

microRNA前体的折叠

南昌大学生物学实验教学中心 汪艳璐

所用软件：

- 1、 mfold(<http://mfold.rna.albany.edu/>)
- 2、 EditSeq

筛选的标准：

- (1) miRNA前体序列要能形成典型的发夹结构，且成熟的miRNA位于茎环结构的一条臂上。
- (2) 成熟的miRNA与已知植物miRNA之间只允许有四个碱基以内的差异。
- (3) miRNA与其互补的miRNA*序列之间的碱基错配应少于6个碱基(包括G:U错配以及其他非经典的碱基错配)。
- (4) pre-miRNA形成的在茎环结构中miRNA*序列不允许有发夹结构和缺口存在。
- (5) pre-miRNA形成的二级结构应具备较低的自由能(小于-15 kcal/mol)

步骤:

- (1) 将候选miRNA的搜索结果保存。当候选miRNA的成熟区与EST的反向互补序列匹配时(plus/minus), 应先将EST序列转变成反向互补序列再保存。
- (2) 在EST序列中找到miRNA的成熟区序列。
- (3) 在成熟区序列上游和下游分别取200-300bp的序列作为准前体序列, 并记下成熟序列的位置。
- (4) 将所取的准前体序列 mfold折叠。
- (5) 将能折叠出典型茎环结构的前体序列、相应的miRNA和折叠结果保留。如不能折出典型的茎环结构, 需进行调整。如多次调整都无法折出, 则淘汰此候选miRNA。

Alignments

Download ▾ GenBank Graphics

▾ Next ▲ Previous ▲ Descriptions

CJ545843 Y.Ogihara unpublished cDNA library Wh_GCPCDAM Triticum aestivum cDNA clone rwhgc14k12 3', mRNA sequence.

Sequence ID: [dbj|CJ545843.1](#) Length: 550 Number of Matches: 1Range 1: 447 to 467 [GenBank](#) [Graphics](#)

▾ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
34.2 bits(17)	0.18	20/21(95%)	0/21(0%)	Plus/Plus

```

Query 1   TGACAGAAGAGAGTGAGCACA 21
          |||
Sbjct 447 TGACAGAAGAGAGAGAGCACA 467
  
```

Related Information

[UniGene](#) - clustered expressed sequence tags

Download ▾ GenBank Graphics

▾ Next ▲ Previous ▲ Descriptions

CJ653710 Y.Ogihara unpublished cDNA library Wh_GCPCDAM Triticum aestivum cDNA clone whgc14k12 5', mRNA sequence.

Sequence ID: [dbj|CJ653710.1](#) Length: 563 Number of Matches: 1Range 1: 337 to 357 [GenBank](#) [Graphics](#)

▾ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
34.2 bits(17)	0.18	20/21(95%)	0/21(0%)	Plus/Minus

```

Query 1   TGACAGAAGAGAGTGAGCACA 21
          |||
Sbjct 357 TGACAGAAGAGAGAGAGCACA 337
  
```

Related Information

[UniGene](#) - clustered expressed sequence tags

Download ▾ GenBank Graphics

▾ Next ▲ Previous ▲ Descriptions

AL810223 d:15 Triticum aestivum cDNA clone D05_d15_plate_8, mRNA sequence.

Sequence ID: [emb|AL810223.1](#) Length: 549 Number of Matches: 1Range 1: 522 to 542 [GenBank](#) [Graphics](#)

▾ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
34.2 bits(17)	0.18	20/21(95%)	0/21(0%)	Plus/Minus

Related Information

[UniGene](#) - clustered expressed sequence tags

53% ↑ 434K/S
↓ 8.5K/S

Taxonomy

UniGene

Recent activity

[Turn Off](#) [Clear](#)

BLAST

BLAST

[See more...](#)

clade: Pooideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 550)

AUTHORS Mochida, K., Kawaura, K., Shimosaka, E., Kawakami, N., Shin-I, T., Kohara, Y., Yamazaki, Y. and Ogihara, Y.

