



Advanced DNA Sequencing Technology

(DNA sequencing in free-solution electrophoresis)



홍익대학교 화학공학과
원종인 교수

서론

2001년 2월 미국 NIH 소속 NHGRI (National Human Genome Research Institute) 와 민간 기업체인 Celera Genomics 에서는 인간의 유전자 (DNA)에 대한 첫 번째 판독 결과를 동시에 공표하였으며 [1, 2], 그 후 2003년 4월에는 99.9% 이상의 정확도를 갖는 인간 유전자 연구 결과를 공동으로 발표했다. 세계 과학사를 돌이켜 보았을 때, 인간 유전자 판독에 대한 연구 결과는 인간이 이루어낸 위대한 과학적 업적들 가운데 하나로서, 이러한 연구 결과물들을 바탕으로 앞으로 많은 유전병의 원인을 규명하고 그에 따른 치료제가 개발되리라 기대하고 있다. 또한 인간의 유전자 외에도 품종 개량 및 치료제 개발의 목적으로 500 종 이상의 여러 동식물 유전자 해독에 대한 연구가 세계적으로 진행 중에 있다.

이러한 유전자 판독 (DNA sequencing)의 결과물들은 물론 많은 과학자들의 노력에 의하여 이루어졌지만, 이를 빠르게 해독 할 수 있는 분석 기술의 개발이 이루어지지 않았더라면 불가능했을 것이다. 인간의 유전자의 경우 약 30억 개 정도의 단량체 (nucleotide) 로 이루어진 거대한

biopolymer로서, DNA sequencing 이란 이렇게 구성된 연속적인 단량체의 염기 (base) 서열을 결정짓는 일을 의미한다. 과거에는 판 전기영동 방식 (slab gel electrophoresis) 에 의하여 몇몇 유전자에 대한 판독이 시도되었으나, 인간 유전자 판독과 같은 현재의 DNA sequencing 작업은 대부분 자동화된 다중 모세관 전기영동 방식 (capillary array electrophoresis, CAE) 에 의하여 이루어지고 있다 (그림 1-A 참조). CAE에 의한 현재의 분석 기술은 먼저 가는 모세관 내부에 자동으로 polymer gel (또는 polymer matrix) 을 충전시킨 후 분석하고자 하는 DNA sample 들을 자동으로 loading 하여 그 크기에 따라 분리하는 방식으로, 기존의 slab gel electrophoresis 의 문제점인 사람이 수동으로 gel 과 DNA sample을 loading 해야 하는 번거로움을 해결하여 반복 작업이 요구되는 거대 분자량의 유전자 판독이 가능하게 되었다. CAE 외에도 현재 작은 기판 위에서 DNA sequencing을 수행하는 초소형 전기영동 방식 ("microfluidic chips" 또는 "microchip electrophoresis") 에 대한 연구가 여러 연구 기관에서 활발히 진

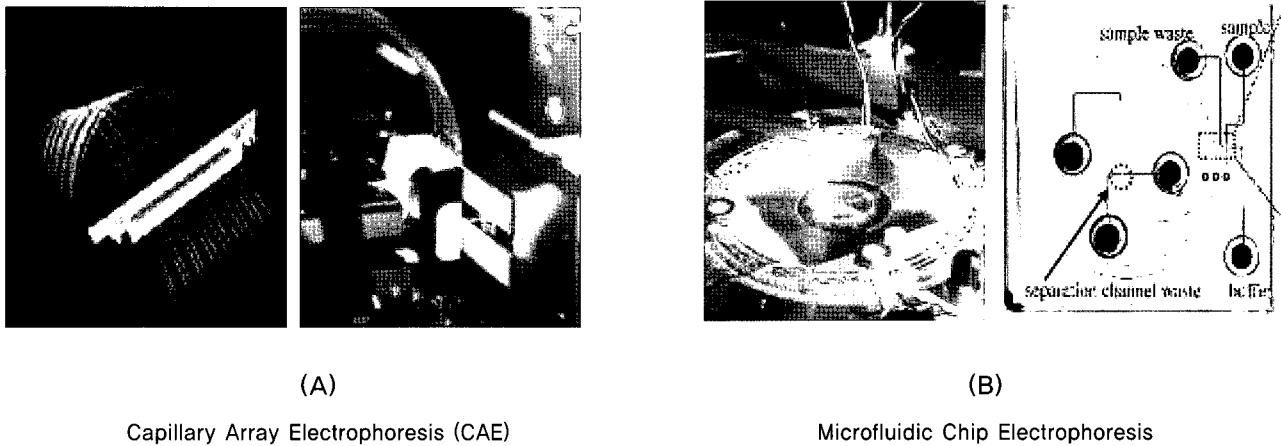


그림 1. Capillary array electrophoresis and microfluidic chip electrophoresis

행 중에 있는데, 이는 작은 기관 위에 보다 가는 모세관을 이용하므로 기존의 CAE 방식보다 더 빠르고 적은 비용으로 분석이 가능하다는 장점 때문이다[3-6] (그림 1-B 참조).

