

효과적인 siRNA의 디자인

Designing An Effective siRNA

구남진, 조광휘*

승실대학교, 생명정보학과, 서울*

초 록

Short interfering RNA(siRNA)는 특별한 gene의 발현을 막는데 사용될 수 있고 그 gene의 기능과 치료의 적용에 많은 가능성을 가지고 있지만, 효과적인 siRNA를 디자인하는 방법은 아직까지 명확하지 않다. 효과적인 siRNA는 서열적인 경향을 가지고 있는데 낮은 G/C content, Sense strand의 3' 끝에 적은 안정성과 1번 위치에는 G/C, 19번 위치에는 A/U의 존재 여부를 들 수 있다. 이러한 특성 말고도 최근에는 mRNA의 2차구조가 RNAi 작용에 중요한 역할을 하게 되는데 복잡한 구조(hairpin, multi loop)를 가지고 수소결합을 많이 하여 안정한 상태에 있는 부분은 siRNA의 기능을 크게 줄어뜨리게 한다. 또한, siRNA가 특정한 mRNA에 작동하도록 BLAST 검색을 하여 부작용의 가능성을 배제한다.

Abstract

Short interfering RNA (siRNA) can be used to silence specific gene expression and have many potential therapeutic applications. However, how to design an effective siRNA is still not clear. Highly effective siRNA has sequence-specific properties which are low G/C content, low internal stability at the sense strand 3'-terminus, sense strand base bias(position 1 is G/C, position 19 is A/U). Recently, mRNA secondary structure plays an important role in RNAi. Target site of siRNA in high-ordered structure (i.e. hairpin loop, multi loop) or base pair of many hydrogen bonds dramatically reduce function of siRNA mediated gene silencing. Possible off-target effects of siRNA is detected from BLAST search.

서 론

RNA interference(RNAi)는 double stranded RNA(dsRNA)에 의해 촉진되어 mRNA를 분해시킴으로써 target gene의 발현을 막는 RNA silencing 과정이다[그림 1]. 이 작용은 exogenous, transposon, virus gene에서 생산된 aberrant RNA에 작용하여 세포 자신의 genome을 보호하는 기능을 한다.[1] 이 RNAi 과정을 일으키는 물질 중 하나를 siRNA라고 부르고 이것은 21~23 base pair 길이의 sense strand와 anti-sense strand로 구성된 dsRNA로 2개의 nucleotides가 3'에 overhang 되게 디자인된다. [2]

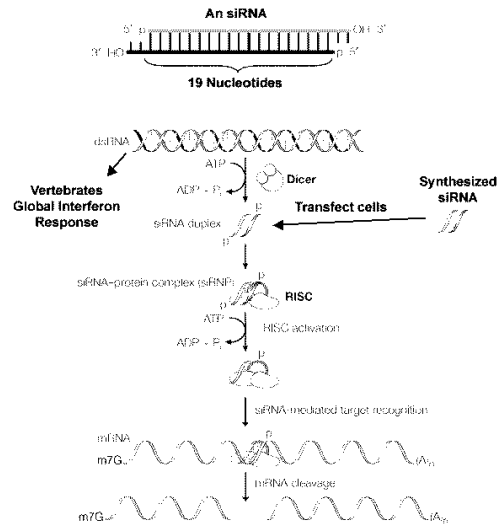


그림 1. RNA interference 과정

(Modified from Dykxhoorn, D.M., Novina, C.D. and Sharp, P.A. 2003. Killing the Messenger: Short RNAs that Silence Gene Expression. Nature Revs. Mol. Cell. Biol. 4: 457-467)

교신저자 조광휘 (Email:chokh@ssu.ac.kr)

This work was supported by Soongsil University Research Fund.

siRNA에 의해 유도되는 RNAi 과정에는 몇 가지 components가 관여를 하는데 우선 긴 dsRNA를 siRNA duplex로 만드는 RNase III enzyme인 Dicer가 존재한다. Dicer는 N-terminal에 RNA helicase인 PAZ domain과 두 개의 RNase domain을 가지며 C-terminal에는 dsRNA-binding motif를 가진다[2]. 다음은 siRNA를 mRNA로 운반하고 mRNA의 분해에 관여하는 RNA-induced silencing complex(RISC)가 있다. [1,2] 이 RISC의 형성은 Dicer가 유발시키고 RNA duplex를 한 가닥으로 떨어뜨려 anti-sense siRNA만을 가지고 mRNA와 결합하게 된다. 결합 후 RISC의 component 중 PAZ domain을 갖고 있는 argonaute가 mRNA를 분해시켜 특정한 gene의 발현을 막게 된다.

