Chemometrics and Intelligent Laboratory Systems 138 (2014) 133–141

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Review

Experimental designs for the Adair model

J. López-Fidalgo *, M.M. Rodríguez-Hernández

Department of Mathematics, University of Castilla-La Mancha, Institute of Mathematics Applied to Science and Engineering, Avda. Camilo José Cela 3, 13071 Ciudad Real, Spain

ARTICLE INFO

Article history:
Received 31 March 2014
Received in revised form 12 July 2014
Accepted 18 July 2014
Available online 30 July 2014

Keywords:
Adair equation
Free ligand
Macro-molecule reactions
Optimal experimental design
Saturation rate

ABSTRACT

The Adair equation is used to model biological macro-molecule reactions. This equation relates the saturation rate to the free ligand concentration. But, the latter is not a variable completely under the control of the experimenter. The ligand is a random variable depending on an initial ligand added by the experimenter, which can be designed, but the dependence of the saturation rate on the initial ligand has not been considered in the literature. In this paper a transformed model based on the Adair model of first order (monomer) is derived in order to obtain a proper model that depends on the initial ligand. This model will allow proper fitting and optimal designs using the initial ligand. It will be called the transformed Adair model (TAM). Optimal designs as well as seven-point quasi-optimal designs forced to follow a harmonic, geometric or uniform progression, are computed. The parameters are estimated and compared for simulated data from these designs. A sensitivity analysis against the choice of nominal values of the parameters is also performed for the TAM. The analytic version of the transformed model is only possible for the first class model. But the good efficiencies of the optimal designs obtained directly from the monomer model for fitting the TAM justify doing something similar for the second order model (dimer). Designs were computed numerically in this case.

Contents

1. Introduction . . . 133
2. Deriving the Adair model . . . 134
   2.1. Optimal design background . . . 134
   2.2. Outline of the paper . . . 135
3. Optimal design for the monomer model . . . 135
   3.1. Optimal design for a transformed monomer model . . . 136
   3.2. Suboptimal sequence designs . . . 137
   3.3. Simulations to compare estimates and designs for both models . . . 137
      3.3.1. Example . . . 137
   3.4. Robustness with respect to the choice of the nominal value of the parameter . . . 138
4. Optimal designs for the dimer model . . . 138
5. Discussion . . . 140
Conflict of interest . . . 140
Acknowledgments . . . 140
References . . . 141

1. Introduction

The Adair model is used in chemical reactions once the balance between a biological macro-molecule and a ligand is reached. A ligand is usually a smaller molecule and sometimes another macro-molecule. These types of reactions are usually reversible. Adair [1] proposed a sequential equation to model how the oxygen molecules are bound to subunits of hemoglobin. The hemoglobin is a quaternary structure protein, whose main function is the transport of oxygen. The sigmoidal shape of the curve describing the union of the hemoglobin to the oxygen means that it has a relatively low affinity to capture the first molecule. The increasing slope means that the binding of the first molecule of oxygen facilitates the binding of another oxygen molecule to the second subunit and so on. This is known as binding cooperation. The

* Corresponding author. Tel.: +34 926 295212; fax: +34 926 295361.
E-mail addresses: jesus.lopezfidalgo@uclm.es (J. López-Fidalgo), MMercedes.Rodriguez@uclm.es (M.M. Rodríguez-Hernández).

http://dx.doi.org/10.1016/j.chemolab.2014.07.014
0169-7439/

hemoglobin binds a fourth oxygen molecule with an affinity 100 times greater than the first one. Obtaining the balance constants is a fundamental issue (see e.g. [12]). The known antigen binding, such as viral antigens, bacterial antigens, and drugs, is of common use in biomedical practice through immunoassays. Tetin and Hazlett [15] considered the interaction antibody–ligand and the determination of the balance binding constant with different models, including the Adair model. The protein–ligand models are remarkably analogous to metal–ligand complexation models that are dependent on the initial ligand concentration [12]. However, the positive cooperativity behavior is unique to these biological systems.

In this paper, optimal experimental designs are computed in order to better fit the Adair model. The theory of optimal experimental designs has been developed considerably, especially after finding the celebrated equivalence theorem [9]. This theorem gives a powerful tool to check whether a particular design is D-optimal (equivalently G-optimal). Whittle [16] generalized this to a more general class of optimality criteria. The General Equivalence Theorem (GET) provides also a tool to construct optimal designs using different algorithms [5,17] for the popular Wynn–Fedorov algorithm. A suitable choice of the experimental conditions can improve inferences on the model. The Adair equation properly describes a particular microbiological phenomenon, in which a random error with zero mean needs to be added to condense the uncertainty. The variance of an observation (or the error) may be constant or dependent on the experimental conditions through a mathematical model, e.g. depending on the mean model.

2. Deriving the Adair model

Let Y be the saturation rate of a macro-molecule. This random variable is decomposed into a form depending on the concentration of the free ligand, $[L]$, plus a random error.

A typical experiment to describe the chemical process consists of introducing a solution of concentration of a macro-molecule or a protein, $[P_0]$, into a semi-impermeable dialysis bag. Then, the bag is put in a recipient with initial ligand concentration, $[L_0]$. Fig. 1 (left) shows the initial situation for the monomer case with just one binding site.