이러한 현재의 DNA 분석 방법의 경우, 먼저 가는 모세관 내부에 polymer gel을 채운 다음 DNA sequencing이 이루어지는데, 이는 일반적으로 gel을 사용하지 않을 경우 크기에 따른 DNA 분리가 불가능하기 때문이다 (즉, DNA는 수용액 상태에서 크기에 관계없이 같은 이동 속도를 가짐). 그러나 polymer gel을 이용한 이러한 분석 방법에는 몇 가지 커다란 단점이 있다. 높은 점도의 polymer gel을 머리카락보다 가는 모세관 내부에 충전하고 사용 후 세척해 내는 일이 가장 큰 어려움으로 여겨지고 있으며, 또 점도가 높은 polymer gel을 충전할 때 높은 압력이 발생해 가는 모세관이 파손되어 버리는 어려움이 있다. 또한 polymer gel을 사용하는 기존의 분석 방법은 높은 전장에 의해 발생하는 해상도 저하 (electric field-induced band broadening) 및 DNA 편향 효과 (DNA orientation effect in polymer gel) 때문에 높은 효율의 DNA sequencing을 기대하기 어려웠다.

이 글에서는 polymer gel (또는 polymer matrix) 을 사용하지 않고 수용액 상태에서 DNA sequencing 이 가능한 새로운 분석 기술에 대하여 소개하고, 이 기술의 특성을 살펴보고자 한다.

본 론

1. 수용액과 polymer gel 에서의 DNA 전기 영동

수용액과 polymer gel 에서 수행되는 전기영동의 이론적 배경은 다음과 같다. 전기영동 (electrophoresis) 이란 전기장 안에서 생화학물질의 이동속도, 즉 이동도 (electrophoretic mobility, μ), 의 차이에 의하여 분리하는 기술이다. 정상상태에서 분자의 이동도는 전하 (charge, q) 와 저항 (friction, f) 의 비로 정의되어질 수 있으며 식 (1) 과 같이 나타낼 수 있다.

$$\mu = q / f \tag{1}$$

그러나 DNA의 경우 전하의 크기가 대부분 phosphate ion의 수에 의하여 결정되어지므로 전하 (q)는 염기 서열에 관계없이 DNA 길이 (N) 에 비례하게 된다. 또한, 전기장 안에서 수용액 상태인 DNA는 “free-draining coil” 구조를 가지므로 저항 (f) 역시 DNA 길이에 비례하게 된다. 따라서 분자량인 q 값이 DNA 길이에 비례하고 분모항인 f 값 역시 DNA 길이에 비례하므로 이들의 비로 정의되는 이동도 (μ) 는 DNA 길이에 관계없이 항상 일정한 값 (μ_0)

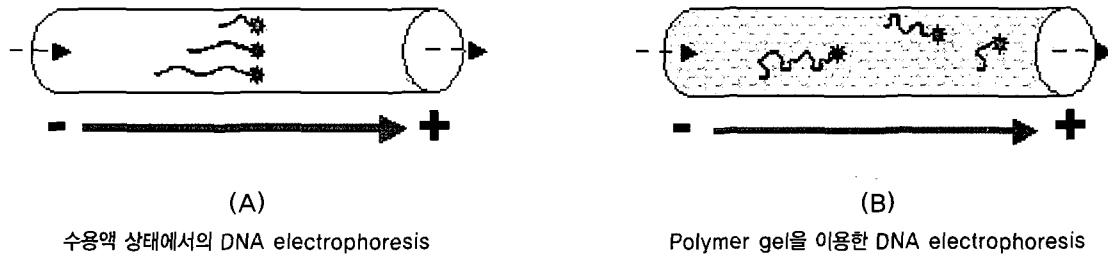


그림 2. 수용액과 polymer gel에서의 DNA 전기영동 방식

을 갖게 된다[7]. 이는 왜 DNA가 수용액 상태 (또는 buffer 상태) 에서 크기에 따른 분리가 불가능한가를 설명해 주고 있다[8] (그림 2A 참조).

$$\mu = q / f = (\rho N) / (\xi N) = \rho / \xi = \mu_0 \quad (2)$$

(여기서, ρ 와 ξ 는 각각 DNA monomer의 전하와 저항 값이다.)