보통 mRNA는 천 개 이상의 길이를 가지고 있기 때문에 siRNA에 의한 RNAi과정을 발생시키기 위하여 고려해야 할 siRNA의 후보들이 많을뿐더러 직접 실험해보면서 그 siRNA가 얼마나 효과적인지 알아보기 위해서는 시간과 돈이 많이 소비되게 된다. 이런 문제로 인해 효과적인 siRNA를 디자인을 하기 위한 연구가 활발히 진행 중이다. 처음으로 Tuschl et. al.가 효과적인 siRNA를 선택하는 규칙을 제시하였다.[2] 이 규칙은 mRNA 서열 중에서 AA로 시작하여 TT로 끝나는 부분을 골라 19개의 nucleotide 길이를 가진 곳을 선정하고 선택된 부분이 30%~ 70%의 GC-content범위 안에 있어야 siRNA가 효과적인 기능을 할 수 있다고 말한다. [2, 3] 더 나아가 RISC의 형성이 siRNA의 기능에 중요한 요소로 생각되면서 siRNA의 열역학적인 특성과 특정한 서열의 위치를 고려한 디자인 방법이 Renolds와 co-worker에 의해 도입되었다.[4] 이들은 2개의 서로 다른 gene의 mRNA를 목표로 하는 180개의 siRNA를 관찰하여 효과적인 기능을 가진 siRNA의 특징을 발견했다. 그 특징을 바탕으로 low G/C content, sense strand의 3'끝의 low internal 안정성, inverted 서열의 반복 결합, sense strand의 위치(3,10,13,19)별 선호도라는 8가지의 기준을 마련하였다. 이러한 연구들은 단지 siRNA 자체만의 특징에 초점을 맞추었지만 다른 연구자들은 mRNA의 구조에 관심을 갖고 siRNA에 의한 gene의 발현을 억제하는 실험을 하였다. [5,6,7] 이 연구들에 의하면 siRNA와 RISC가 형성되어 mRNA에 결합하기 위해서는 mRNA의 2차구조가 접근 가능성이 커야 RNAi과정이 작용할 수 있다고 본다. 그래서 이들은 siRNA가 좋은 서열을 가지고 있어도 결합될 mRNA의 부분이 hairpin과 같은 복잡한 구조를 가지고 있거나 많은 수의 nucleotide가 수소결합을 통해 base pair를 이루어 안정적인 상태를 이루고 있다면 그 siRNA가 제대로 기능하지 못한다는 것을 설명한다.

이 같은 발견들을 토대로 효과적인 siRNA의 디자인을 위해 siRNA 서열의 성향과 mRNA의 2차 구조를 함께 적용시킨 연구를 시작하였다.

실 험

siRNA library 구축과 분석

www.mainterference.org에서 증명된 siRNA를 수집하여 siRNA library를 구축하였다. 이들은 다양한 개수를 가지고 있기 때문에 19개의 길이만을 가진 것들로 추려서 모았는데 21개의 길이는 Tuschl et. al 규칙을 따라서 처음이 AA로 시작하거나 끝이 TT로 끝나는 부분을 잘라서 모았고 23개 길이는 처음이 AA로 시작하고 끝이 TT로 끝나거나 끝이 TTTT로 끝나는 부분을 잘라서 462개의 siRNA library를 완성하였다. 그리고 나서 증명된 siRNA의 서열이 어떤 특성을 갖는지 서열의 위치별로 pattern을 분석하였다. 그 결과 첫 번째 위치에 Guanosine으로 시작되는 siRNA가 반 이상을 차지하였고 A/U는 다른 위치에 비해 적게 분포하는 것을 보였다. 다른 위치에서는 첫 번째 위치처럼 눈에 띄는 pattern을 보이는 않았지만 상대적으로 base마다 약간의 차이를 보였다.

