Titration Simulations for Understanding of the Binding Equilibrium

Dae Hyup Sohn*

ABSTRACT

The reliability evaluation of the predicted binding constants in numerous models is a challenge for supramolecular host-guest chemistry. Here, I briefly formulate binding isotherm with the derivation of the multivalent equilibrium model for the chemist who wants to determine the binding constants of their compounds. This article gives an in-depth understanding of the stoichiometry of binding equilibrium to take divalent binding equilibria bearing two structurally identical binding sites as an example. The stoichiometry of binding equilibrium is affected by (1) the cooperativity of complex, (2) the concentration of titration media, and (3) the equivalents of guests. The simulations were conducted with simple Python codes.

From a biological point of view, the multivalent binding equilibrium is necessary to study its allosteric effects. These interactions are related to a variety of cellular functions, such as buffering, transport, and signal transduction[1]. Understanding the interaction of 'supramolecules' also requires the knowledge of their properties, such as binding affinity at each site and the effects that binding at one site impacts on binding at other sites for the determination of their binding constants.

How to determine the binding constant with high reliability is an essential point in supramolecular host-guest chemistry. However, these cannot be measured directly in a single titration and are usually used to be estimated by fitting physical changes into several models. Over the years, through the quantitative analysis, binding constants of interests have been

measured. There are several issues to be carefully addressed to ensure the reliability of results. Mathematical approaches might provide a guide for the proper design and analysis of titration experiments for supramolecular host-guest chemistry. Out-of-dated models (i.e., Benesi-Hilderbrand, Scott/Hanse-Woolf, Scatchard plot) that do not provide any reliable parameters by abusing shortcuts like false identifications without proper logics[2].

This article formulates binding isotherm with the derivation of the multivalent equilibrium model. These derivations can provide insights through simulations performed by a numerical approach called the Newton-Raphson method (Supplementary Information). From these simulations, not only the concentration of titration media, but the cooperativity of binding equilibrium also affects its stoichiometry at every titration point. The practicality of the Job plot frequently used in supramolecular host-guest chemistry is also considered.

Derivation of Binding Isotherm for Multivalent Equilibrium Model

The mechanical parameters cannot be measured directly but must be estimated by fitting a quantitative model of binding or structural data. Reasonable binding models that take into account factors such as different binding modes, different binding guest molecules (heterotropic), and co-operation depending on the host's various states, require quite a lot of variables. Determination of the binding model based on the data is necessary for progress to impose significant constraints on the value of the variable to improve the quality of the data. From these reasons, cooperative interactions caused by non-identical binding sites in a host are also possible but not considered in this research. On the multivalent complexation, we usually express the first stepwise binding constant as $K_1 = \frac{[HG]}{[H][G]}$. The mole fraction of HG (f_{HG}) can be presented by rearranging the equation of first stepwise binding constant (K_1) to $[HG] =$

$K_1[G][H]$ and divide both sides by $[H]_t$.

$$f_{HG} = K_1[G]f_H$$

Express the second stepwise binding constant as $K_2 = \frac{[HG_2]}{[G][HG]}$. To calculate the overall binding

constant (β), multiply K_1, and K_2.

$$K_1K_2 = \frac{[HG_2]}{[G]^2[H]}$$

$$[HG_2] = K_1K_2[G]^2[H]$$

$$f_{HG_2} = K_1K_2[G]^2f_H$$

In a similar manner, we could generalize this relationship as

$$K_1K_2\cdots K_n = \frac{[HG_n]}{[G]^n[H]}$$

$$[HG_n] = K_1K_2\cdots K_n[G]^n[H]$$

$$f_{HG_n} = K_1K_2\cdots K_n[G]^nf_H$$

The total concentration of host $[H]_t$ is equal to the sum of the concentration of complexes and

free (unsaturated) host $[H]$, respectively. The expression of total host concentration as $[H]_t = [H] + [HG] + [HG_2] + \cdots + [HG_n]$. Substitute the expression of total host concentration using

$[HG_n]$.

$$[H]_t = [H] + K_1[G][H] + K_1K_2[G]^2[H] + \cdots + K_1K_2\cdots K_n[G]^n[H]$$

$$\frac{[H]_t}{[H]} = 1 + K_1[G] + K_1K_2[G]^2 + \cdots + K_1K_2\cdots K_n[G]^n$$