TITLE Tissue expression map of a large number of expressed sequence tags and its application to in silico screening of stress response genes in common wheat

JOURNAL Mol. Genet. Genomics 276 (3), 304-312 (2006)

PUBMED [16832693](#)COMMENT Contact: Yasunari Ogihara
Graduate School of Agriculture
Kyoto Prefectural University
Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan.FEATURES Location/Qualifiers
source 1..550
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="DT4B (CS)"
/db_xref="taxon:4565"
/clone="rwhgc14k12"
/tissue_type="anther at meiosis"
/clone_lib="LIBEST_019551 Y.Ogihara unpublished cDNA library Wh_GCPCDAM"

ORIGIN

```

1 cagtcnaga catagaaca tgaaaaagag agaataaaca cagaaagaac gggacgtagg
61 cctgatcact aaaacgggca cagctagcta cagatgatca ccaaaagaac ccaagaacag
121 cccctgacgc tcaactcaaga cttctactgc cagcagaag ggagggacgg cgacgtgccc
181 tcggaggagc catcggcgcc gacgtggaag atcccggtgc agttcatctc gttgtcgttg
241 gagctcggga ccagcgcgcc gcaggcgttg agctgctgc cgtccatgcc gccggcgaag
301 gcggcgaagc cggtggcgcc gctggaggcc tgggagcagg cgaaccagga ggagctgccc
361 aactgctgca gtttggggac ggaggggatg tgctcgttg gacggacat ccgctgacg
421 tcgacgctgg aggagttggc cgggatgac agaagagaga gagcacagtc ggagtcgagc
481 acgtgcgtca actgtccatc ggagaacatc ttgttctgc tgctgctcgc gggcggcggg
541 cgggcggcgc

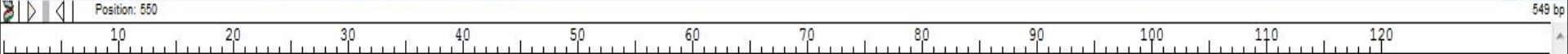
```

//

45%

↑ OK/S

↓ OK/S



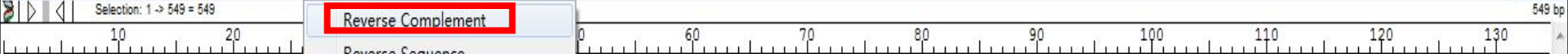
```

gggctgcaggccccgcgtcaccgtcgccggccgcgagatgcgcttctgccaacagtgcagcaggtttcacctgcttacggagtttgatgaggccaaacgcagctgtagaaagcgtctt 120
gatgggcaccaaccgtcgccgcaggaaccgcagccagatgtgatgaattctgcaagttttatgacaagtcaacaaggaacaaggttttcatctcattccaactccaagaccggagcaaa 240
actggccagggatcattaaaactgaggagaaccctattacgcacatcaactccctctaggcatcagcaacaggcagcattttgggtggctctgcgtcgacttacgccaagaaggacggc 360
gatttcctttcctgcaggaaggcgaataaaactttgccactggtgtggcacttgagccttcagtgtgccaaccgctcctcaagacggtagctcctcccagagcagcagcagcagca 480
agatgttctctgatgggtgactccagtgtggactcagactgtgctctctctcttctgtcagctccgg 549

```

Created: Monday, October 14, 2013 10:51 PM

53% ↑ 849K/S
 ↓ 13.9K/S



```
gggctgcaggccccctgcgtcacggttccaacagtcagcaggtttcacctgcttacggagttgatgaggccaaacgcagctgtagaaagcgtcttgatgggcacc130  
aacgctgcgccgaggaaaccgcagcctatgacaagtcaacaaggacaaggtttcatctcattccaactccaagaccggagcaaaactggccagggatcatataa260  
actgaggagaaccctattacgcacacaggcagcattttggggctctgctcgacttacgccaagaaggacggcgatttccttctctgcaggaaggcgaataa390  
actttgccactggtgtggcacttgagaagacggtagctcctcccagagcagcagcagcagcaagatggttctctgatgggctgactccagtgctggactcaga520  
ctgtgctctctctcttctgtcagctc549
```

