In situ Monitoring of Poly(L-glutamine)-Bound dsDNA Selection on a Quartz-Crystal Microbalance

Hiroyuki Furusawa, Yasuhito Ebara, and Yoshio Okahata*
Department of Biomolecular Engineering, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology,
4259 Nagatsuda, Midori-ku, Yokohama 226-8501

(Received May 19, 1999; CL-990404)

We applied a highly sensitive quartz-crystal microbalance as a device of dsDNA *in vitro* selection. When poly(L-glutamine)-immobilized QCM was employed, both the selection and evaluation processes could be monitored *in situ* as mass changes, and the dsDNAs having the continuous sequences of 3-5 AT base-pairs were selected after the 4th cycle from a random dsDNA pool.

Recently, systematic evolution of ligands by exponential enrichment (SELEX) or *in vitro* selection was devised for the identification of high-affinity oligonucleotide ligands to target molecules such as DNA binding proteins.[1-3] It consists of selection of target-binding sequences from a pool of random sequence-DNAs, amplification of selected sequences by PCR methods and repetitions of this process to enrich for target-bound sequences.[4] These selections have been carried out on a filter paper or a gel column on which target molecules are immobilized with radio-isotope or fluorescent labeled random DNAs or RNAs. In these conventional methods, it is not easy *in situ* to monitor the selection process and to evaluate the binding behavior of the selected DNA to the target molecule.

We newly apply a highly sensitive 27 MHz QCM to in vitro selection of dsDNAs that binds selectively to poly(L-glutamine) immobilized on a QCM plate to investigate interactions between a helical polypeptide and dsDNA. A 27 MHz QCM is very sensitive mass measuring device in which the mother frequency of the QCM decreases linearly with increase of the mass on the electrode in nanogram level (-1 Hz = 0.62 ng/cm^2).[5] The QCM is expected to have many advantages as an *in vitro* selection devise. Selection process can be followed by mass changes due to the binding of random DNAs on a QCM, as well as evaluation process whether the selected DNA can bind to a target molecule or not, on the same device without labeling any radio-isotope or fluorescent molecule. The reason why poly(L-glutamine) was chosen as a simple target molecule is that the amide residues of L-glutamine or L-asparagine are expected to recognize AT base-pairs, because those residues in DNA-binding proteins are found frequently from X-ray crystallography to interact with AT base pairs.[6]

Oligonucleotides (R90) having 5'-ACTATCCGACTG-GCACCGAT-(N)$_{50}$-CTAGGCGTTCCGGTCATTAG-3' were prepared by phosphoroamidite method with ARK32 as a DNA synthesizer and purified by an oligo purific column (Sawady Technology, Co.). Random sequences of 50 nucleotides (N)$_{50}$ were synthesized by a mixture of each phosphoramidate with molecular ratios of 3:3:2:2 (A:C:G:T), considering differences of each phosphoramidate reactivity.[7] Other oligonucleotides of a PCR primer were prepared as a same manner (P1: 5'-ACTATCCGACTGGCACCGAT-3' and P2: 3'-GAT-CCGCAAGGCCAGTAATC-5'). On Au electrode (diameter: 2.5 mm) of a commercially available 27 MHz QCM (a quartz diameter: 8 mm, Showa Crystal, Co., Chiba), bis(succiimidyl)-3,3'-dithiodipropionate was immobilized by using Au-S interactions. N-terminal of poly(L-glutamine) (average Mw: 6300, average DP: 20-100, Sigma, Co.) was covalently bonded to the activated carbonyl groups by using a water-soluble carbodiimide. The immobilization amount of poly(L-glutamine) was calculated to be 30 ng/cm^2 (5 pmol). Poly(L-glutamine was confirmed from FT-IR spectra to exist as α-helix structures in 60% on a QCM. The procedure of in vitro selection of poly(L-glutamine)-bound dsDNA from random DNAs is shown