If you flip both sides to get the fraction of un-saturated host (f_H)

$$f_H = \frac{1}{1 + K_1[G] + K_1K_2[G]^2 + \cdots + K_1K_2\cdots K_n[G]^n}$$

Define $1 - f_H = f_{HG} + f_{HG_2} + \cdots + f_{HG_n}$ as *the mole fraction of complexes,*

$$1 - f_H = \frac{K_1[G] + K_1K_2[G]^2 + \cdots + K_1K_2\cdots K_n[G]^n}{1 + K_1[G] + K_1K_2[G]^2 + \cdots + K_1K_2\cdots K_n[G]^n}$$

$$\therefore f_{HG} + f_{HG_2} + \cdots + f_{HG_n} = \frac{K_1[G] + K_1 K_2 [G]^2 + \cdots + K_1 K_2 \cdots K_n [G]^n}{1 + K_1[G] + K_1 K_2 [G]^2 + \cdots + K_1 K_2 \cdots K_n [G]^n}$$

For the sake of dealing with the physical changes to general binding isotherm in spectroscopic titration, additional derivation of formula should be needed. Through *Beer-Lambert law*, we can express the physical change (Y_{HG_n}) as

$$Y_{HG_n} = f_{HG_n} \cdot \varepsilon_{HG_n} = (K_1 K_2 \cdots K_n [G]^n f_H) \cdot \varepsilon_{HG_n}$$

$$Y_{HG} + Y_{HG_2} + \cdots + Y_{HG_n} = \left[\varepsilon_{HG} K_1[G] + \varepsilon_{HG_2} K_1 K_2 [G]^2 + \cdots + \varepsilon_{HG_n} K_1 K_2 \cdots K_n [G]^n \right] \cdot f_H$$

$$= \left[\frac{\varepsilon_{HG} K_1[G] + \varepsilon_{HG_2} K_1 K_2 [G]^2 + \cdots + \varepsilon_{HG_n} K_1 K_2 \cdots K_n [G]^n}{1 + K_1[G] + K_1 K_2 [G]^2 + \cdots + K_1 K_2 \cdots K_n [G]^n} \right]$$

Now, we can express $Y_{HG} + Y_{HG_2} + \cdots + Y_{HG_n}$ as ΔY

$$\therefore \Delta Y = \sum_{i=1}^{n} Y_{HG_i} = \sum_{i=1}^{n} K_1 K_2 \cdots K_i [G]^i f_H \varepsilon_{HG_i}$$

The determination of the order of a multivalent binding equilibrium by setting up an adequate binding model is an essential task for titration experiments. Here, I express four factors for the stoichiometry in the titration media. These factors can easily visualize the binding equilibrium, even in multivalent binding equilibria, on the same principle.

$$\alpha = \frac{[H]_t - [H]}{[H]_t} = \frac{[HG] + [HG_2]}{[H]_t} = \frac{K_1[G] + K_1 K_2 [G]^2}{1 + K_1[G] + K_1 K_2 [G]^2}$$

$$v = \frac{[G]_t - [G]}{[H]_t} = \frac{1 \cdot [HG_1] + 2 \cdot [HG_2]}{[H]_t} = \frac{K_1[G] + 2K_1 K_2 [G]^2}{1 + K_1[G] + K_1 K_2 [G]^2}$$

To know the fraction of the host participating in the binding equilibrium, define α as the mole fraction of complexes. In brief, the value of α is an expression of the percentage of host occupied with the guest among the total (saturated + unsaturated) host understandably. The symbol v describes the mole fraction of complexes weighted by the number of guests. Utilizing α and v into the expression of the multivalent binding equilibrium yields the

stoichiometry of binding equilibrium. Express overall binding constant (β_{mn}),

$$\beta_{mn} = \frac{[H_m G_n]}{[H]^m [G]^n}$$

$$[H]_t = [H] + m \cdot [H_m G_n], \qquad m = \frac{[H]_t - [H]}{[H_m G_n]}$$

$$[G]_t = [G] + n \cdot [H_m G_n], \qquad n = \frac{[G]_t - [G]}{[H_m G_n]}$$

$$\frac{v}{\alpha} = \frac{[G]_t - [G]}{[H]_t - [H]} = \frac{n}{m}$$

$$\frac{\alpha}{\alpha + v} = \frac{[H]_t - [H]}{[H]_t - [H] + [G]_t - [G]} = \frac{m}{m + n}$$

These induced factors $(\frac{v}{\alpha}, \frac{\alpha}{\alpha + v})$ presents the stoichiometry of the binding equilibrium in titration media. The visualization of these four factors will be mainly discussed.

