

CHARACTERIZATION OF PARVALBUMIN AND POLCALCIN DIVALENT ION BINDING BY ISOTHERMAL TITRATION CALORIMETRY

Michael T. Henzl*

Contents

1. Introduction	260
2. Practical Aspects of Data Collection	262
2.1. Buffer selection	263
2.2. Standardization of metal ion and chelator solutions	263
2.3. Removal of metal ions from buffers and protein solutions	264
2.4. Preparation of EDTA-agarose	264
2.5. Removal of Ca^{2+} from protein samples	265
2.6. Binding parameters for competing chelators	267
2.7. ITC data collection	268
2.8. Data set preparation	268
2.9. Preparation of the parameter file	271
2.10. General comments on ITC model development	271
2.11. Binding in the presence of a competing chelator	275
2.12. Binding in the presence of a competing metal ion	276
2.13. Least-squares minimization	277
2.14. Error analysis	280
3. Illustrative Global ITC Analyses of Divalent Ion Binding	281
3.1. The independent two-site model	281
3.2. Competing chelator	283
3.3. Competing metal ion	284
3.4. Analysis of the divalent ion binding by the S55D/E59D variant of rat α -parvalbumin	285
3.5. Analysis of positively cooperative divalent ion binding	288
3.6. Modeling divalent ion binding by Phl p 7	291
4. Conclusion	295
Acknowledgment	295
References	295

* Department of Biochemistry, University of Missouri, Columbia, Missouri, USA

Abstract

The elucidation of structure-affinity relationships in EF-hand proteins requires a reliable assay of divalent ion affinity. In principle, isothermal titration calorimetry (ITC) should be capable of furnishing estimates for Ca^{2+} - and Mg^{2+} -binding constants in these systems. And because the method yields the binding enthalpy directly, ITC can provide a more detailed view of binding energetics than methods that rely on $^{45}\text{Ca}^{2+}$ or fluorescent indicators. For several reasons, however, it is generally not possible to extract reliable binding parameters from single ITC experiments. Ca^{2+} affinity is often too high, and Mg^{2+} affinity is invariably too low. Moreover, least-squares minimization of multisite systems may not afford a unique fit because of strong parameter correlations. This chapter outlines a strategy for analyzing two-site systems that overcomes these obstacles. The method—which involves simultaneous, or global, least-squares analysis of direct and competitive ITC data—yields binding parameters for both Ca^{2+} and Mg^{2+} . Application of the method is demonstrated for two systems. The S55D/E59D variant of rat α -parvalbumin, noteworthy for its elevated metal ion affinity, binds divalent ions noncooperatively and is amenable to analysis using an independent two-site model. On the other hand, Phl p 7, a pollen-specific EF-hand protein from timothy grass, binds Ca^{2+} with positive cooperativity. Divalent ion-binding data for the protein must be analyzed using a two-site Adair model.

1. INTRODUCTION

The ability to selectively replace amino acid residues within a protein provides a powerful tool for investigating its function. Effective incorporation of site-specific mutagenesis into a structure-function study, however, requires an effective assay for altered function. For an enzyme, perturbations can be readily diagnosed with a standard activity assay, and the nature of the perturbation can often be interpreted from its relative impact on k_{cat} and or $k_{\text{cat}}/K_{\text{m}}$. However, the biological activity of many physiologically significant proteins is restricted to noncovalent interaction with one or more target ligands. For those systems, it is important to have an accurate, convenient assay for altered ligand affinity.

Isothermal titration calorimetry (ITC) offers a particularly attractive method for studying protein-ligand reactions (Freyer and Lewis, 2008; Lewis and Murphy, 2005; Leavitt and Freire, 2001; Ladbury and Chowdhry, 1996). In an ITC experiment, automated additions of the ligand are made at regular intervals to a solution of the protein at constant temperature. The modern commercial instruments used to study protein-ligand interactions are so-called power-compensation calorimeters. They maintain a constant, minuscule temperature difference between the sample cell and a reference cell filled with buffer (or water). The magnitude of the heat flux that accompanies an injection of titrant is estimated from the

change in electrical power required to maintain the temperature difference. Because the vast majority of protein-ligand interactions are accompanied by detectable absorption or evolution of heat, ITC is a nearly universal method for monitoring binding reactions.