[그림 2]

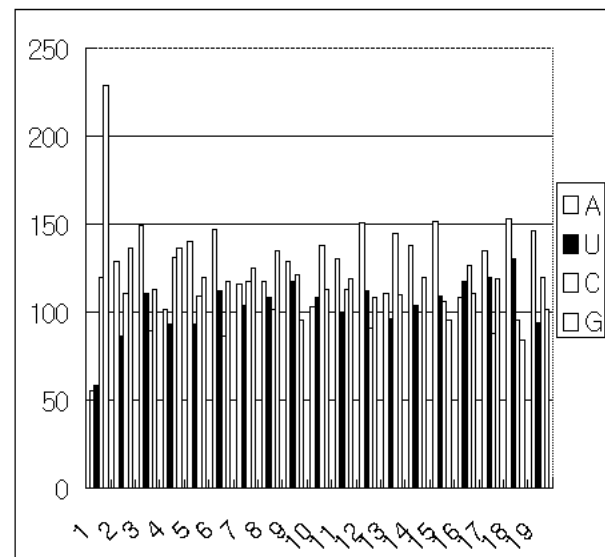


그림 2. siRNA library의 위치별 분석

siRNA 의 서열 점수 기준

위에서 언급하지 않은 siRNA의 디자인에 대한 규칙이 현재 다양하게 개발되어 있어 모아진 siRNA dataset에 적용해 보았지만 규칙에 잘 맞지 않는 결과를 보였다. 그래서 siRNA library를 바탕으로 서로 다른 규칙들을 융합하고 수정하여 새로운 함수를 마련하였다. 이것을 바탕으로 siRNA에 점수를 주는 방법은 표1과 같다.

표 1. siRNA 서열 점수 표

정의		점수	
		만족	불만족
1	낮은 G/C content (30% ~ 60%)	+1	제거
2	Sense strand 3' 끝에 적은 안정성 (14~18번 위치에 A/U 개수)	개수만큼 +1	
3	1번 위치에 G/C	+1	-1
4	19번 위치에 A	+1	
5	Start and Stop Codon 에서의 거리(최소 100 nucleotide)	선택	제거

첫 번째, 낮은 G/C content는 예전부터 효과적인 siRNA silencing에 필수적인 조건이 되어왔다. 높은 G/C content의 siRNA duplex는 서로 강한 결합을 갖고 있기 때문에 한 strand로 분리되기에 열역학적으로 불리하고 특히, RNAi 과정에 관여하는 RISC와의 결합에 방해받는다.[2] 그래서 새로운 함수에는 G/C content가 30~60%를 만족하는 경우 +1을 주고 범위에서 벗어나는 것은 제거하도록 하였다. 앞에서 언급한 것처럼 RISC와 상호작용하기 위해서는 siRNA의 sense strand의 3' 끝이 덜 안정해야 효과적인 것으로 알려졌다. [3,8] 그래서 3'에 A/U의 개수가 많은 것을 찾아 열역학적으로 덜 안정한 서열을 고르게 했다. siRNA library 분석에서는 A/U에 대한 위치별 특별한 경향을 보이지는 않았지만 상대적인 비율로 본다면 G/C에 비해 A/U가 앞서가는 것을 볼 수 있었다. 그것에 대한 점수는 그 개수만큼 +1 점수를 주어 다른 기준들과는 차이를 두었다. siRNA의 정해진 위치에 특정한 서열이 와서 효과적인 RNAi 경로를 이끌어 오는 규칙도 존재한다. [3,9] 이 규칙은 sense strand의 5'에 G/C가 오고 anti-sense strand의 5'에 A/U가 존재하면, unwinding이 잘되어 RISC의 형성에 크게 관여한다고 본다. siRNA library에서는 1번째 위치에 G가 많이 오는 것을 확인 하였지만, 19번째 위치에서는 A 서열이 약간의 비율로 앞서는 것을 확인 하였다. 그래서 1번째 위치에 G/C가 온다면 +1을 주고 다른 서열이 오면 -1을 주는 감점요소를 도입하였고 19번째에 A가 오면 +1을 주는 방식을 이용하여 차별을 두었다. 마지막으로 선택되는 siRNA가 start codon과 stop codon에서 최소한 100 nucleotide보다 멀어지도록 하였다. 3'의 untranslated region(UTR)의 경우 UTR-binding proteins이 붙고 start codon과 stop codon 부분에는 translation complex가 붙게 되어 siRNA와 RISC가 결합하는 것을 방해할 수 있다. [10] 특히 3'의 UTR에서 골라진 siRNA는 부정확한 mRNA에 결합하여 원하지 않는 RNAi 반응을 일으키기도 한다. [11, 12] 이러한 경우들을 피하기 위하여 100 nucleotide 만큼 떨어진 부분을 선택하도록 규칙을 정하였다.