따라서 현재 DNA 분석 방법은 대부분 polymer gel (또는 polymer matrix)를 이용하는데, 가는 모세관 내부에 polymer gel을 충전시킨 후 DNA를 loading 시킴으로써 DNA의 저항 (f) 값이 수용액 상태의 저항과 다른 값을 갖게 되고 (즉 gel 상태에서의 f 값은 N 에 선형적으로 비례하지 않음), 그 결과 크기에 따른 DNA 분리가 가능한 것이다. 이 때 polymer gel은 분자체 (molecular sieve) 역할을 함으로써 크기가 작은 DNA는 빠르게 이동하고 크기가 큰 DNA는 느리게 이동 한다 (그림 2B 참조).

그러나 polymer gel을 이용한 현재의 분석 방법에는 다음과 같은 커다란 단점들이 있다. 먼저, 높은 점도의 polymer gel을 머리카락보다 가는 모세관 내부에 충전하고 사용 후 세척해 내는 일이 가장 큰 어려움으로 여겨지고 있다. 최근 Nanotechnology 의 급속한 발전에 힘입어 수십 nm 의 직경을 갖는 초미세 모세관을 제작할 수 있으나, 이러한 초미세 모세관에 고정도의 polymer gel을 충전하는 어려움 때문에 아직까지 초미세 모세관을 이용하여 자동화된 microchip system이 상용화되지 못하고 있는 실정이다. 또한 polymer gel 내에서 DNA의 편향 효과 (DNA orientation effect in polymer gel)도 polymer gel

electrophoresis를 사용했을 때의 한계점으로 지적 받고 있다. 이는 polymer gel electrophoresis의 본질적인 문제점으로 (intrinsic problem), 높은 전기장에서 polymer gel 내의 DNA는 편향 효과를 보이게 되고 이 때문에 DNA의 분리 해상도가 떨어지게 된다[9]. 마지막으로 polymer gel의 비싼 비용도 polymer gel electrophoresis의 큰 문제점이다. 2002년 NHGRI Workshop 연구 보고에 의하면 인간 유전자를 해독하는데 필요한 비용은 약 72%의 고정비 (fixed cost) 와 38%의 변동비 (variable cost) 로 구성되어 있으며, 이 중 polymer gel의 가격이 전체 변동비의 약 33%에 해당된다고 발표하였다[10].

2. 새로운 DNA 전기영동 방법 (End-Labeled Free-Solution Electrophoresis)

1992년 한 연구 보고에 의하면, 단백질과 같은 균일한 크기의 저항 물질 (drag-tag)을 DNA fragments에 연결시킬 경우, 수용액 상태에서 크기에 따른 DNA 분리 (또는 sequencing)가 가능하다는 것이 발표되었다[11]. 분석하고자 하는 서로 다른 길이의 DNA fragments에 균일한 크기의 저항 물질 (“drag-tag”)을 연결시키면, 상대적으로 저항감을 덜 느끼는 긴 DNA는 빨리 이동할 것이고, 짧은 DNA는 저항감을 크게 느껴 느리게 이동할 것이므로 수용액 상태에서 크기에 따른 DNA 분리가 가능할 수 있다는 내용이다. 이 방법은 End-Labeled Free-Solution Electrophoresis (ELFSE) 이라 불리며, 비교적 최근에 개발

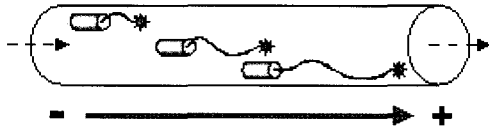


그림 3.
End-Labeled Free-Solution Electrophoresis (ELFSE)

된 전기영동 방식이다 (그림 3 참조). 이는 수용액 상태에서 DNA 분리가 이루어지는 것으로서, polymer gel을 사용할 때 보다 분석 시간을 현저히 단축시킬 수 있으며, 또한 점도가 높은 polymer gel을 사용하지 않으므로 아주 가는 모세관을 가지는 microchip 시스템에 보다 효율적으로 사용되어 질 수 있다는 큰 장점이 있다.