Then, the ligand molecules start penetrating the bag but the protein molecules cannot get out. The equilibrium is reached once there is equal free ligand concentration in and out of the bag, say $[L]$. Fig. 1 (right) displays this situation for the monomer. In order to measure this concentration the bag is withdrawn and the free ligand within the container is measured. If the protein has n binding sites and the ligand can bind the protein through 0, 1, 2, …, n sites, the concentration of protein bound to the ligand will be denoted by $[M_n^i]$, with $i = 1, 2, \ldots, n$ the number of sites of the protein bound to the ligand. Thus, in the interior of the bag there is the free ligand, $[L]$, the bound ligand (to the protein or macro-molecule), $[M_n^i]$, $i = 1, 2, \ldots, n$, and the unbound protein, $[M_n^0]$.

Let r be the observed average number of busy sites, that is, the number of moles of the bound ligand per mol of protein, $0 < r < n$. The saturation rate is defined as $Y = r/n \in [0, 1]$. The mean value of the busy sites is

$$r = \frac{[\text{Bound ligand}]}{[\text{Total of macro-molecule}]} = \frac{\left[M_n^1\right] + 2\left[M_n^2\right] + \ldots + n\left[M_n^n\right]}{\left[M_n^0\right] + \left[M_n^1\right] + \left[M_n^2\right] + \ldots + \left[M_n^n\right]} = \frac{K_1[L] + 2K_1K_2[L]^2 + \ldots + nK_1\ldots K_n[L]^n}{1 + K_1[L] + K_1K_2[L]^2 + \ldots + K_1\ldots K_n[L]^n},$$

where the asymptotic value of r is the average number of bound sites in each macro-molecule divided by the number of macro-molecule binding sites, n. Thus, the model can be formulated as

$$Y = \frac{K_1[L] + 2K_1K_2[L]^2 + \ldots + nK_1\ldots K_n[L]^n}{n\left(1 + K_1[L] + K_1K_2[L]^2 + \ldots + K_1\ldots K_n[L]^n\right)} + \varepsilon. \quad (1)$$

The parameters may be estimated using Least Squares Estimates (LSE), Maximum Likelihood Estimates (MLE) or by doing some transformation to linearize the model for the monomer. For more details see [6]. Tetin and Hazlett [15] stressed that either these linear transformations or the deletion of experimental data introduce uncontrolled bias in the estimation process. In order to use these linear equations they suggested using weighting optimization. Throughout this paper the MLE will be used assuming normality and so they are LSE. No linearization by transforming the equation will be considered in this paper.

Monod's equation is a classic microbiological model describing the kinetics of batch microbial growth. Another typical model used in this field to describe the kinetics of fast equilibrium of enzymatic systems and the analysis of data from drugs, neurotransmitters and assays with hormonal receptors is the popular Michaelis–Menten model [11]. The Monod equation has the same form as the Michaelis–Menten equation, but the Monod equation is empirical while the latter is based on theoretical considerations. López–Fidalgo et al. [10] computed optimal designs for population growth models.

The aim of this paper is to compute and compare suitable experimental designs for the Adair model for the saturation ratio with macro-molecules with one and two binding sites (monomer and dimer). An optimal design will provide both, good estimates and savings in the use of experimental resources [5]. The theory of optimal experimental design provides tools for computing good estimates from different points of view. One of the criticisms made to this theory is that optimal designs frequently claim for extreme points not allowing the detection of some features of the data as is the case of possible curvature. This is the reason why in this paper a greater number of different points in the design are forced through regular sequences as classes of harmonic, geometric or uniform progressions.

2.1. Optimal design background

For nominal values of the parameters locally optimal designs will be obtained [3] maximizing the determinant of the Fisher Information Matrix (FIM). For the monomer the information matrix is a scalar and an analytic expression is available for the one-point optimal design. For a dimer two parameters are in the model and therefore two or three points will be needed. For this case the equivalence theorem will be used to compute optimal designs.

An (*exact*) experimental design is a collection of points $[L]_1, \ldots, [L]_N$ from a design space χ. Some of them may be repeated, which means several experiments are realized for the same experimental condition. The total number of observations is N and this number is limited by cost, time or feasibility restrictions. A good experimental design may reduce this number, and so as cost and time, preserving a good estimation of the model. Taking into account that possible replicates of some points are possible a probability measure may be defined identifying the

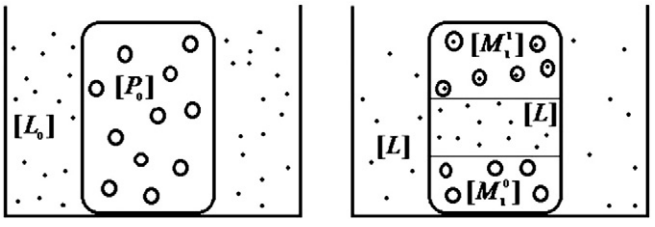

Fig. 1. The initial phase of the experiment (left) and the experiment at equilibrium (right) for the monomer.