ITC is the only technique that yields binding enthalpy directly. Because the enthalpy is measured as a function of the ligand concentration, the titration can also provide estimates for the overall binding free energy and stoichiometry. Given the free energy change and the reaction enthalpy, the entropic contribution can be calculated by difference. Protonation events can be readily diagnosed from the observation of a buffer-dependent reaction enthalpy. Moreover, the apparent ΔC_p for the interaction, which is obtained from the temperature dependence of the reaction enthalpy, can furnish insight into the molecular details of the protein-ligand interaction.

For a number of years, this laboratory has sought to understand the physical and structural basis for variations in divalent ion affinity in EF-hand proteins (Celio, Pauls, and Schwaller, 1996; Kawasaki and Kretsinger, 1995; Kretsinger, 1980; Strynadka and James, 1989), using the parvalbumin molecule as a model. Parvalbumins are small (M_r 12,000), vertebrate-specific proteins that contain two EF-hand binding motifs (Heizmann and Kagi, 1989; Kretsinger and Nockolds, 1973; Pauls, Cox, and Berchtold, 1996). Although they are generally viewed as interchangeable Ca^{2+} buffer proteins, in fact there are significant differences in divalent ion-binding properties among the various parvalbumin isoforms. We have attempted to exploit these differences to further our understanding of determinants of divalent ion affinity. We have also conducted numerous site-specific mutagenesis studies. The ability to facilely measure divalent ion affinity has been crucial to this project.

The lab purchased a microtitration calorimeter from Hart Scientific in 1993. However, we continued to rely on $^{45}\text{Ca}^{2+}$ flow-dialysis to evaluate Ca^{2+} affinity until 2001. The calorimeter was used exclusively for measuring binding enthalpies, employing the binding constants determined by corresponding flow-dialysis measurements in least-squares modeling. Mg^{2+} affinity was largely ignored, although occasionally an attempt was made to extract Mg^{2+} affinity by competition in flow-dialysis assays. This approach seems archaic in retrospect. However, it was simply not possible to extract reliable estimates for the parvalbumin divalent ion-binding constants from single ITC experiments or to accurately assess small changes in these values. The Ca^{2+} affinities in these systems approach or exceed the upper limit for accurate estimation by ITC. By contrast, the Mg^{2+} affinities are generally too low to determine unambiguously. For both ions, the least-squares treatment often fails to find a unique solution, because of correlations between the binding parameters for the two binding sites.

In recent years, we have employed an ITC-based strategy that surmounts the obstacles. The method involves simultaneous, or global, nonlinear least-squares modeling of direct and competitive ITC data. The

decision to adopt this approach was, in large part, motivated by the success of global NLLS methods for modeling sedimentation equilibrium (Johnson *et al.*, 1981), sedimentation velocity (Philo, 1997; Schuck, 2000), and time-resolved fluorescence (Beechem, 1992) data.

Analysis of single-site systems by ITC is generally straightforward, in the absence of confounding ancillary issues (e.g., limited protein availability or solubility, high ligand affinities). Although the complexity of the problem increases exponentially with the number of binding sites, we herein demonstrate that the treatment of two-site systems remains quite tractable.

We currently employ a MicroCal VP-ITC for our calorimetric analyses. However, the first application of the global ITC approach, to measure the divalent ion-binding parameters for wild-type rat β -parvalbumin, was actually performed with data collected on a Hart 4209 microtitration calorimeter. Although the 4200 series has been vastly improved in the interim, our 1993-vintage instrument had a signal-to-noise ratio fully an order of magnitude lower than the VP-ITC. That the analyses with data from the two calorimeters yielded comparable parameter values is testimony to the robustness of the global fitting strategy.

The motivation for writing this chapter was to provide a reasonably detailed road map for implementing the global ITC analysis of calcium-binding proteins harboring two binding sites. The chapter begins with a discussion of some practical issues related to data acquisition. We then outline the general approach for modeling ITC data, including descriptions of the in-house-generated software that we employ for our analyses. The chapter concludes with two applications of the method. In the first, we describe the analysis of the S55D/E59D variant of rat α -parvalbumin, noteworthy for its extremely high affinity for Ca^{2+} and Mg^{2+} . The binding of divalent ions in this case is macroscopically noncooperative, permitting the use of an independent two-site model. In the second example, we analyze divalent ion binding by Phl p 7, a member of the polcalcin family. Ca^{2+} binding in this system is positively cooperative, necessitating the application of a general two-site model.