mRNA의 2차 구조 도입

mRNA의 2차 구조가 siRNA서열의 특성보다 siRNA의 활성에 중요한 역할을 한다는 사실이 많은 연구들에 의해 지지되고 있는데 Luo 와 Chang은 목표가 되는 mRNA 부분마다 수소결합의 확률을 계산하여 그것을 바탕으로 수소결합을 하지 않아 unpaired nucleotide의 수가 많을수록 siRNA의 효과가 커진다는 것을 제시하였다. [13] 앞의 관계를 더 발전시켜 base-pair가 되지 않아서 개방된 부분과 mRNA의 위상학적인 구조를 energy적으로 분석하여 low negative free energy라는 개념을 만들어 이 수치를 바탕으로 부분적으로 mRNA 가진 energy와 siRNA의 silencing 선행적인 관계에 있다고 밝힌 연구도 있다. 여기서는 또한 표적이 된 mRNA의 부분이 hairpin 구조를 가지고 있다면 gene silencing의 효능이 크게 줄어든다고 말한다. [6]

이 energy적인 고려와 비슷하게 Heale et al.은 mRNA 2차 구조의 free energy와 anti-sense가 mRNA와 결합하여 보이는 이론적으로 최고 안정한 상태의 energy 차이를 계산하여 접근 가능한 표적 위치를 예측한다. [7] Energy적인 관점은 빠고 siRNA가 작용할 mRNA의 부분이 이웃하는 가지에 둘러싸여 있다면 siRNA 효과가 크게 줄어든다고 하였다. 이들은 이러한 부분을 repelling loop라고 정의하였고 그 부분을 통한 제거 알고리즘을 완성하였다. [14] 이러한 mRNA의 2차 구조에 대한 실험적인 증거와 컴퓨터적인 예측을 토대로 효과적인 siRNA를 디자인하는 프로그램에 적용시켰다. 우선 디자인할 gene의 mRNA 2차 구조를 Zuker's Mfold software를 통하여 예측한다. [15] 예측한 결과를 통해 3가지를 고려하였는데 opened base pair의 수, local free energy, mRNA가 가진 2차 구조를 도입하였다. Opened base pair의 수는 base가 서로 수소결합 했는지를 판단하고 Vienna RNA 표시법 [그림 3]을 이용하여 siRNA가 될 부분에 base-pair를 이루지 않는 base의 개수를 계산하였다.[16]



1. 실제 RNA 2차 구조

(((((.((((...)))))))))

2. Vienna RNA 2차 구조

그림 3. Vienna Structure 표시법 (pair된 base는 () 로 unpaired base는 . 으로 표시)

Local free energy는 mfold에서 pair된 base에 제공되는 ddG 값을 이용하여 부분마다 계산하였다. Unpaired 된 base의

ddG는 0으로 초기화하였다.

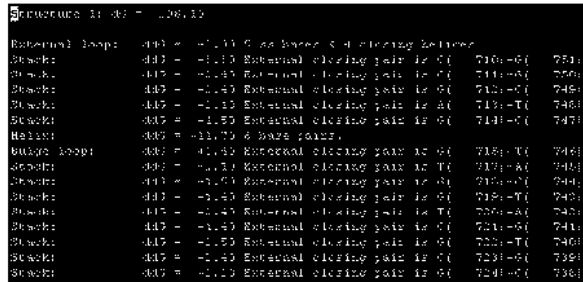


그림 4. mfold 에서 실행 후 나오는 *.det 파일의 모습

Mfold에 의해 예측되는 구조는 서열의 길이마다 다르긴 하지만 free energy 값에 따라 서로 다른 구조를 가진 후보들이 평균 20개정도가 나온다. 이 후보들 중 가장 낮은 energy 값을 가진 5개를 뽑아서 그 부분이 가진 local free energy의 평균값을 계산하였다. 그리고 가장 작은 free energy를 가진 구조에서 형성되는 RNA secondary structure의 정보를 모아 각 base가 어떤 loop를 형성하는지 찾는 알고리즘을 적용하였다. [그림 4]그 결과를 토대로 위에서 말한 siRNA의 효과에 큰 영향을 주는 Hairpin 구조와 repelling loop가 될 여지가 있는 Multi-loop 구조를 형성하는 부분은 제거하였다.[그림 5]

