innis, M. A. and Gelfand, D. H. (1990) Optimization of PCRs, in *PCR Protocols* (Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J., eds.), Academic, New York, pp. 3–12.
6. Breslauer, K. J., Frank, R., Blocker, H., and Markey, L. A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **83,** 3746–3750.