2. PRACTICAL ASPECTS OF DATA COLLECTION

To obtain estimates of the binding enthalpies and binding constants for both Ca^{2+} and Mg^{2+} , aliquots of the protein of interest are subjected to a battery of titrations. Typically, Ca^{2+} titrations are performed at two or more protein concentrations. Mg^{2+} titrations are likewise performed at one or more protein concentrations. The protein is titrated with Ca^{2+} at several fixed levels of Mg^{2+} . It is also titrated with Ca^{2+} in the presence of

competitive chelators—typically EDTA, EGTA, and NTA. Finally, the protein is titrated with Mg^{2+} in the presence of EDTA.

2.1. Buffer selection

The analysis buffer should have low affinity for Ca^{2+} and Mg^{2+} . Phosphate and bicarbonate are unsuitable. Although di- and tricarboxylic acids (e.g., malonate and citrate, respectively) could potentially be used as buffers between pH 5 and 6, they have the potential to chelate divalent ions and should be avoided. At pH 7.4, the pH at which we conduct our analyses, Hepes is an effective buffer. Other buffers in this series (e.g., Mes, Pipes, Mops) would be logical candidates at other pH values.

Protonation phenomena associated with binding can be diagnosed by titration of the protein with Ca^{2+} in two or more buffers differing in ionization enthalpy. A buffer-dependent binding enthalpy signals (de)protonation. If the apparent binding enthalpy is plotted against the buffer ionization enthalpy, the slope equals the number of protons involved at the pH of the analysis. Ionization enthalpies have been tabulated for many common buffer systems (Fukada and Takahashi, 1998; Goldberg, Kishore, and Lennen, 2002).

The composition of titrant and sample buffers should be as closely matched as possible, to minimize artifactual mixing heats. Unfortunately, when dealing with low-molecular-weight ligands, dialysis of the protein and ligand against a common pool of buffer is not an option. Alternatively, the protein could be dialyzed to equilibrium against a pool of buffer, which is subsequently used to prepare the titrant solutions. However, this approach requires time-consuming standardization of titrant solutions for every protein preparation. In our experience, with exercise of due care in their preparation, the variation between individual buffer solutions is acceptably low.

2.2. Standardization of metal ion and chelator solutions

Careful consideration should be given to the preparation of reagents. In this lab, 1.0 M solutions of analytical reagent-grade $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ and $\text{MgCl}_2 \cdot 2\text{H}_2\text{O}$ are prepared gravimetrically in the analysis buffer. Preparing the concentrated stock solutions in the analysis buffer helps to ensure that heat-of-mixing artifacts arising from adding the diluted titrant to the protein solution will be minimal. Aliquots of the stock Mg^{2+} solution are diluted in analysis buffer to obtain 100 mM and 2.0 mM solutions. Then 20 mM and 1.0 mM solutions of Ca^{2+} are similarly prepared. We store aliquots of the stock solutions and dilutions in tightly sealed polypropylene containers at -20°C .

We treat analytical-grade $\text{Na}_2\text{EDTA}\cdot 2\text{H}_2\text{O}$ as a primary standard. A 100.0 mM stock solution is prepared in the analysis buffer and then appropriately diluted, with the analysis buffer, to obtain 10.0 mM and 0.15 mM solutions. The latter is titrated with the 1.0 mM Ca^{2+} and 2.0 mM Mg^{2+} solutions to obtain more precise estimates of the respective metal ion concentrations.

Additionally, 100 mM stock solutions of EGTA and NTA are likewise prepared and diluted with buffer to obtain 10.0 mM solutions. These are used in the sample preparation. Additionally, 0.15 mM EGTA and 1.0 mM NTA solutions are prepared. The former is standardized by titration with 1.0 mM Ca^{2+} , the latter by titration with 20 mM Ca^{2+} .

2.3. Removal of metal ions from buffers and protein solutions

For the majority of cases, metal ions can be sequestered from protein preparations by passage over EDTA-derivatized agarose. The chelating matrix is prepared by carbodiimide-mediated coupling of EDTA to aminohexyl agarose.