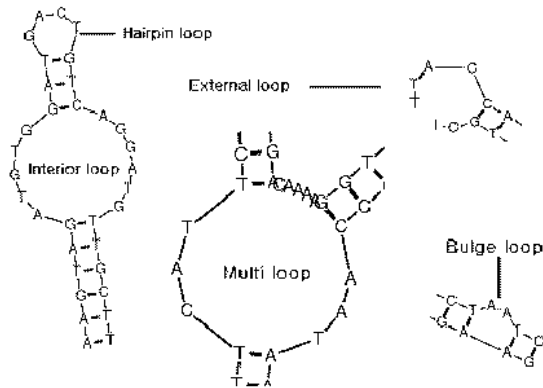


그림 5. RNA 2차 구조의 종류

BLAST 검색

siRNA 후보들을 아무리 줄인다고 하여도 mRNA의 길이가 siRNA의 서열에 비해 굉장히 크기 때문에 원하지 않게 다른 gene의 silencing을 유발시킬 지도 모른다. 이 문제점을 보완하기 위하여 다른 서열의 mRNA와 비교하여 그 gene에 작용하는 것을 피할 수 있도록 BLAST 검색을 해주는 것이 효과적이다. Gene은 transcription을 거쳐 RNA로 바뀐 후 다

시 splicing 과정에 의해 exon 부분만 mRNA를 만들기 때문에 BLAST를 이용할 때는 일반적인 nucleotide database를 이용하면 안 된다. 그래서 mRNA나 EST database를 구축하여 검색을 실행해야 하는데 EST는 mRNA 중에서 부분적인 서열만 있는 것이 대부분이기 때문에 Reference_Sequence_RNA database에서만 검색을 실행했다. 서버에는 human genomic, EST human, nucleotide database도 따로 구축되어 있기 때문에 언제든지 비교해서 실행할 수 있다.

BLAST 검색은 RNA 2차 구조 방식처럼 목표가 아닌 다른 mRNA에 작용할 수 있는 가능성을 보인다면 그 siRNA 후보들을 제거하는 것이 아니라 match되는 결과만을 보여주도록 만들었다. siRNA가 약으로 개발될 경우는 부작용을 고려하여 원하지 않는 mRNA에 작용할 수 있는 siRNA 서열을 꼭 걸러내야 한다. 그렇지만 siRNA는 gene의 기능을 알아보기 위한 knockdown에도 많이 이용되기 때문에 조건에 따라 선택할 수 있는 여지를 남겨두었다.

결 과

mRNA의 길이가 약1000개 정도 되고 밝혀진 siRNA가 2개 이상인 것을 선택해서 만들어진 프로그램을 통해 디자인해 보았다. [표2] 그래서 나온 siRNA 후보들 중에서 밝혀진 siRNA가 결과에서 나오는지를 확인하고 다른 웹에서 제공해 주고 있는 siRNA 디자인 도구들과 비교하였다. [17]

표 2. 실험에 쓴 mRNA의 siRNA 들

mRNA ID	발혀진 siRNA Sequence		
NM_006142	5'-GAGCGA AACCCUGCU CUCAG-3'	5'-GGGUGACUA CUACCGCUAC- 3'	5'-AGACAGCAC CCUCAUCAUG-3
NM_004219	5'-GAUCUCAAGUUU CAACACC-3'	5'-GUCUGUAAAAGACCAA GGGA-3'	
NM_012331	5'-CCCCUGUAGCGG CCAAACA-3'	5'-CGGGAGGGACAGACU UUUCU-3'	

다른 디자인 도구를 이용한 결과 mRNA의 길이에 따라서 siRNA 후보들이 확연히 증가하는 숫자를 보이기도 하였고 Library에 존재하는 siRNA가 모두 나와 완벽한 결과를 얻는 디자인 도구도 있기는 하였지만 그 후보들이 너무 많아서 어떤 것이 효과적인 siRNA가 맞는지 확인하기 힘들었다. 후보들이 작은 경우의 디자인 도구에서는 기능을 하는 siRNA가 한 개도 안 나오는 것이 대부분 이었다.