- Reverse Complement
- Reverse Sequence
- Translate DNA
- Reverse Translate Protein
- Genetic Codes
- Edit Genetic Code
- DNA Statistics
- Protein Statistics

```
Reverse Complement DNA Sequence Untitled Seq #2(1,549)  
Reverse Complement DNA Sequence Untitled Seq #1(1,549)  
Created: Monday, October 14, 2013 11:15 PM  
...  
...
```

45% ↑ 0K/S
↓ 0K/S

mfold

360安全浏览器 6.2

文件 查看 收藏 工具 帮助

http://mfold.rna.albany.edu/

收藏 谷歌 网址大全 游戏中心 Links Microso Microso MSN W MSN 网 Window 购物

扩展 影视大全 网银 翻译 截图 游戏 登录管家

UNAfold | mfold.rna.albany.edu

THE RNA INSTITUTE
COLLEGE OF ARTS AND SCIENCES
UNIVERSITY AT ALBANY State University of New York

The UNAFold Web Server

Home DINAMelt Web Server **Mfold Web Server** Forum

Software

- [UNAFold](#)
- [mfold](#)
- [OligoArrayAux](#)
- [OligoArray 2.1 \(external link\)](#)
- [Ensemble Calc](#)

References

- [UNAFold](#)
- [mfold](#)

About

- [About](#)

Contact

"UNAFold"

- "Unified Nucleic Acid Folding and hybridization package"

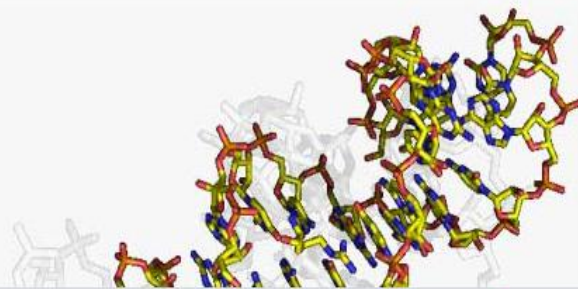
The UNAFold web server is currently an amalgamation of two existing web servers: mfold & DINAMelt. The aim of this web site is to integrate the existing servers and to expand by developing algorithms and software that will provide new services to the scientific community.

Created by

- Nicholas R. Markham
- Andrew Wright
- Leslie S. Zuker
- Michael Zuker

Maintained and supported by

- Paul F. Agris
- Eric F. Warnke



35% 0K/S
0K/S

浏览器医生 下载 100%

The mfold Web Server

[Home](#)[DINAMelt Web Server](#)[Mfold Web Server](#)[Forum](#)

Applications

- [RNA Folding Form](#)
- [DNA Folding Form](#)
- [Structure Display and Free Energy Determination](#)
- [RNA Folding Form \(version 2.3 energies\)](#)

View Folding Results

- [Folding Results](#)

Documentation

- [Mfold References](#)
- [FAQs](#)
- [Folding & output options](#)

The mfold Web Server

The mfold web server is one of the oldest web servers in computational molecular biology.

- It has been in continuous operation since the fall of 1995 when it was introduced at Washington University's School of Medicine.
- It operated at Rensselaer Polytechnic Institute from October 2000 to November 5, 2010, when it was relocated to the RNA Institute web site.
- As of the relocation date, an article describing it that was published in the first web server issue of Nucleic Acids Research in July 2003 had been cited 2893 times.
- In October 2005, [in-cites](#) ranked this article number 3 in a list of 103 "super hot papers" in science published since 2003.



35% ↑ 0K/S ↓ 0K/S

RNA Folding Form

- RNA Folding Form
- DNA Folding Form
- Structure Display and Free Energy Determination
- RNA Folding Form (version 2.3 energies)

View Folding Results

- Folding Results

Documentation

- Mfold References
- FAQs
- Folding & output options
- Folding with constraints

Software

- Mfold

About

- About

Contact

- Contact

M. Zuker

mfold web server for nucleic acid folding and hybridization prediction.

