

hybridization signal과 Gibbs Free Energy의 관계성 검토

유영수, 박선아, 박수영, 이경아, 이준석, 문일[†]
연세대학교 화공생명공학부

A case study on relationship between hybridization signal and Gibbs free energy

Young-su Ryu, Sun-ah Park, Soo-young Park, Kyoung-ah Lee,
Joon-suk Lee, and Il Moon[†]
School of Chemical Engineering and biotechnology, Yonsei University

서론

지금까지는 DNA chip에서 probe와 target DNA간의 hybridization시 ΔG° 가 hybridization signal에 영향을 줄 것이라는 추측만을 하고 있을 뿐 실제로는 여러 번의 실험을 통해서 얻은 실험 데이터를 database화해서 hybridization signal 분석에 지침만을 제시하고 있다. 따라서 이 방법은 정확하긴 하지만 저장된 데이터의 정확성과 많은 자료를 요구한다. 여기에서는 용액 상에서 hybridization한 결과로부터 signal의 세기와 ΔG° 의 관계를 재분석해보고, 이 결과를 DNA chip상에서의 최적화된 probe 설계와 같이, hybrid 결과 design에 적용하고자 한다.

1. DNA의 안정성과 DNA chip에서의 signal 분석과정

DNA 분자의 열적 안정성은 용매의 염의 농도, 용매의 pH, % C+G 에 따라 다르며 T_m (melting temperature)로 나타낼 수 있다. 간단히 말하자면 T_m 이 높을수록 duplex의 열적 안정성도 높다는 것이다. T_m 으로 DNA의 열적 안정성을 얘기한다면 ΔG° 는 DNA분자의 결합 안정성을 나타낸다. 즉 ΔG° 의 값이 낮다는 것은 안정한 결합을 의미한다. 따라서 PCR의 primer, DNA chip의 probe를 설계할 때 대부분 ΔG° 를 기준으로 한다. DNA chip은 최적화 알고리즘으로 선택된 이 단일 가닥(probe)을 형광 표지를 한 상보적인 서열과 hybridization을 시킨 후 결합되지 않은 것들을 씻어내고 레이저 형광 스캐너로 그 세기를 읽는다. Probe의 강도는 형광 세기(signal)와 관계가 있으므로 형광 세기의 표준화를 통해 signal과 background를 구별해 낼 수 있게된다. 즉 baseline과 실험데이터를 비교해서 signal의 타당성을 보는데, 이때 baseline은 실험 데이터에만 의존하고 있으며 현재 이 data를 분석하는 s/w의 개발이 진행중이다.

2. NN 변수 및 위치에 따른 ΔG° 계산

여기에서 사용된 열역학적 변수들은 Nearest-neighbor model(NN model)을 사용하여 구한 값이다[1]. 그러나 여기에서의 변수들은 37°C에 해당하는 값이므로 Eqn. (1-1)에 의해 주어진 실험온도인 12°C에 맞는 변수들로 바꾸어 사용했다(Table 1). 그리고 각 서열에 해당하는 ΔG° (total)은 다음 식 Eqn. (2)에 의해 계산했다. 핵산의 ΔC_p° 는 온도변화에 거의 변화가 없으므로[2, 3] ΔH° 와 ΔS° 도 온도영향을 무시하고 일정한 값을 사용할 수 있다.

따라서 먼저 각 서열의 ΔG_{37}° 을 구하고 Eqn. (1-2)를 사용해서 ΔG_{12}° 를 구했다.

$$\Delta G_T^\circ = \Delta H^\circ - T\Delta S^\circ \quad (1-1)$$

$$\Delta G^{\circ}_T = \Delta G^{\circ}_{37} + (37 - T)\Delta S^{\circ} \quad (1-2)$$

$$\Delta G^{\circ} (total) = \sum_i n_i \Delta G^{\circ} (i) + \Delta G^{\circ} (init.w/term G \cdot C) + \Delta G^{\circ} (init.w/term A \cdot T) \quad (2)$$

sequence	ΔG°_{37} kcal/mol	ΔH° kcal/mol	ΔS° cal/k · mol	ΔG°_{12} kcal/mol
AA/TT	-1.00	-7.9	-22.2	-1.555
AT/TA	-0.88	-7.2	-20.4	-1.390
TA/AT	-0.58	-7.2	-21.3	-1.113
CA/GT	-1.45	-8.5	-22.7	-2.018
GT/CA	-1.44	-8.4	-22.4	-2.000
CT/GA	-1.28	-7.8	-21.0	-1.805
GA/CT	-1.30	-8.2	-22.2	-1.855
CG/GC	-2.17	-10.6	-27.2	-2.850
GC/CG	-2.24	-9.8	-24.4	-2.845
GG/CC	-1.84	-8.0	-19.9	-2.338
Init.w/term.G · C	0.98	0.1	-2.8	0.91
Init.w/term.A · T	1.03	2.3	4.1	1.133

Table 1. ΔH° , ΔS° and ΔG°_T NN parameter in 1M NaCl

Perfect matched된 서열의 형광 세기에 대한 데이터는 Z. Strezoska(1991)[4]로부터 얻었다. Table 2. 는 논문의 데이터 값과 실제 계산한 값을 정리한 것의 일부이다. G+C content를 42%정도로 하여 한 개의 probe당 2~3개의 G+C 염기가 포함된 octamer와 nonamer를 12°C에서 3h동안 hybridization을 시켰다. Table 2.에서 Hfm은 가장 강한 signal을 10으로 주었을 때의 full-match DNA dot의 상대적인 값이다.