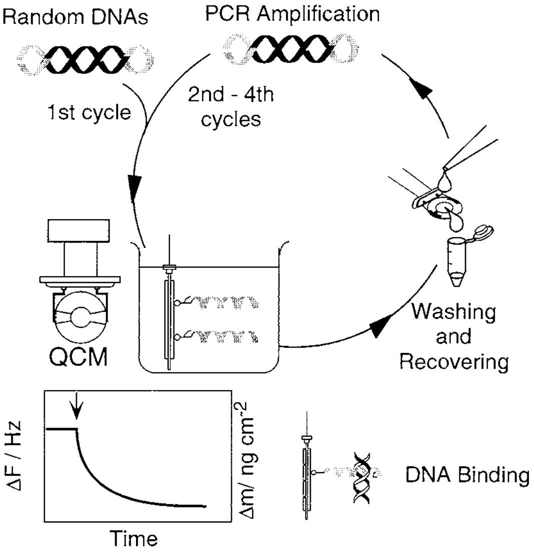

Figure 1. Schematic illustration of *in vitro* selection of dsDNA on a poly(L-glutamine)-immobilized QCM.

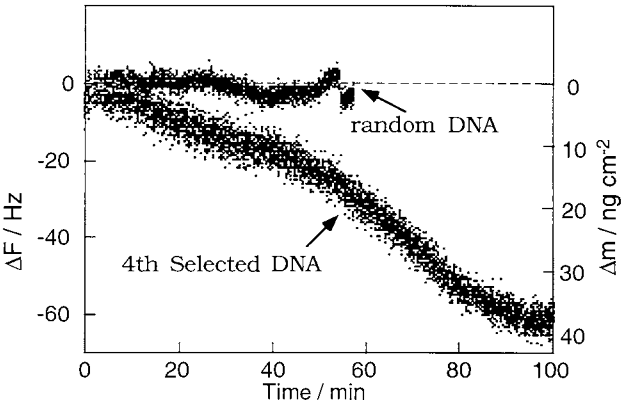

Figure 2. Evaluation of the 4th selected dsDNAs on a poly(L-glutamine)-immobilized QCM (10 mM Tris-HCl, pH 7.8, 0.2 M NaCl, 25 °C).

Table 1 . Sequencing of dsDNAs after 4th selection

10 20 30 40 50

5'-CCACA CTTCA CGGTG TATTC TCAAC CCAGC TATAA CATTG CTATC GCAGG-3'
3'-GGTGT GAAGT GCCAC ATAAG AGTTG GGTCG ATATT GTAAC GATAG CGTCC-5'

5'-CCATT CCTCG TTTTA CCTCA ATTCT CCCAC ACGCC TCTAG TAGTC ATGGC-3'
3'-GGTAA GGAGC AAAAT GGAGT TAAGA GGGTG TGCGG AGATC ATCAG TACCG-5'

5'-ACCGG CCTTG CTAAC TGGTA ATATA CCCCC TTTCA TTCAG CTTGT TCACG-3'
3'-TGGCC GGAAC GATTG ACCAT TATAT GGGGG AAAGT AAGTC GAACA AGTGC-5'

5'-AGTTC GGCAT CTACT CTGCT TTTTT TCATC TTACA AGCTT AATAG GTGCG-3'
3'-TCAAG CCGTA GATGA GACGA AAAAA AGTAG AATGT TGGAA TTATC CACGC-5'

5'-CGCTG CGTAC CCACT ATGAT TGATG TTTCA CGCTC TCGAC CTCTC GCTAC-3'
3'-GAGAC GCATG GGTGA TACTA ACTAC AAAGT GCGAG AGCTG GAGAG CGATG-5'

5'-CCCAG CTATT ACTTT ATTTT TTTTT TTTTT CCAAT ATATC TCTCG GCGG-3'
3'-GGGTC GATAA TGAAA TAAAA AAAAA AAAAA GGTTA TATAG AGAGC CGCC-5'

5'-CCATA ACGTA GATTT ACATT TATAC TTTCC TTCGA CCCGC TCACA GAGA-3'
3'-GGTAT TGCAT CTAAA TGTAA ATATG AAAGG AAGCT GGGCG AGTGT CTCT-5'

5'-CCAGT TTTCG AGCGA GATAC ACCGT TCACT GATAA CTCTC TTCAT CGTTG-3'
3'-GGTCA AAAGC TCGCT CTATG TGGCA AGTGA CTATT GAGAG AAGTA GCAAC-5'

5'-ACCCA CTCCT GGATC TTTCG AAATA ACTTC TATCG TTTTA CTGCG CTGTC-3'
3'-TGGGT GAGGA CCTAG AAAGC TTTAT TGAAG ATAGC AAAAT GACGC GACAG-5'