Titration Simulation for Visualization of the Binding Equilibrium

Mathematical and computational approaches are essential tools to provide reasonable guidelines. Newton-Raphson method is a tool in the numerical analysis that approximates most equations. Sessler, Anslyn, and their colleagues have already introduced this method to calculate binding constants using commercial software[3]. The application of this numerical algorithm is possible to produce the virtual system of titration. Modulating the derivation of binding isotherm on the basis of the host conducts hypothetical titration experiments with a simple code. These can be monitored the concentration of overall molecules at each titration point.
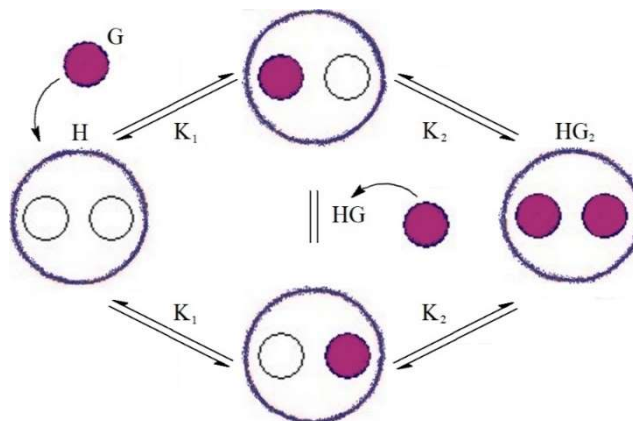
Figure 1. The brief diagram of stepwise guest binding equilibrium with two structurally identical binding sites. The illustration represents the divalent binding equilibrium. (H : host bearing two structurally identical binding sites, G : Guest, HG : first stepwise complexation, HG_2 : second stepwise complexation)

The diagram in **Figure 1** shows the divalent binding equilibrium to be addressed. Model parameters (K_1, K_2, f), first (K_1), second (K_2) stepwise binding constant, and cooperativity factor (f) for divalent binding equilibrium, an indicator of allosteric effect, is shown below[2].

$$f = 4\frac{K_2}{K_1}$$

The cooperativity factor (f) is classified as positive ($f > 1$), none ($f = 1$) and negative ($f < 1$). The value of the cooperativity factor increases as it promotes (positive) or decreases as it inhibits (negative) of the next binding step. Non-cooperative binding means that the host-guest complex changed due to the first stepwise binding does not affect the second stepwise binding.

Theoretical titration simulations illustrated by these four factors ($\alpha, v, \frac{v}{\alpha}, \frac{\alpha}{\alpha+v}$) provide insights for titration experiments in the laboratory. Three kinds of cooperativities on the same overall binding constants (β_{12}) are examined (**Table 1**). Detailed information on the simulations is given in Supplementary Information. The red line represents the equilibrium with negative cooperativity ($f = 0.04$), the black line with non-cooperative ($f = 1$), and the green line with

positive cooperativity ($f = 400$) as below.

	Red (negative-)	Black (none-)	Green (positive-)
f	0.04	1	400
K_1	100,000	20,000	1,000
K_2	1,000	5,000	100,000