ELFSE의 이론적인 배경을 살펴보면 다음과 같다[12]. 식 (2)에 의하면 ssDNA (single-stranded DNA) 의 이동도는 다음과 같이 정의될 수 있고 $[\mu = (\rho N) / (\xi N)]$, DNA와 연결된 저항 물질인 drag-tag의 이동도는 식 (3)과 같이 정의될 수 있다.

$$\mu_{\text{drag-tag}} = q / f = (\rho \beta) / (\xi \alpha) \quad (3)$$

(여기서 $\xi\alpha$ 와 $\rho\beta$ 는 각각 drag-tag의 저항과 전하의 크기)

따라서 DNA와 drag-tag의 연결물질 (conjugated molecule) 의 이동도는 DNA와 drag-tag의 segregation이 발생하지 않는 전기장 하에서 (아주 높은 전기장 하에서는 segregation 이 발생) 각각의 저항과 전하를 합한 값인 식 (4)와 같이 정의될 수 있다 (그림 4 참조).

$$\mu = q / f = \rho(N + \beta) / \xi(N + \alpha) = \mu_0 \cdot [(N + \beta) / (N + \alpha)] \quad (4)$$

식 (4)에 의하면 연결물질의 이동도는 DNA 길이인 N의 함수이며 따라서 수용액 상에서 DNA 길이에 따라 이동도 값이 달라지므로 분리가 가능한 것이다.

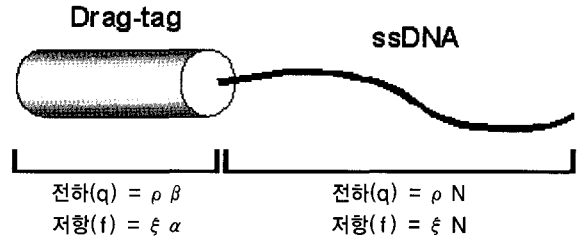
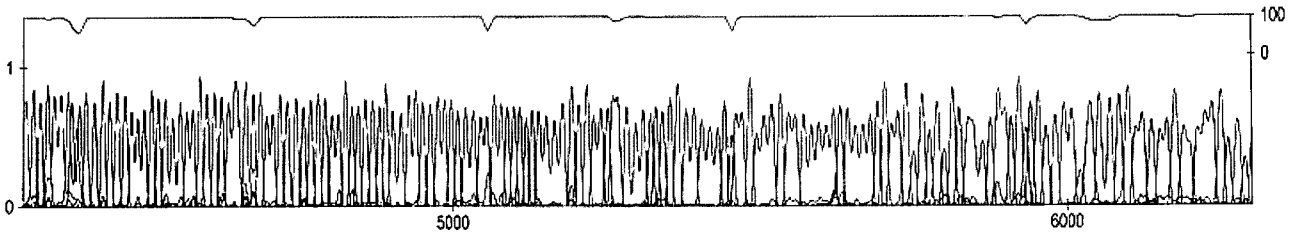


그림 4.
DNA와 drag-tag이 연결된 물질의 저항과 전하

3. 이상적인 ELFSE drag-tag 조건 및 drag-tag 후보 물질들

높은 효율의 DNA 분리를 위해서는 저항 물질 (drag-tag)의 물성이 대단히 중요한데, 이상적인 저항 물질 (drag-tag)의 조건으로는 균일하고 (monodisperse), 분자량이 크며 (large frictional drag), 수용성이고 (hydrophilic), 전기적으로 중성인 물질이어야 한다[13,14]. 각각의 조건에 대하여 간단히 살펴보면, 먼저 ELFSE 반응은 수용액 상태에서 이루어지므로 저항 물질은 수용성 이어야 하고, DNA의 분리 효율을 높이기 위해서는 저항 물질의 저항력 (drag force) 이 커야 하므로 분자량이 큰 물질이 이상적인 저항 물질이라 생각될 수 있다. 또한 DNA를 크기에 따라 분리하기 위해서는 DNA fragments와 연결된 저항 물질은 균일한 크기의 물질이어야 한다. 만일 저항 물질이 균일하지 않다면 일정한 크기의 DNA에 대하여 저항 물질의 비균일성 때문에 서로 다른 이동 속도를 갖게 되므로 DNA의 크기에 따른 분리가 불가능할 것이다. 저항 물질의 전하도 매우 중요한 결정 요소이다. 만일 저항 물질이 (-) 전하를 갖는 물질이라면 DNA와 같은 방향으로 이동하려는 driving force 때문에 큰 저항력 (drag force) 을 가지지 않게 된다. 반면에 저항 물질이 (+) 전하를 갖는 물질이라면 DNA와 저항 물질간의 정전기적 인력 (electrostatic interaction) 이 발생해 서로 엉키게 되어 그 결과 DNA가 더 이상 "free-draining coil" 구조를 갖지 않게 되므로 DNA의 크기에 의한 분리가 불가능하게 된다. 따라서 전기적으로 중성인 물질이 가장 이상적인 저항 물



