Mathematical Biosciences 247 (2014) 59–68

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs

Estimation of statistical binding properties of ligand population during *in vitro* selection based on population dynamics theory

Takuyo Aita [a,*], Koichi Nishigaki [a], Yuzuru Husimi [b]

[a] Graduate School of Science and Engineering, Saitama University, 255 Shimo-okubo, Saitama 338-8570, Japan
[b] Innovation Research Organization, Saitama University, 255 Shimo-okubo, Saitama 338-8570, Japan

ARTICLE INFO

Article history:
Received 22 October 2012
Received in revised form 9 September 2013
Accepted 31 October 2013
Available online 15 November 2013

Keywords:
Aptamer
In vitro evolution
Directed evolution
SELEX
Binding constant
Monitoring method

ABSTRACT

During *in vitro* selection process, it is very valuable to monitor the binding properties of the ligand population in real time, particularly the population average of the association constant in the population. If this monitoring can be realized, the selection process can be controlled in a rational way. In this paper, we present a simple method to estimate the binding properties of the ligand population during *in vitro* selection. The framework of the method is as follows. First, the number of all the collected ligand molecules, which are eluted after incubation and washing, is measured. Ideally, this number corresponds to the number of all the ligand molecules bound with the target–receptor or other materials in a test tube. This measurement is performed through several successive rounds of selection. Second, the measured numbers of molecules are subjected to a theoretical analysis, based on the mathematical theory of population dynamics in the selection process. Then, we can estimate the probability density of the binding free energy in the ligand population. The validity of our method was confirmed by several computer simulations based on a physicochemical model.

1. Introduction

In the last two decades, *in vitro* selection has been developed as one of the essential processes in evolutionary molecular engineering; serving for the study on the origin of life, the development of advanced biomolecules, etc. [1–5]. Many studies have dealt with the ligand-receptor binding system. A ligand library is prepared by random synthesis of biopolymers. The ligand library is mixed with the target molecules immobilized on the magnetic beads in a test tube, and incubated up to the equilibrium state. A part of the ligand population is bound with the target (specific binding) or non-target materials (non-specific binding). Free ligands and a part of the bound ligands are removed by the washing process which is repeated several times. Then, the ligands bound with the target and non-target are eluted from the mixture. The collected ligands are amplified and subjected to the following rounds of the selection cycle.

In our previous paper [6], we focused on the concentration of the free target–receptor molecules, [R], which governs the selection dynamics drastically, and examined its optimal concentration to perform the rapid selective enrichment. We concluded that the concentration [R] should be equal to the dissociation constant of the fittest sequence (with the highest affinity) in the library. In this previous study, we found the necessity for monitoring the binding properties of the ligand population in real time, particularly the population average of the association constant in the population. If this monitoring can be realized, the selection process can be controlled in a rational way [7–9].

In this paper, we present a simple method to estimate the binding properties of the ligand population during *in vitro* selection. The framework of the method is as follows. First, the number of all the collected ligand molecules, which are eluted after incubation and washing, is measured by using the quantitative polymerase chain reaction (PCR) or other methods. Ideally, this number corresponds to the number of all the ligand molecules bound with the target–receptor or other materials in a test tube. This measurement is performed through several successive rounds of selection. Second, the measured numbers of molecules are subjected to theoretical analysis, based on the mathematical theory of population dynamics in the selection process. Then, we can estimate the probability density of the binding free energy in the ligand population and the population average of the association constant in the population (=statistical moments of association constant in the population).

Several methods for monitoring the statistical binding properties of a ligand population have been developed. For example, the filter-binding assay combined with radioactivity-labeled ligand molecules [7] and flow cytometry combined with FITC (fluorescein isothiocyanate)-labeled ligand molecules [8]. Horisawa et al. [10]

* Corresponding author.
E-mail address: taita@mail.saitama-u.ac.jp (T. Aita).

0025-5564/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.mbs.2013.10.012

Nomenclature

Notation of main variables and parameters

M	diversity, that is the number of different sequences in a ligand library
N and N'	library size, that is the number of ligand molecules in a ligand library, and the number of all the collected ligand molecules after a single selection round, respectively
s	serial number of ligand sequences in a ligand population ($s = 1, 2, \ldots, M$)
$s = 1$	fittest sequence, that is a particular ligand sequence with the highest affinity in a ligand population
$\tilde{q}$	$\equiv N'/N$ with no target molecules
K_{ds} (mol/l) and K_{as} (l/mol)	dissociation constant of the ligand with sequence s with the receptor molecules and association constant, respectively $K_{d1} < K_{d2} < K_{d3} < \cdots < K_{dM}$
$X \equiv \ln K_d$ and $X_s \equiv \ln K_{ds}$	non-dimensional binding free energy, and that for sequence s, respectively
r	serial number of rounds and generations through a selection process ($r = 1, 2, 3, \ldots$). The ligand population resulting after the rth round of selection is defined as being in the rth generation
$[R]_r$ (mol/l)	molar concentration of free target–receptor molecules at equilibrium state for the rth round
ω (mol/l)	$\equiv 1/k_{as}t$, where k_{as} is the association rate constant for sequence s and t is the wash time
D_s	dilution factor for sequence s in washing process
$A_s(r)$	$\equiv (D_s - \tilde{q})/(K_{ds} + [R]_r)$
$x_s(r)$	mole fraction of ligand molecules with sequence s in the rth generation
$\psi_r(X)$	probability density function of X over all the ligand sequences in the rth generation
$\langle Q \rangle_r$	the average of Q over all the ligand sequences in the rth generation $\langle Q \rangle_r \equiv \sum_{s=1}^{M} Q_s x_s(r) = \int_{-\infty}^{\infty} Q(X)\psi_r(X)dX$