Nucleic Acids Res. **31 (13)**, 3406-15, (2003)

[\[Abstract\]](#) [\[Full Text\]](#) [\[Supplementary Material\]](#) [\[Additional Information\]](#)

The folding temperature is fixed at 37°. You may still fold with the older *version 2.3* RNA parameters, which allow the temperature to be varied.

[DNA mfold server.](#) [Quikfold.](#) Fold many short RNA or DNA sequences at once.

Enter sequence name:

Enter the sequence to be folded in the box below. All non-alphabet characters will be removed.

FASTA format may be used.

```
ctgc cacgagaaag ggaggggacgg cgacgtgcc
    181 tcggaggagc catcggcgcc gacgtggaag atccccgtgc
agttcatctc gttgtcgttg
    241 gagctcggga ccagcgcgcc gcaggcgctg agctgctcgc
ctccatgcc gccggcgaag
    301 gcggcgaagc cggtgggcgc gctggaggcc tgggagcagg
cgaaccagga ggagctgccg
    361 aactgctgca ggttggggac ggaggggatg tgctcgcttg
cacggaccat ccggctgacg
    421 tcacactaa aggaattac cgaagataac aaaaagaaa
```

前体序列

Format Sequence

Clear Constraints

Check Constraints

Enter [constraint information](#) in the box at the right. (optional) You may:

- force bases $i, i+1, \dots, i+k-1$ to be double stranded by entering:

Your job can be processed while you wait (the default) or can be submitted for batch processing by pressing the button below. In this case, you will be notified at a later time that the job is finished. If you select a *batch job*, please make sure your E-mail address is correct in the window below.

- Select: job for:
- Choose [image width](#) for png & jpg files: Small: Regular: Medium: Large: XLarge: Huge:
- Choose [structure format](#): Automatic: Bases: Outline:
- Grid lines in [energy dot plot](#): On: Off:
- Choose [structure draw mode](#):
- Choose [exterior loop type](#):
- Choose [base numbering frequency](#):
- Choose [sequence numbering offset](#):
- Choose [regularization angle in degrees](#): (Not used if 0.)
- Choose [structure rotation angle](#):
- Choose [structure annotation](#): None: p-num: ss-count: high-light:
Enter high-light regions(s): x

成熟序列的位置

Current limits: 800 bases for an immediate job, 9000 for batch.

If you wish to make comments, please select an appropriate [forum](#).

http://mfold.rna.albany.edu/results/11/13Oct14-11-58-32-65691deb35/

收藏 谷歌 网址大全 游戏中心 Links Microso Microso MSN W MSN 网 Window 购物 »

扩展 影视大全 网银 翻译 截图 游戏 登录管家

RNA folding results on 130 x Nucleotide BLAST: Search n x NCBI Blast:Nucleotide Sequ x AL810223 d:15 Triticum aes x

Applications

- [RNA Folding Form](#)
- [DNA Folding Form](#)
- [Structure Display and Free Energy Determination](#)
- [RNA Folding Form \(version 2.3 energies\)](#)

View Folding Results

- [Folding Results](#)

Documentation

- [Mfold References](#)
- [FAQs](#)
- [Folding & output options](#)
- [Folding with constraints](#)

Software

- [Mfold](#)

About

- [About](#)

Contact



13Oct14-11-58-32-65691deb35 is the 1719309th nucleic acid sequence folded on the RNA Institute mfold server.