Table 1. The simulation studies of divalent binding equilibrium performed by varying the concentrations of the solution media with three different cooperativities.
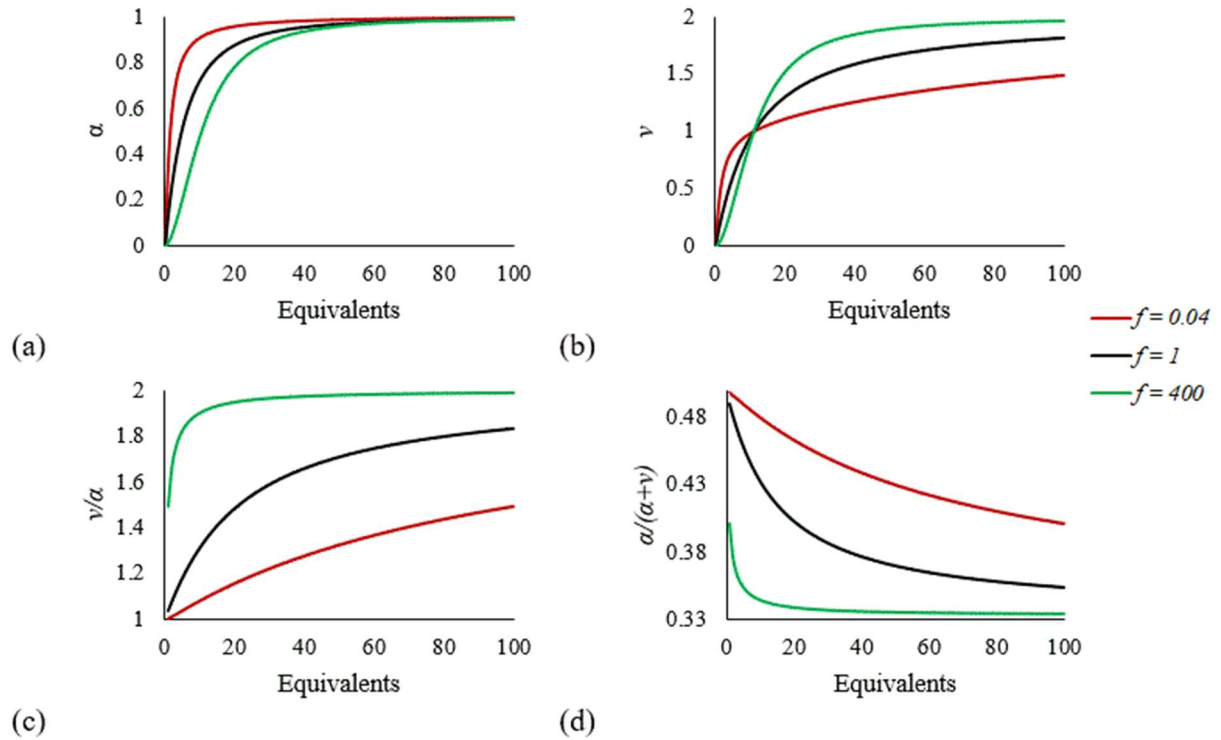

(a) (b) (c) (d)

Figure 2. Titration simulations about four factors $(\alpha, v, \frac{v}{\alpha}, \frac{\alpha}{\alpha+v})$ examine the divalent binding equilibrium of the three different systems.

At first, the simulations are illustrated by four factors $(\alpha, v, \frac{v}{\alpha}, \frac{\alpha}{\alpha+v})$ in **Figure 2**. Before reaching the range of over-saturation in **Figure 2 (a)**, the first stepwise association (K_1) mainly concludes its mole fraction of complexes (α) on the same overall binding constants (β_{12}). Even though hosts have two binding sites, **Figure 2 (b)** shows that v is not appropriate to describe the order of binding equilibrium. From **Figure 2 (c)**, even if it is divalent binding equilibrium, it might not be perfectly H : G = 1 : 2 stoichiometry in the system. From the interpretation of $\frac{v}{\alpha}$, we could know the exact ratio of bound guests to a host. However, even if the host has two