표 3. 다양한 디자인 도구에서 검색한 siRNA 후보

Accession 번호	Ambion	Jack	Emboss	MPI	GenScript	Qiagen	Invitrogen	Deqor	MPI+Rational	Gu
NM_004219	1/33	0/32	2 /129	2 /129	0/6	0/4	0/1	0/2	0/20	1/9
NM_006142	3/21	1/26	3 /158	3 /170	0/10	0/1	0/9	1/60	0/9	1/10
NM_012331	0/66	1/61	2 /244	2 /148	0/10	0/6	0/10	0/2	0/13	0/8

설명: Library 에 있는 siRNA 개수/ 나열된 총 siRNA

새롭게 만든 Gu 프로그램으로 돌린 결과는 NM_004219에서는 9개의 후보를 보여주고 1개를 맞추었다. NM_006142에서는 10개의 후보를 보여주고 1개를 맞추었다. NM_012331에서는 8개의 후보 중에서 단 하나도 맞추지 못하였다.

```

[cho2]@occh108su:paper1$ python test.py NM_004219
Post Sequence Score Open. max Energy
148 UAGUUUUUUUUUUUUUUUUUU 13 7 3.4
Post Sequence Score Open. max Energy
149 AUAUUUUUUUUUUUUUUUUUU 13 7 -8.4
Post Sequence Score Open. max Energy
150 GAAUUUUUUUUUUUUUUUUUU 14 6 -9.0
Post Sequence Score Open. max Energy
151 AUUUUUUUUUUUUUUUUUUU 15 6 -9.0
Post Sequence Score Open. max Energy
152 UUUUUUUUUUUUUUUUUUU 13 5 -7.7
Post Sequence Score Open. max Energy
153 GUUUUUUUUUUUUUUUUUUU 16 5 2.1
Post Sequence Score Open. max Energy
154 UUUUUUUUUUUUUUUUUUU 15 5 1.3
Post Sequence Score Open. max Energy
155 UUUUUUUUUUUUUUUUUUU 15 5 3.4
Post Sequence Score Open. max Energy
156 UUUUUUUUUUUUUUUUUUU 14 4 -0.0
Total siRNA Number is 9
[cho2]@occh108su:paper1$
    
```

그림 6. 하나의 mRNA를 프로그램 돌린 후 나온 결과의 예

BLAST 검색은 원하는 siRNA의 서열만 뽑아서 따로 실행했다. 아래의 경우 같은 종과 완전히 match되는 예로서 같은 mRNA에 대해서는 제외하고 판단해야 한다.

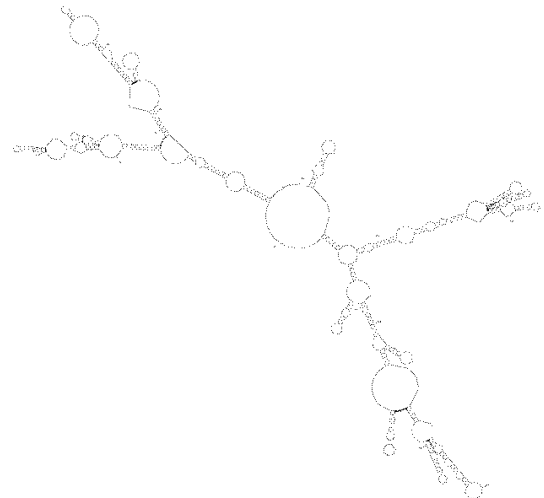
```

> seq NM_004219.1 Homo sapiens insulin1 (insulin1) (FIT1), mRNA
Length = 243

> blast -v -dbs db1 -tbl -expect 0.114
Libraries = db1 (1.0M)
strand = plus / plus

Output 1: gatctcaagtttcacacac 13
||||| ||| ||| |||
Object 158 gatctcaagtttcacacac 176
    
```

그림 7. 위의 siRNA 하나를 BLAST 검색한 결과



dG = -163.27 [initially -198.4] gi|11038651|ref|NM_004219.2
 그림 8. Mfold로 예측된 2차구조의 예

[그림 8]은 NM_004219에 대한 예측된 2차 구조 중 가장 안정한 energy를 띄고 있는 예이다. 하나의 서열을 넣게 되면 energy에 따라 평균적으로 20개 정도의 후보가 파일로 만들어진다. 프로그램 상에 어떠한 구조를 가지는지 문자로 보여주지만 [그림 8]을 확대해서 살펴보는 것이 가장 정확하게 어떤 구조를 가지고 있는지 알 수 있었다.