- Monday, October 14 11:58:33 EDT 2013 -

Folding 13Oct14-11-58-32-65691deb35 at 37° C. (3.5)

- Computed for 59.52.156.5  -

```
Linear RNA folding at 5%, window = 8, max folds = 50
69 A's, 110 C's, 165 G's, 60 U/T's and 0 N's.
      10      20      30      40      50
cugccacgag aaagggaggg acggcgacgu gccucggag gagccaucgg
      60      70      80      90     100
cgccgacgug gaagaucceg uggcaguuca ucucguuguc guuggagcuc
      110     120     130     140     150
gggaccagcg cgccgcaggc guugagcugc ucgccgucca ugccgccggc
      160     170     180     190     200
gaaggcggcg aagccggugg cggcgcugga ggccugggag caggcgaacc
      210     220     230     240     250
aggaggagcu gccgaacugc ugcagguugg ggacggaggg gaugugcucg
      260     270     280     290     300
cuuggacgga ccauccggcu gacgucgacg cuggaggagu uggccgggga
      310     320     330     340     350
ugacagaaga gagagagcac agucggaguc gacgacguc gucaacuguc
```

48%

↑ 0K/S

↓ 3K/S

Averaging window Magnification Base to magnify about Plot format

View Individual Structures:

Circular structure Plots

◆ [Structure 1](#) : Initial $\Delta G = -188.30$ kcal/mol, ([Thermodynamic Details](#)).

Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).

◆ [Structure 2](#) : Initial $\Delta G = -187.90$ kcal/mol, ([Thermodynamic Details](#)).

Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).

◆ [Structure 3](#) : Initial $\Delta G = -187.50$ kcal/mol, ([Thermodynamic Details](#)).

Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).

◆ [Structure 4](#) : Initial $\Delta G = -187.30$ kcal/mol, ([Thermodynamic Details](#)).

Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).

◆ [Structure 5](#) : Initial $\Delta G = -187.20$ kcal/mol, ([Thermodynamic Details](#)).

Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).

◆ [Structure 6](#) : Initial $\Delta G = -185.50$ kcal/mol, ([Thermodynamic Details](#)).

Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).

◆ [Structure 7](#) : Initial $\Delta G = -185.20$ kcal/mol, ([Thermodynamic Details](#)).

Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).

◆ [Structure 8](#) : Initial $\Delta G = -184.60$ kcal/mol, ([Thermodynamic Details](#)).