binding sites for guests, the response of the stoichiometry is insufficient to be H : G = 1 : 2 in negative and none cooperative binding equilibrium. Therefore, we must think about stoichiometry does not provide the capability of how many guests can be bound, and the possible binding sites in a host. In other words, obtaining stoichiometry to determine the order of binding equilibrium is not a good option. The variations of stoichiometry in terms of three different cooperativities is shown in **Figure 2 (d)**.
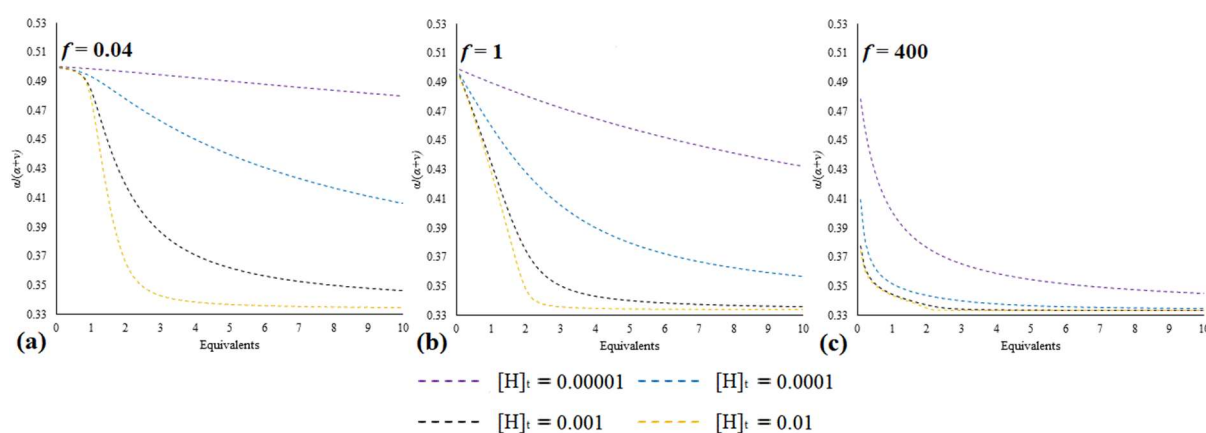

Figure 3. The visualizations of $\frac{\alpha}{\alpha+v}$ at **(a)** $f = 0.04$, **(b)** $f = 1$, and **(c)** $f = 400$ show the changes of stoichiometry over the course of titration experiments in various host concentration. The purple dotted line represents the host concentration of 0.00001 M. The blue dotted line represents the host concentration of 0.0001 M. The black dotted line represents the host concentration of 0.001 M. The yellow dotted line represents the host concentration of 0.01 M.

As an extension of **Figure 2 (d)**, I also conducted simulations by varying the concentration of solution media (**Figure 3**). In previous, Jurczark[4] already confirmed that the interpretation of Job plots varies with concentration changes. **Figure 3** explains host-guest complexation increases corresponding to the number of quest equivalents. And the stoichiometry of binding equilibrium is also affected by the concentration of titration media.

The Practicality of the Job Plot

The continuous variation method (Job plot) is mainly one of the analysis methods of the

stoichiometry performed. However, according to Jurczark and colleagues[4], while the existing Job plot method is useful for very stable metal-ligand complexes, the approach in supramolecular chemistry gives incorrect results. And recently, Thordarson and Hibbert announced the end of the Job plot in supramolecular chemistry[5]. In this part, I take a different perspective to Job plot, not even providing the order of multivalent binding. Still, it will be an essential experimental method that can give the reliability of the binding models.
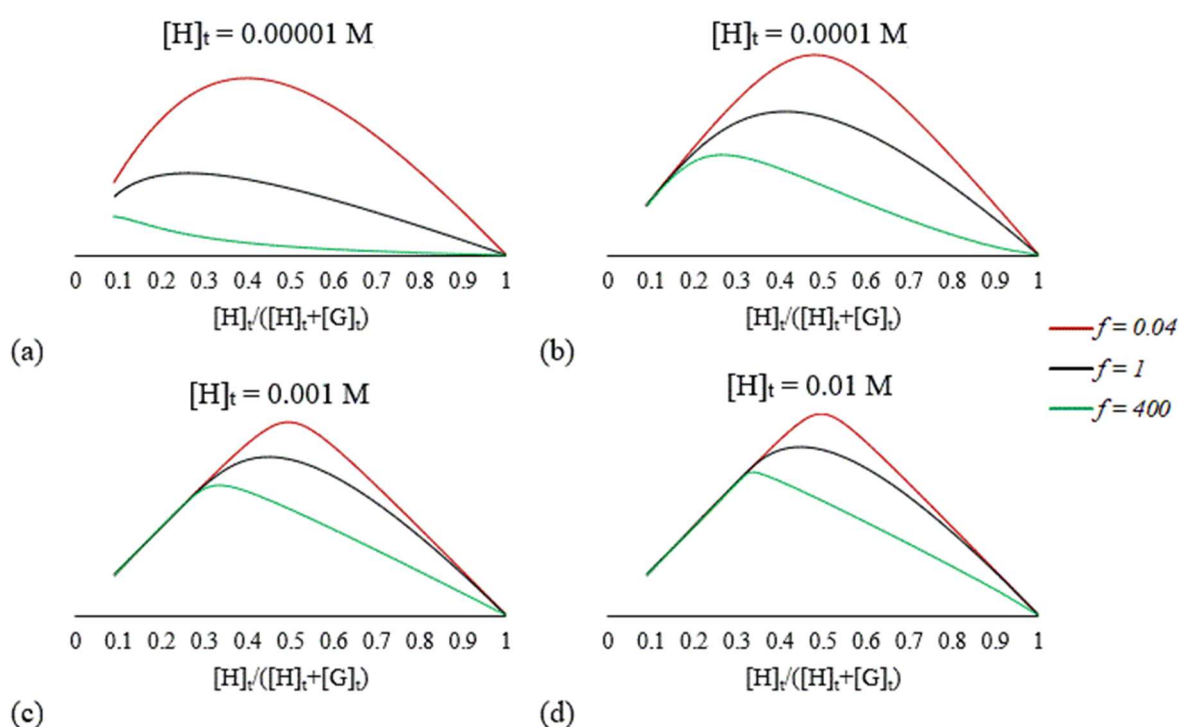