5'-CCCGC GTGCA TATCT CTTTT GTTTC TTTTT TTTTT ACACT CAGAA CCCAC C-3'
3'-GGGCG CACGT ATAGA GAAAA CAAAG AAAAA AAAAA TGTGA GTCTT GGGTG G-5'

5'-GCGCG CCTGA ACAGA TTTTT TATCC CGGTT GATGA CAAAT TAACA TTGTG-3'
3'-CGCGC GGACT TGTCT AAAAA ATAGG GCCAA CTACT GTTTA ATTGT AACAC-5'

5'-GCCTT AAGGG TCATG CGTTC GTTTA TAATA ATCGA TACAG TCCTG CTCGG-3'
3'-CGGAA TTCCC AGTAC GCAAG CAAAT ATTAT TAGCT ATGTC AGGAC GAGCC-5'

5'-CGCAA CTGCT ATTTA ACACA AACAA ATAAC CGCAA CTAAC CCAGA TTGG-3'
3'-GCGTT GACGA TAAAT TGTGT TTGTT TATTG GCGTT GATTG GGTCT AACC-5'

5'-TCTCC CTCGG ACCCA TTTTT TAAGC GGAAA TCTGC TCACT TCGCG TCGA-3'
3'-AGAGG GAGCC TGGGT AAAAA ATTCG CCTTT AGACG AGTGA AGCGC AGCT-5'

5'-CCTCG TTACT TCTAC TTTAC TAATA AACCT CGTAA ATGCC TCGGC CGAC-3'
3'-GGAGC AATGA AGATG AAATG ATTAT TTGGA GCATT TACGG AGCCG GCTG-5'

5'-GTCCC CAAGA CTGCG GCTCC TAATT CTGCT AATTC TACAA GGCGC CCCC-3'
3'-CAGGG GTTCT GACGC CGAGG ATTAA GACGA TTAAG ATGTT CCGCG GGGG-5'

5'-GGCAT ACTCA CCCTT CATTA TTTAC GAACC ATATT CCTCA CTTTC GC-3'
3'-CCGTA TGAGT GGGAA GTAAT AAATG CTTGG TATAA GGAGT GAAAG CG-5'

5'-CCAAT TGGGA GTCTT TTTTT TTTCT CCTAC GCGGT GCCTT ACGTC GC-3'
3'-GGTTA ACCCT CAGAA AAAAA AAAGA GGATG CGCCA CGGAA TGCAG CG-5'

5'-CAAAC TCTCC CCACA CACCC AGCAA CAACA TGCCA AGCGG TGTC-3'
3'-GTTTG AGAGG GGTGT GTGGG TCGTT GTTGT ACGGT TCGCC ACAG-5'

5'-TGGGT ACCGC GAATA TACCT ATGTC ATTTC ATCAA CCTTG CACTC ACCGT-3'
3'-ACCCA TGGCG CTTAT ATGGA TACAG TAAAG TAGTT GGAAC GTGAG TGGCA-5'

in Figure 1. The poly(L-glutamine)-immobilized QCM was immersed into a R90 dsDNA pool solution (*ca.* 1 μg/2 mL, *ca.* 10 nM, 10 mM Tris-HCl, pH 7.8, 0.2 M NaCl) and incubated for 3 h at room temperature. The picked-up QCM plate was slightly rinsed with the binding buffer and washed out with 20 μl of distilled water for 3 min (4 times) to recover the poly(L-glutamine)-bound DNAs. The recovered DNA was amplified by 20-25 cycles of PCR (polymerase chain reaction, 1 cycle:94°C for 45 s and 72 °C for 2 min) to *ca.* 10 nM and used as the DNA for the next round of selections. The selection was repeated 4 cycles. The binding period was decreased as 3 h, 3 h, 30 min, and 5 min and the rinsing period was increased as 1, 1, 15 and 30 min, for the 1st, 2nd, 3rd, and 4th selection, respectively. When the QCM was picked out from the binding solution, the frequency decrease (mass increase) was observed clearly with proceeding selections from 1st to 4th cycle. The frequency changes, however, were not so large to qualify the binding amount (data not shown).