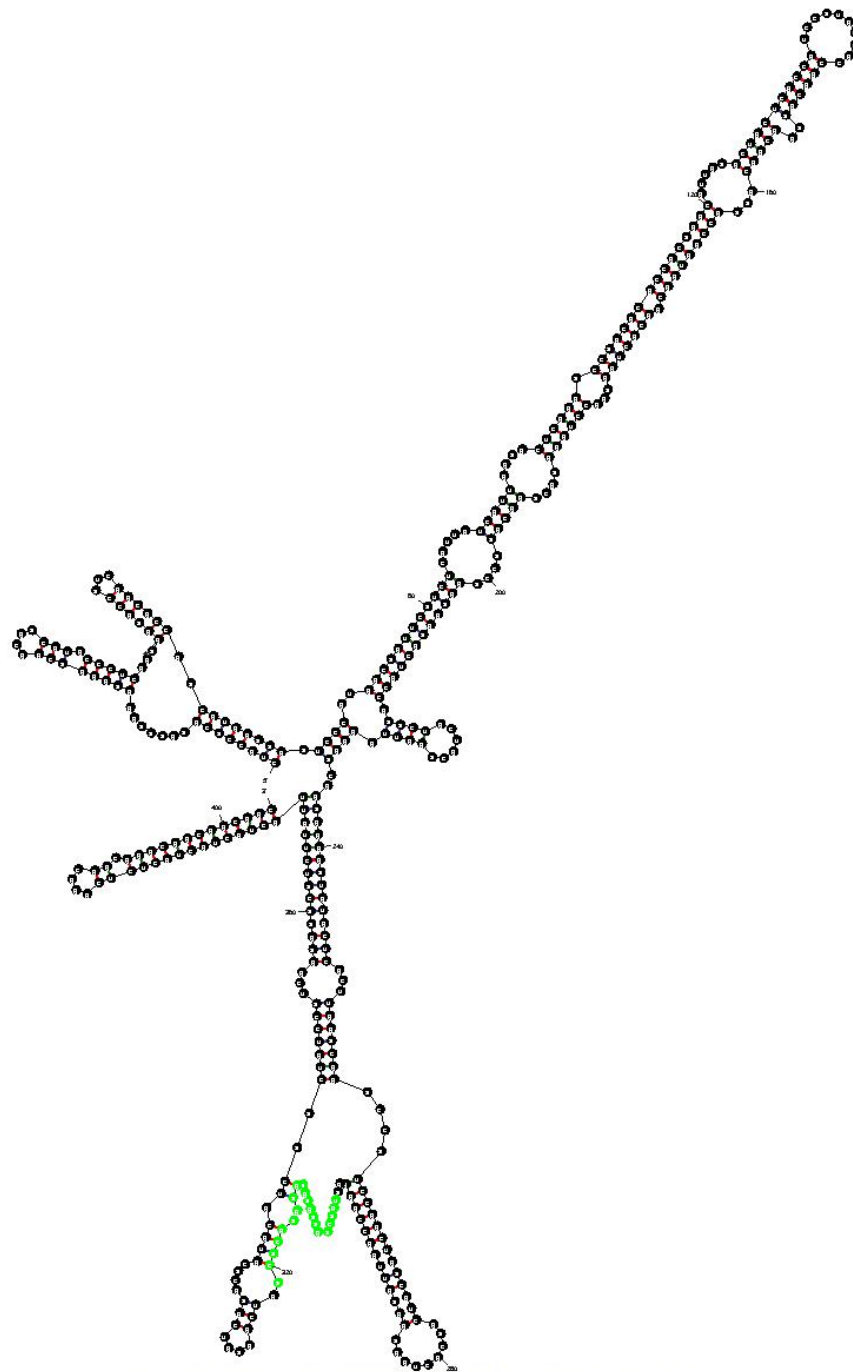
Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).

◆ [Structure 9](#) : Initial $\Delta G = -184.60$ kcal/mol, ([Thermodynamic Details](#)).