질이라 생각될 수 있다. 이 외에도 이상적인 저항 물질의 조건으로는 DNA와 1:1 결합을 해야 하고, 모세관 표면과의 비특성 반응 (nonspecific interaction) 이 적은 물질이어야 한다는 점이다.

4가지 후보 물질들이 ELFSE 저항 물질로 연구되어 왔다. (i) 화학적으로 합성된 고분자; (ii) 펩타이드 합성기 (peptide synthesizer) 에 의하여 생산된 oligopeptide (또는 그 유사체); (iii) natural protein; 그리고 (iv) nonnatural protein (repetitive protein) 들이 후보 저항 물질들이다. 화학적으로 고분자 물질을 합성할 경우 오직 하나의 분자량만을 갖는 고분자 물질을 얻기는 실제로 불가능하고, 어떤 분자량 분포를 갖는 고분자 물질을 얻게 된다. 분자량 분포에 기인한 저항 물질의 비균일성은 높은 해상도의 DNA 분리 효과를 기대할 수 없게 하는데, 예로써 polydispersity index (M_w/M_n) 가 1.01 인 비교적 균일한 물질이라 여겨지는 polyethylene glycol (PEG) 도 DNA sequencing 의 저항 물질로 사용되기에는 부적절했다 [15]. 펩타이드 합성기에 의하여 생산되고 HPLC로 정제된 oligopeptide (또는 그 유사체) 의 경우에는 작은 크기의 DNA 분리에는 효과적이었으나 합성기로 합성할 수 있는 oligopeptide의 제한된 크기 때문에 긴 DNA의 sequencing 에 이용되기에는 부적절했다[16, 17]. Natural protein을 ELFSE 저항 물질로 이용한 연구는 캐나다의 Ottawa 대학 연구팀을 중심으로 수행되었으며, 전기적으로 중성 단백질인 streptavidin (natural protein) 을 저항 물질로 사용하여 18분에 ~110 base 정도의 DNA sequencing 결과를 얻을 수 있었다. 그러나 전체적인 streptavidin의 전하는 중성이었으나, (+) 와 (-) 전하를 띠는 아미노산들이 혼재되어 있는 단백질 표면의 국부 전하 (localized surface charge) 와 단백질 접힘 현상 (folding) 때문에 110 base 이상의 높은 sequencing 효율을 기대하

기 어려웠다[18]. 최근에는 이와 같은 natural protein의 문제점을 해결하고자, 전기적으로 중성이고 수용성인 아미노산만으로 이루어진 (예로써 Gly, Ser 등) nonnatural protein (repetitive protein) 을 저항 물질로 이용한 연구가 진행 중에 있다[19]. 유전자 재조합 기술을 사용하여 반복적인 아미노산 code를 갖는 DNA multimer를 클로닝하고 이를 대장균과 같은 미생물에서 발현한 후 정제함으로써 이상적인 저항 물질이 요구하는 조건에 가장 가까운 물질을 얻을 수 있었다. 비록 nonnatural protein 저항 물질의 생산이 비교적 복잡한 과정을 요구하고 또 해결해야 할 어려움이 아직 남아 있지만, 이 어려움을 극복한다면 높은 효율의 DNA sequencing 이 가능하리라 기대되고 있다.

결론

ELFSE에 적절한 저항 물질이 개발되어 이를 유전자 판독과 같은 DNA sequencing에 실용화 시킬 수 있다면 이 연구는 앞으로 커다란 파급 효과를 불러일으킬 것이다. 이 방법은 수용액 상에서 수행되므로 polymer gel을 이용한 기존의 방법보다 최소 5 배 이상의 분석 속도의 향상을 가져올 수 있으며, 또한 점도가 높은 polymer gel을 사용하지 않으므로 아주 가는 모세관을 이용한 초소형 전기영동 방식에 보다 효과적으로 이용되어 질 수 있을 것이다. 또한 최근 Slater 그룹의 연구 결과에 의하면 전기적으로 완전히 중성이고, streptavidin 보다 10 배 정도 큰 저항력을 갖는 저항 물질이 개발될 경우 ELFSE 분석 방법에 의하여 12 분 내에 ~1300 base 정도의 DNA sequencing 효율이 가능하다고 예측되었다[19]. 이는 기존의 분석 방법 (gel-based electrophoresis) 에 의한 DNA sequencing 효율인 약 1 시간에 ~700 base 보다 크게 향상된 값으로서,