Figure 2 shows typical frequency changes when the selected DNA was injected to the poly(L-glutamine)-immobilized QCM. The first random dsDNA (20 nM of R90) hardly bound to poly(L-glutamine) on a QCM. On the contrary, the dsDNA after the 4th selection (20 nM) was observed to decrease frequencies (increase mass) and the binding amount was calculated to be 37 ng/cm^2 (0.3 pmol). The binding amounts saturated with increase of the concentration of the 4th selection DNA in this solution. The binding constant could be roughly estimated to be 10^7 M^{-1}. The 4th selected DNA hardly bound to the poly(L-serine)- or polyglycine-immobilized QCM, as well as the bare gold surface of the QCM or the carboxyl group-immobilized QCM plate. Thus, the obtained dsDNA is expected to have the specific DAN sequences to recognize glutamine residues.

The 4th selected DNA was cloned by the T/A cloning method and the nucleotide sequences in random region of 20 clones were determined by a standard dideoxynucleotide method. The obtained sequences of a random region (50 bp) are summarized in Table 1, in which continuous base pairs are hi-lightened and the continuous A·T base pairs are underlined. In many sequences, the continuous A·T base-pairs were found frequently as compared with the continuous G·C base-pairs. The probabilities of the continuous $(A\cdot T)_{3\text{-}5}$ or $(G\cdot C)_{3\text{-}5}$ base-pairs were calculated and shown in Table 2. Probability of random DNA was obtained by sequencing their 10 clones. Probabilities of $(A\cdot T)_n$ and $(G\cdot C)_n$ in random DNA should be consistent within experimental errors, although they were slightly scattering. The probabilities of the continuous A·T base pairs of the selected DNAs were apparently higher than those of the first random DNA (R90), and the probabilities of the continuous G·C base pairs were almost equal to both the selected and random DNAs. Probability of alternative sequences of A/T was also higher than the alternative G/C sequences. These results indicate that the continuous A·T base pairs are selected as a binding site of poly(L-glutamate), consisting with the results of X-ray crystallography that amide residues of glutamine in DNA-binding proteins interact with 6-N and 7-N of adenine moieties of DNAs. The continuous 3-5 base-pairs are reasonable to recognize glutamine residues on polymer chains, because an α-helix peptide can interact with 3-5 bases when the helix binds to the major groove of DNA duplex.[6] Double stranded oligonucleotides of $d(A\cdot T)_{20}$ and $d(G\cdot C)_{20}$ were synthesized separately and evaluated the binding behavior to the poly(L-glutamine)-immobilized QCM. As expected, $d(A\cdot T)_{20}$ could bind selectively to poly(L-glutamine), as compared with $d(G\cdot C)_{20}$ (data not shown).

Table 2. Probabilities of continuous A·T or G·C base pairs on DNAs

Continuous number of base pairs	Probability / %	
	4th selection DNA	random DNA
$(A\cdot T)_3$	7.5	1.7
$(A\cdot T)_4$	4.2	0.2
$(A\cdot T)_5$	3.0	0
$(G\cdot C)_3$	3.0	2.3
$(G\cdot C)_4$	0.8	0.9
$(G\cdot C)_5$	0.3	0.2

In conclusion, glutamine residues are confirmed to be important to recognize A·T base-pairs of DNA-binding proteins. The QCM technique will become a simple and useful tool as *in vitro* selection of DNAs, since both the selection and evaluation processes can be monitored directly without labeling any radioisotope or fluorescent probes.

References

1 M. Famulok and J. W. Szostak, *Angew. Chem. Int. Ed. Engl.*, **31**, 979 (1992).
2 J. W. Szostak, *Trends Biochem. Sci.*, **17**, 89 (1992).
3 S. E. Osborne and A. D. Ellington, *Chem. Rev.*, **97**, 349 (1997).
4 W. E. Wright and W. D. Funk, *Trends Biochem. Sci.*, **18**, 77 (1993).
5 a) K. Niikura, H. Matsuno, and Y. Okahata, *J. Am. Chem. Soc.*, **120**, 8537 (1998). b) Y. Okahata, K. Niikura, Y. Sugiura, M. Sawada, and T. Morii, *Biochemistry*, **37**, 5666 (1998).
6 M. Suzuki, *Structure*, **2**, 317-326 (1994).
7 C. Wilson and J. W. Szostak, *Nature*, **374**, 777 (1995).