No.	sequence	Hfm	Tm	ΔG°_{37} kcal/mol	ΔS° cal/k · mol	ΔG°_{12} kcal/mol
1	GTCTGAAA	7	27.9	-6.76	-153.6	-10.6
2	TGTCTGAA	4	29.7	-7.16	-147.2	-10.84
3	TTGTCTGA	1.5	29.7	-7.16	-147.2	-10.84
4	CTTGTCTG	3	30.7	-7.24	-159.8	-11.235
5	TCTTGTCT	1	28.7	-6.99	-145.5	-10.6275
6	ATCTTGTC	4	27	-6.64	-151.8	-10.435
7	AATCTTGTC	10	33.1	-7.64	-174	-11.99
8	GAATCTTG	10	25.3	-6.25	-158.5	-10.2125

Table 2. probe used, their target composition, obtained hybridization, and caculated free energy

각 염기쌍은 위치에 따라서 duplex의 안정성에 영향을 미치는 정도가 다르다. 그래서 위치에 따른 ΔG° 를 계산해보았다. 즉, 위의 데이터는 용액 상에서 실험을 한 것이므로 terminal에 있는 dimer는 같은 영향을 미친다고 가정하고 양끝으로부터 같은 위치에 있는 dimer의 ΔG° 의 값을 합했다. 예를 들어 Figure. 1의 경우에 양쪽 terminal로부터

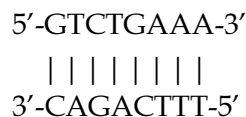


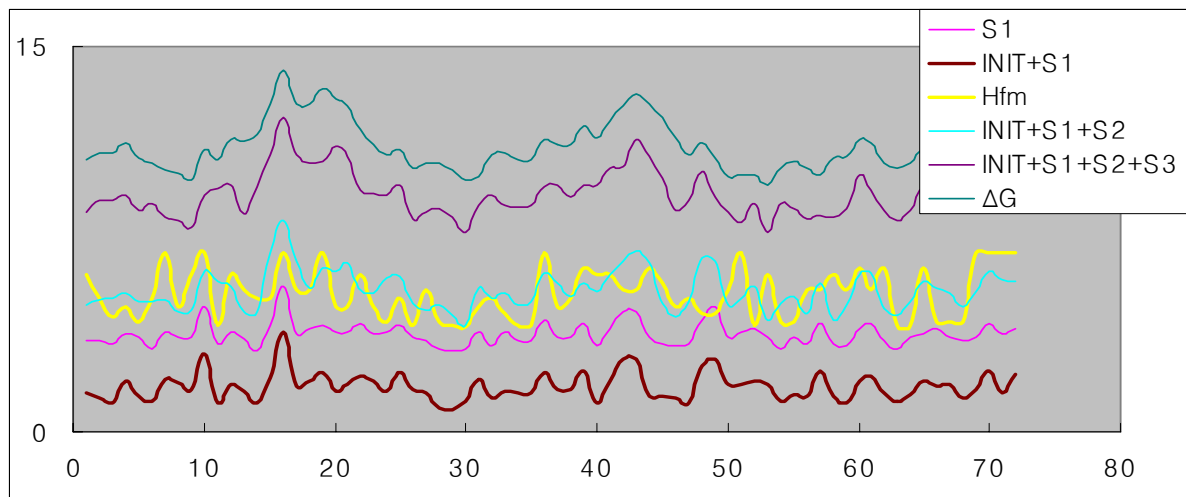
Figure 1. NN model dimer인 GT/CA, TT/AA 의 ΔG° 의 값인 -1.44과 -1.00을 합해 'S1'이라 하고 그 다음 dimer인 TC/GA와 TT/AA의 ΔG° 합을 S2'라 했다. 'S3'는 CT/GA + TC/AG, 'S4'는 TG/AC이다. Table 3은 위 Table 2의 서열에 위치에 대한 ΔG° 값의 일부이다.

No.	sequence	Hfm	Tm	INIT.	S1	S2	S3	S4
1	GTCTGAAA	7	27.9	2.0425	-3.555	-3.41	-3.66	2.0175
2	TGTCTGAA	4	29.7	2.265	-3.5725	-3.855	-3.8725	-1.805
3	TTGTCTGA	1.5	29.7	2.265	-3.41	-4.035	-3.805	-1.855
4	CTTGTCIG	3	30.7	1.82	-3.8225	-3.36	-3.8725	-2
5	TCTTGTC	1	28.7	2.265	-3.66	-3.66	-3.555	-2.0175
6	ATCTTGTC	4	27	2.0425	-3.245	-3.855	-3.8225	-1.555
7	AATCTTGTC	10	33.1	2.0425	-3.41	-3.99	-3.8725	-3.36
8	GAATCTTG	10	25.3	1.82	-3.8725	-3.11	-3.195	-1.855