applied the quantitative real-time PCR for evaluation of the sequences obtained in *in vitro* selection for monitoring the process of enrichment of selected clones in each round of selection. While their aim is to quantify the amount of each of the selected clones, our aim is to quantify the amount of population of ligand molecules bound with the target–receptor or other materials in a test tube. Even if ligand molecules are peptides or proteins, this quantification is possible by using genotype(DNA or RNA)-phenotype(protein) linkage techniques [11,12].

This paper is constructed as follows. In Section 2, we introduce a model of *in vitro* selection. The model is based on the law of mass action. In Section 3, we describe how the transition of ligand population occurs through successive selection rounds. In Section 4, we describe a simple method to estimate binding properties of the ligand population, that is the probability density of the binding free energy in the ligand population. The validity of our method was confirmed by several computer simulations based on a physicochemical model.

2. Model of *in vitro* selection

We consider an ensemble of ligand peptides that bind to a single kind of target receptor immobilized on the magnetic beads in a test tube. It is possible for each ligand to bind to non-target materials in the test tube (non-specific binding). We denote: L_s, free ligand molecule with sequence s; R, free target–receptor molecule; $L_s \cdot R$, molecular complex of L_s and R; U, non-target materials (such as wall of the test tube and other substances); $L_s \cdot U$, ligand L_s bound with U, respectively. For each case, [X] and $[X]^\circ$ represent the numerical value of molar concentration (mol/l) of the molecule "X" at equilibrium and that in the preparation stage, respectively. $\{L_s\}$ represents a set of heterogeneous ligands over all sequences.

2.1. Protocol of in vitro selection

2.1.1. Preparation of a ligand library

A library of ligand molecules with a large variety of sequences is prepared by random synthesis. In the resulting library, the number of all the ligand molecules is denoted by N (="library size"), and the number of different sequences is denoted by M ("diversity"). For example, $N = 10^8 \sim 10^{14}$ molecules and $M = 20^5 \sim 20^{10}$ sequences in typical cases. In addition, let n_s be the number of ligand molecules with sequence s, that is $\sum_s n_s = N$, where $\sum_s$ means the sum over all the M sequences. Particularly, each peptide sequence in the initial library is generated by random DNA synthesis and we neglect the bias originating from the genetic code table for simplicity.

2.1.2. Pre-screening of the ligand library

The initial ligand library is subjected to a pre-screening process without target protein receptors R but with non-target materials U. Then, the library after the pre-screening does not contain the ligand molecules that show specific binding with U.

2.1.3. Selective enrichment of the ligand library

The ligand library ($\{L_s\}$) is mixed with the target receptor molecules (R) and the mixture is incubated to an equilibrium state. In cases where the target molecules are immobilized in a local area in the bulk, the mathematical framework in these cases is equivalent to that in cases where the same amount of the target molecules are distributed uniformly over the bulk [6]. At equilibrium, a part of the ligand library is bound with the receptors ($\{L_s \cdot R\}$) (specific binding), or bound with non-target materials ($\{L_s \cdot U\}$) (non-specific binding) in the test tube. Next, the magnetic beads are washed with a buffer by several times. Finally, the ligand molecules are collected by separating them from the binding complex ($\{L_s \cdot R\} + \{L_s \cdot U\} \rightarrow \{L_s\} + R + U$) and eluting from the mixture. We denote the number of the collected ligand molecules with sequence s by n'_s, and we denote the number of all the collected ligand molecules by $N' : \sum_s n'_s = N'$.

The collected ligands are amplified by PCR up to the given library size N. That is, the number of the ligand L_s is amplified up to $n'_s \times N/N'$. The new ligand library is subjected to the next round of selective enrichment. Here, we assume that mutations do not occur in amplification of genotypes by PCR.

2.2. Formulation of the selective enrichment

2.2.1. Equilibrium of binding reaction

The binding-reaction scheme is given by

$$L_s + R \overset{K_{ds}}{\leftrightarrows} L_s \cdot R \quad \text{(for specific binding)} \tag{1}$$

$$L_s + U \overset{\text{randomly}}{\rightarrow} L_s \cdot U \quad \text{(for non-specific binding)}, \tag{2}$$