토 의

RNA interference를 통한 gene silencing은 gene의 기능연구와 치료적인 적용에 큰 도움을 줄 수 있는 도구이다. [18] 하지만 mRNA의 길이가 siRNA에 비해 굉장히 커서 효과적인 siRNA를 쉽게 찾기 어렵고 gene에 대한 특이성이 떨어지게 되면 예기치 않는 작용을 하게 된다. 그렇기 때문에 부작용을 줄이고 효과적인 기능을 가진 siRNA를 디자인하기 위한 기준은 꼭 필요하다. siRNA에 대한 mRNA의 2차 구조의 영

향이 뚜렷하게 보이지만 [5,6,7] 아직까지 효과적인 siRNA에 대한 디자인 도구들은 보통 **sequence**의 특성에만 중점을 두고 있다. 그것 또한 자신들만의 다양한 규칙을 가지고 있으며 일관되는 기준이 없는 상태이다.

위의 연구들처럼 mRNA의 2차 구조를 적용한 도구들이 웹에서 제공되고 있는 것도 있다[19]. 그렇지만 한 도구는 아직까지 목표된 siRNA에 대한 mRNA의 2차 구조를 사용자에게 보여주는 정도만 제시하고 있고 또 다른 도구는 mRNA의 길이가 800 nucleotide로 제한적이고 제거되지 않는 부분이 거의 대부분일 경우가 많다.[14] 2차 구조에 대한 알고리즘을 적용해 본 결과 Library에서 밝혀진 siRNA가 나오지 않는 경우는 대부분 알고리즘에서 제거되는 Multi loop 구조를 포함하고 있었다. 이러한 문제의 원인은 Repelling loop에 대한 정의[14]를 도입하지 않은 알고리즘에서 모든 Multi loop를 제거하였기 때문에 발생할 수 있고 또 다른 가능성은 mRNA의 길이가 커질수록 Mfold에서 예측되는 구조가 다양하기 때문에 실제 구조와 어느 정도 차이가 발생하기 때문이다. [20] 이러한 이유로 인하여 다른 디자인 도구와 비교한 결과 평균적인 결과 수치를 보이는 것으로 예상되고 2차 구조에 대한 예측의 정확률이 95% 이상으로 증가되고 제거하는 2차 구조의 알고리즘을 더욱 정교화 한다면 더욱 좋은 결과를 얻을 수 있을 것이다.

BLAST의 검색은 mRNA에 대한 Database를 더욱 확장하여 다른 gene의 mRNA에 작동하는 것을 미리 예방하고 match되는 개수에 따라 점수를 도입하여 몇 개의 siRNA 서열이 match되지 않더라도 RNAi 과정을 일으킬 수 있는 가능성을 배제해야한다.