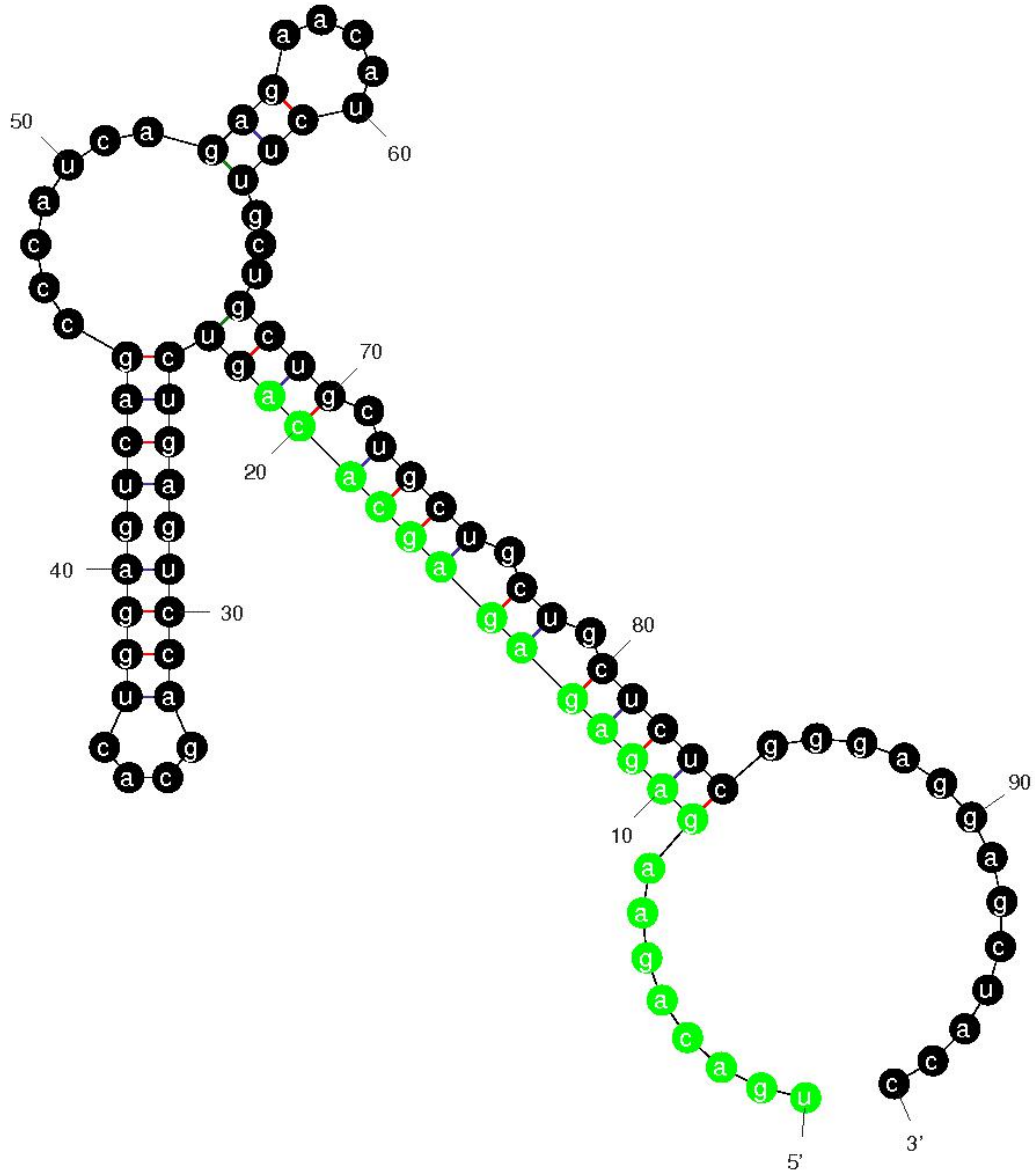
Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).

◆ [Structure 10](#) : Initial $\Delta G = -184.50$ kcal/mol, ([Thermodynamic Details](#)).

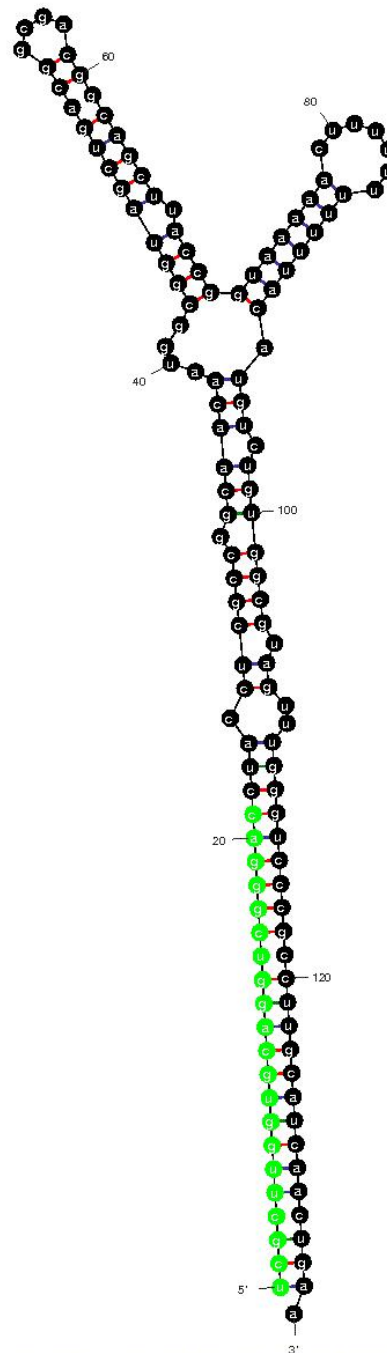
Different file formats: [PostScript](#), [pdf](#), [png](#), [jpg](#), [.ct file](#), [Vienna](#), [RNAML](#), [RnaViz ct](#), [Mac ct](#), [RNAdraw](#), [XRNA ss](#).