Figure 4. The host concentration of titration media makes a difference in Job plots illustrated by simulations. *x-coordinate : $\frac{[H]_t}{[H]_t+[G]_t}$, *$y$-coordinate : $\frac{[H]_t}{[H]_t+[G]_t} \times ([HG] + [HG_2])$

The simulations in **Figure 4** explains the deviation of the measurement of stoichiometry according to its cooperativity and the concentration of titration media in the Job plots. We could deduce whether multivalent binding or not. If equilibrium is a multivalent binding, the maximum value of the Job plot deviates from 0.5 depending on its cooperativity. We must recognize that the x-coordinate corresponding to the high point of the Job plot does not represent the number of binding sites of a host. In that respect, it is not a desirable method only

to perform a Job plot to measure the ratio of host-guest complexations. But we might take advantage of the Job plot as data supporting the reliability of the binding models in high concentration media. For example, if the point at which the x-coordinate is maximum in the Job plot is close to 0.33, we could judge that the model with the result with positive cooperativity is highly reliable. On the other hand, if the point at which the x-coordinate is maximum is close to 0.5 in the Job plot, we could think of monovalent or negative cooperative binding. Therefore, the need for a Job plot is quite limited.

In conclusion, I briefly formulated the general binding isotherm with the derivation of the multivalent equilibrium model. From this model, the hypothetical titration simulations were conducted in cases of divalent binding equilibrium bearing two structurally identical binding sites with simple Python codes. As a result of these simulations using four factors, the order of host-guest binding equilibrium could be obtained only limited conditions when the titration reaches over-saturation and high concentration of solution media (**Figures 2** and **3**). The limited practicality of the Job plot was also considered.

Abbreviations and Symbols

f : cooperativity factor;

[H] : molar concentration of unsaturated host;

[G] : molar concentration of unbound guest;

$[H]_t$: total concentration of host (unsaturated + saturated);

$[G]_t$: total concentration of guest (unbound + bound);

$[H_mG_n]$: total concentration of complexes;

K_1 : first stepwise binding constant;

K_2 : second stepwise binding constant;

m : the overall ratio of the host in complexes;

n : the overall ratio of guest in complexes;

f_{HG_n} : the mole fraction of HG_n;

β_{mn} : the overall binding constant;

α : the mole fraction of complexes;

v : the mole fraction of complexes weighted by the number of guests;

$\frac{v}{\alpha} = \frac{n}{m}$: the mean number of bound guests in the complexes;

$\frac{\alpha}{\alpha+v} = \frac{m}{m+n}$: the fraction of host in the binding equilibrium;

References

1. T. R. Middendorf, R. W. Aldrich, *J. Gen. Physiol.*, **2017**, 149 (1), 105–119.

2. P. Thordarson, *Chem. Soc. Rev.*, **2011**, 40, 1305–1323.

3. A. E. Hargrove, Z. Zhong, J. L. Sessler, and Eric V. Anslyn, *New J. Chem.*, **2010**, 34, 348–354.

4. F. Ulatowski, K. Dąbrowa, T. Bałakier, and J. Jurczak, *J. Org. Chem.*, **2016**, 81, 1746−1756.

5. D. B. Hibbert, and P. Thordarson, *Chem. Commun.*, **2016**, 52, 12792–12805.