Table 3. ΔG° of dimer position

결과 및 토론

Table. 2에서 볼 수 있지만 ΔG° 와 hybridization signal은 그다지 관계가 있어 보이지는 않는다. 그 동안의 많은 연구에서조차도 염기서열의 ΔG° 가 작을수록 hybridization signal이 강해지는 경향이 있지만 상당한 오차를 보인다고만 분석이 되어졌다[4]. 그러나 분명히 형광 세기는 클수록 duplex가 안정하다고 말할 수 있다. Table 3과 같은 방법으로 계산된 73개의 probe의 ΔG° 와 실제 실험결과인 hybridization signal 값을 가지고 그래프를 그려보았다.

Figure 2. hybridization signal 과 sequence의 position에 따른 ΔG° 의 영향

Octamer의 경우는 Figure. 2에서 볼 수 있듯이 terminal의 dimer와 initiation의 값을 합한 경우와 형광의 상대값 (가장 강한 signal을 3으로 정함)은 가장 비슷한 경향성을 띠고 center의 합이 더해질수록 점점 더 그 경향성이 달라지고 있음을 볼 수 있다. DNA chip의 경우에 mismatched가 DNA 가닥의 center보다는 terminal 부분에서 많이 생기는 이유도 hybride의 안정성을 결정하는 요인이 terminal에 있기 때문일 것이다.

No.	sequence	Hfm	Hp
1	GTCTGAAA	7	6.1
2	TGTCTGAA	4	5.2
3	TTGTCTGA	1.5	4.45
4	CTTGTCTG	3	4.9
5	TCTTGTCT	1	4.3
6	ATCTTGTC	4	5.2
7	GAATCTTG	10	7
8	GATGAATC	3	4.9

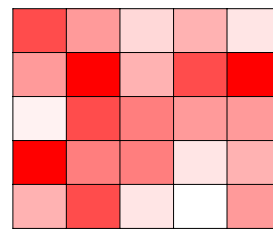


Figure 3.
experimental signal

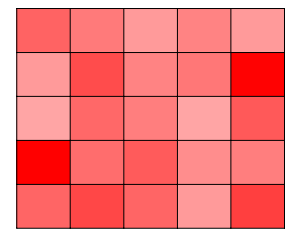


Figure 4.
predicted signal

Table 4. experimental signal과
predicted signal의 계산 값

Hfm : 실험에 의한 signal 의 상대적인 값

Hp : ΔG° 로써 예측된 signal의 상대적인 값

위에서 계산한 값을 근거로 가상의 array를 만들어서 실험 데이터와 terminal+penultimate의 ΔG° 로써 예측된 값을 각각 표시해보았지만 형광세기의 상대적인 값과 terminal + penultimate의 ΔG° 의 값이 정확히 비례관계에 있지는 않다.[Figure. 3,4] 따라서 이 두 값의 관계에는 염기서열의 ΔG° 뿐만 아니라 좀더 복잡한 인자가 있을 것으로 사료된다. 즉 용액에서든 chip에서든 dulpex의 안정성은 단순히 ΔG° 만으로는 판단할 수 없을 것이다. 더구나 octamer의 경우에는 염기서열의 길이가 짧으므로 terminal dimer의 영향 외에 다른 dimer들의 영향이 그리 크지 않지만 염기서열의 길이가 길어질수록 더 많은 고려가 필요하다. 실제로 염기가 한 개 더 많은 nonamer의 경우에는 한 눈에 경향성을 보기란 어려웠다. 여기서 분석해 본 데이터는 용액에서 실험한 것이다. 실제로 chip상에는 용액보다 훨씬 duplex의 안정성이 낮을 뿐만 아니라[5] 일반적으로 probe의 길이가 길기 때문에 실험 데이터를 근거로 좀더 정확한 분석이 요구된다.

참고문헌

1. SantaLucia, J. , A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics, *Proc.Natl.Acad.Sci.USA* , **95**, 1460-1465 (1998)
2. Chaires,J.B. , Possible origin of differences between van't Hoff and calorimetric enthalpy estimates, *Biophys. chem.* , **64** , 15-23 (1997)
3. Liu, Y. Sturtevant, J. M. , Significant discrepancies between van't Hoff and calorimetric enthalpies. III , *Biophys. chem.*, **64** , 121-126 (1997)
4. Z. Strezoska, T. Paunesku, D. Radosavljevic, I. Labat, R. Drmanac, and R. Crkvenjakov , DNA sequencing by hybridization: 100 bases read by a non-gel-based method, *Proc.Natl.Acad.Sci.USA*, **88** , 10089-10093 (1991)
5. Fotin, A.V., Drobyshev, A.L., Proudnikov, D.Y., Perov, A.N., and Mirzabekov, A.D., parallel thermodynamic analysis of duplexes on oligodeoxyribonucleotide microchips, *Nucleic Acids Res.*, **26** , 1515-1521 (1998)
6. genetics.hannam.ac.kr
7. www.affymetrix.com/technology