$dG = -178.52$ [Initially -188.30] 13Oct14-11-58-32-65691deb35

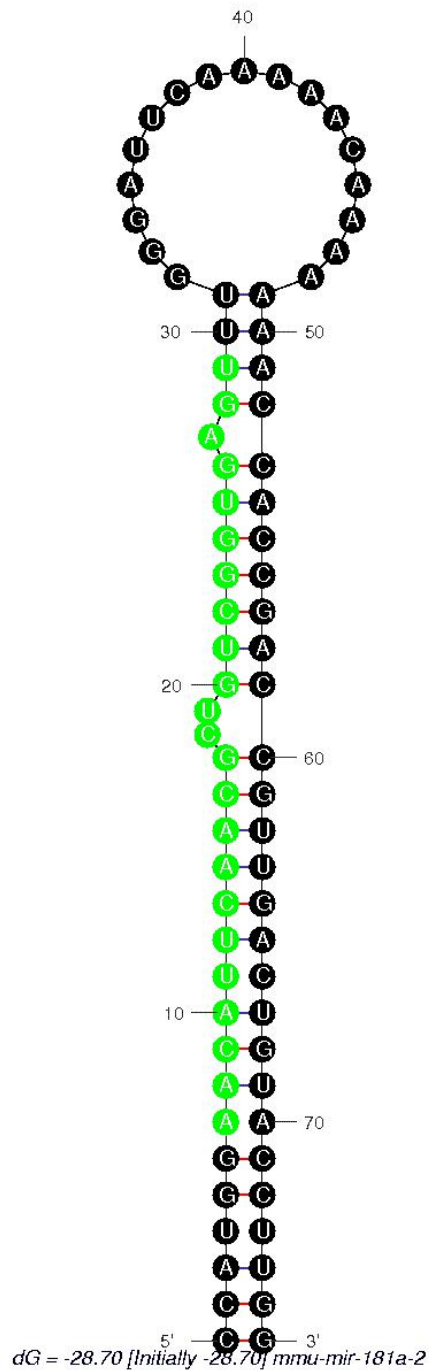


miR168a



dG = -61.80 [Initially -63.00] 12Oct16-12-17-02

mmu-mir-181a-2



EditSeq

The screenshot displays the EditSeq application window. The title bar reads "EditSeq - [Untitled Seq #1 : SEQUENCE]". The menu bar includes "File", "Edit", "Search", "Speech", "Features", "Goodies", "Net Search", "Window", and "Help". Below the menu bar is a toolbar with navigation icons and a "Position: 1" indicator. A ruler at the top shows positions 10, 20, 30, and 40. The main text area contains the text "Created: Tuesday, October 16, 2012 09:57 PM". The status bar at the bottom shows "Unspecified Search" and navigation icons.

EditSeq - [Untitled Seq #1 : SEQUENCE]

File Edit Search Speech Features Goodies Net Search Window Help

Position: 1

10 20 30 40

0

Created: Tuesday, October 16, 2012 09:57 PM

Unspecified Search

Selection: 1 -> 145 = 145

aggaagagtgcac
tgacgaaggcaa
tctagtatatca
ttgccatagaaa

- Reverse Complement
- Reverse Sequence
- Translate DNA
- Reverse Translate Protein
- Genetic Codes
- Edit Genetic Code
- DNA Statistics
- Protein Statistics

aaaagattcgattt
aaatatttgttaca
ttattgctgttgtt

40
80
120
145

Created: Tuesday, October 16, 2012 09:57 PM

筛选出的miRNA需要保留的结果

miRNA名称、 Gene ID、 Gene Source、 LS (EST序列总长度)、 LM (成熟miRNA长度)、 LP (前体miRNA长度) MFE (前体形成二级结构的最小自由能)