참 고 문 헌

- [1] Phillip D.Zamore and Benjamin Haley. Ribo-gnome: The Big World of Small RNAs. *Science* 309 ,1519-1524 (2005)
- [2] Sayda M. Elbashir , Jens Harborth, Klaus Weber, and Thomas Tuschl. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Elsevier* 25, 199-213 (2002).
- [3] Kumiko Ui-Tei,Yuki Naito, Gunitaka Takahashi, Takeshi Haraguchi, Hiroko Ohki-Hamazaki, Aya Juni, Ryu Ueda and Kaoru Saigo. Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids* 32, 936-948 (2004)
- [4] Angela Reynolds, Devin Leake, Queta Boese, Stephen Scaringe, William S Marshall & Anastasia Khvorova. Rational siRNA design for RNA interference. *Nature biotechnology*. 22, 326-330 (2004)
- [5] Marita Overhoff, Martina Alken, Rosel Kretschmer-Kazemi Far, Marc Lemitre, Bernard Lebleu, Georg Sezakiel andlan Robbins. Local RNA target Structure Influences siRNA efficacy: A Systematic Global Analysis. *J. Mol. Biol.* 348, 871-881 (2005)
- [6] Steffen Schubert, ArnoldGrunweller, Volker A. Erdmann and Jens Kurreck. Local RNA Target Structure Influences siRNA Efficacy: Systematic Analysis of Internationally Designed Binding Regions. *J. Mol. Biol.* 348, 883-893 (2005)
- [7] Bret S. E. Heale, Harris S. Soifer, Chauncey Bowers and John J. Rossi. siRNA target site secondary structure predictions using local stable substructures. *Nucleic Acids* 33, 3, e30. (2005)
- [8] Anastasia Khvorova, Angela Reynolds, Sumedha D. Jayasena. Functional siRNAs and miRNAs Exhibit Strand Bias. *Cell* 15, 209-216 (2003)
- [9] Yuki Naito, Tomoyuki Yamada , Kumiko Ui-Tei, Shinichi Morishita and Kaoru Saigo. siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference. *Nucleic Acids* 32, 124-129 (2004)
- [10] Tuschl. The siRNA user guide. (2004) <http://www.rockefeller.edu/labheads/tuschl/sirna.html>
- [11] Xiaoyu Lin, Xiaoran Ruan, Mark G.Anderson, Jeffrey A. McDowell, Paul E. Kroeger, Stephen W.Fesik and Yu Shen. siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. *Nucleic Acids* 33 , 4527-4535 (2005)
- [12] Amanda Birmingham, Emily M Anderson, Angela Reynolds, Diane Ilsley-Tyree, Devin Leake, Yeny Fedorov, Scott Baskerville, Elena Maksimova, Kathryn Robinson, Jon Karpilow, William S Marshall & Anastasia Khvorova. 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nature Methods* 3 199-204 (2005)
- [13] Kathy Q. Luo and Donald C. Chang. The gene-silencing efficiency of siRNA is strongly dependent on the local structure of mRNA at the targeted region. *Biochemical and Biophysical Research* 318, 303-310 (2004)
- [14] WH wong, TW Lam, YC mui, SM Yiu, HF Kung, Marie Lin, YT Cheung. Filtering of Ineffective siRNAs and Improved siRNA Design Tool. *Bioinformatics* 21, 144-151 (2005)
- [15] Michael Zuker. Mfoldweb server for nucleic acid folding and hybridization prediction. *Nucleic Acids* 3 3406-3415 (2003) <http://www/bioinfo.rpi.edu/appli>

- caionts/mfold
- [16] Ivo L. Hofacker. Vienna RNA secondary structure server. *Nucleic Acids* 3 3429-3431 (2003) <http://rna.tbi.univie.ac.at/>
- [17] siRNA Design Software <http://i.cs.hku.hk/~sirna/software/sirna.php>
- [18] Moitreyee Chatterjee-Kishore and Christopher P. Miller. Exploring the sounds of silence: RNAi mediated gene silencing for target identification and validation. *Drug Discovery Today*, 10, 1559-1565 (2005)
- [19] Gerrit Schramm & Rebecca Ramey. siRNA design including secondary structure target site prediction (2005) <http://www.mwg-biotech.com/>
- [20] Jens Reeder, Matthias Hochsman, Marc Rehmsmeier, Bjorn Voss, Robert Giegerich. Beyond Mfold: Recent advances in bioinformatics. *Biotechnology* 124, 41-55 (2004)
- [21] Ambion (2003). Ambion siRNA Target Finder. http://www.ambion.com/techlib/misc/siRNA_finder.html
- [22] Lin, J. (2002). Jack Lin's siRNA Sequence Finder. <http://www.sinc.sunysb.edu/Stu/shilin/rnai.html>
- [23] Williams, G. (2002). The siRNA design tool EMBOSS. <http://www.hjmp.mrc.ac.uk/Software/EMBOSS/apps/sirna.html>
- [24] MIT (2002). siRNA Selection Program. White head Institute for Biomedical Research. <http://jura.wi.mit.edu/pubint/http://iona.wi.mit.edu/siRNAext/>.
- [25] Qiagen (2003) The siRNA design tool Qiagen. http://python.penguindreams.net/Xeragon_Order_Entry/jsp/Index.jsp