저항력이 강하고 균일한 저항 물질의 개발이 선행되어질 경우 ELFSE 분석 방법은 기존의 방법을 대체할 수 있는 매력적인 분석 방법이 될 수 있을 것이다. ㉞

참고 문헌

- [1] E.S.Lander, L.M.Lintons, B.Birren, *et al.*, "Initial sequencing and analysis of the human genome", *Nature* 409 (2001) 860-921
- [2] J.C.Venter, M.D.Adams, E.W.Myers, *et al.*, "The sequence of the human genome", *Science* 291 (2001) 1304+.
- [3] S.R.Liu, H.J.Ren, Q.J.Gao, *et al.*, "Automated parallel DNA sequencing on multiple channel microchips", *PNAS* 97 (2000) 5369-5374
- [4] A.Manz, D.J.Harrison, E.M.J.Verpoorte, *et al.*, "Planar chips technology for miniaturization and integration of separation technique into monitoring systems-Capillary electrophoresis on a chip", *Journal of Chromatography* 593 (1992) 253-258
- [5] S.C.Jacobson, R.Hergenroder, L.B.Koutny, J.M.Ramsey, "High-speed separation on a microchip", *Analytical Chemistry* 66 (1994) 1114-1118
- [6] A.T.Woolley, R.A.Mathies, "Ultra-high-speed DNA fragment separations using microfabricated capillary array electrophoresis chips", *PNAS* 91 (1994) 11348-11352
- [7] L.S.Lerman, H.L.Frisch, "Why does the electrophoretic mobility of DNA in gels vary with the length of the molecule?", *Biopolymers* 21 (1982) 995-997
- [8] B.M.Olivera, P.Baine, N.Davidson, "Electrophoresis of the nucleic acids", *Biopolymers* 2 (1964) 245-257
- [9] G.W.Slater, G.Drouin, "Why can we not sequence thousands of DNA bases on a polyacrylamide gel?", *Electrophoresis* 13 (1992) 574-582
- [10] NHGRI Workshop entitled "Sequencing and Resequencing the Biome"
- [11] J.Noolandi, "A new concept for sequencing DNA by capillary electrophoresis", *Electrophoresis* 13 (1992) 394-395
- [12] P.Mayer, G.W.Slater, D.Drouin, "Theory of DNA sequencing using free-solution electrophoresis of protein-DNA complexes", *Analytical Chemistry* 66 (1994) 1777-1780
- [13] R.J.Meagher, J-I.Won, L.C.McCormick, *et al.*, "End-labeled free-solution electrophoresis of DNA", *Electrophoresis* 26 (2005) 331-350
- [14] G.W.Slater, C.Desruisseaux, S.J.Hubert *et al.*, "Theory of DNA electrophoresis: A look at some current challenges", *Electrophoresis* 21 (2000) 3873-3887
- [15] W.N.Vreeland, C.Desruisseaux, A.E.Karger, *et al.*, "Molar mass profiling of synthetic polymers by free-solution capillary electrophoresis of DNA-polymer conjugates", *Analytical Chemistry* 73 (2001) 1795-1803
- [16] W.N.Vreeland, A.E.Barron, "Free-solution capillary electrophoresis of polypeptoid-oligonucleotide conjugates" Abstr. Pap. Am. Chem. Soc. 219 (2000), 555-556
- [17] W.N.Vreeland, R.J.Meagher, A.E.Barron, "Multiplexed high-throughput genotyping by single-base extension and End-Labeled Free-Solution Electrophoresis (ELFSE)", *Analytical Chemistry* 74 (2002) 4328-4333
- [18] H.Ren, A.E.Karger, F.Oaks *et al.*, "Separation DNA sequencing fragments without a sieving matrix", *Electrophoresis* 20 (1999) 2501-2509
- [19] J-I.Won, R.J.Meagher, A.E.Barron, "Protein polymer drag-tags for DNA separations by end-labeled free-solution electrophoresis", *Electrophoresis* 26 (2005) 2138-2148