2.4. Preparation of EDTA-agarose

Aminohexyl agarose is prepared by activation with 1,4-butanedioldiglycidyl ether (Pepper, 1998), followed by reaction with 1,6-diaminohexane at pH 13. For example, 100 mL (moist cake) of Sepharose 4B is placed in a 500-mL polypropylene reaction bottle. In a fume hood, 100 mL of 0.6 N NaOH and 100 mL of 1,4-butanedioldiglycidyl ether are added, and the suspension is incubated overnight at 25 °C with vigorous agitation (required to keep the minimally soluble ether in suspension). The suspension is filtered through a sintered-glass funnel and washed extensively with hot tap water (e.g., 10 vols) until the smell of the 1,4-butanedioldiglycidyl ether is faint. The resulting epoxy-activated agarose is ready for modification with 1,6-diaminohexane.

Then, 1.0 g of 1,6-diaminohexane is dissolved in 100 mL of 1 M sodium carbonate, and the pH is adjusted to 13 with concentrated NaOH. This solution is added to the epoxy-activated gel cake, and the pH is readjusted to 13, if necessary. The resulting suspension is incubated for 24–48 h with vigorous agitation at 25 °C. At the end of the incubation, excess reagents are removed by filtration, and the gel is washed extensively with water.

After removing excess water, the aminohexyl agarose is resuspended in 250 mL of 0.5 M EDTA, pH 6.0. Then 5.0 g of EDAC are added to the gel suspension, which is then incubated at room temperature with constant agitation. After 4 h, the pH is measured and readjusted to 6.0 if necessary. A second 5.0-g aliquot of EDAC is added, and incubation is continued for another 4 h. The resulting material is washed extensively with water and then 10 mM Ca^{2+} to remove excess EDTA.

The binding capacity is determined by saturating an aliquot of the gel with Ca^{2+} , rinsing off the excess Ca^{2+} , eluting the bound Ca^{2+} with dilute HCl (pH 1), and measuring the Ca^{2+} content of the eluate by flame atomic absorption spectrometry. One mL of EDTA-agarose thus prepared binds 20–30 μmol Ca^{2+} .

The EDTA-derivatized matrix retains a net positive charge due to the presence of unreacted amino groups, and it can act as an anion exchanger at low ionic strength. To prevent nonspecific electrostatic interactions between the acidic EF-hand proteins and the EDTA agarose, a solvent ion concentration of at least 0.15 M is recommended.

2.5. Removal of Ca^{2+} from protein samples

Prior to loading the protein, residual divalent ions are stripped from the column with four bed volumes of dilute HCl (pH 2). The HCl in the head space is then removed and replaced with the analysis buffer, and the column is eluted with the analysis buffer until the pH of the eluate matches the buffer pH. The protein (2.5 μmol , at a concentration of 80 μM) is loaded onto the column at 0.5 mL/min. After the protein has loaded, elution is continued at the same flow rate with buffer. Fractions of the eluate exhibiting significant UV absorbance are combined, and the resulting solution is assayed for residual Ca^{2+} by flame atomic-absorption spectrometry.

Although the protein is diluted approximately 50% by passage over the column, recovery exceeds 90%. The residual Ca^{2+} concentration is typically less than 0.5 μM . Thus, if the protein concentration is 25 μM , removal is $\geq 99\%$ complete. Using this protocol, we have successfully removed the Ca^{2+} from rat α -parvalbumin, which exhibits an average Ca^{2+} -binding constant of $1.2 \times 10^8 \text{ M}^{-1}$ in Hepes-buffered saline.

The key to quantitative removal of divalent cations is to avoid overloading the column. Because a carboxylate is sacrificed in the coupling reaction, immobilized EDTA exhibits fairly modest affinity for Ca^{2+} . The titration of an aliquot of EDTA-agarose with 20 mM Ca^{2+} , in Hepes-buffered saline at pH 7.4, is displayed in Fig. 10.1. Least-squares analysis of the integrated data indicates the existence of two populations of binding sites. The major population binds with an average enthalpy of -3.8 kcal/mol and average binding constant of just $2.7 (0.4) \times 10^5 \text{ M}^{-1}$. A minor population ($<5\%$) binds with marginally higher affinity, $7.7 (1.0) \times 10^5 \text{ M}^{-1}$, and greater exothermicity, $\Delta H = -9.5 \text{ kcal/mol}$. The latter may represent adjacent chelating groups that are capable of functioning cooperatively. Clearly, the ability of the EDTA-agarose matrix to extract Ca^{2+} from proteins with binding constants exceeding 10^8 M^{-1} is strongly dependent on mass action. Assuming 25 μmol of chelator per mL of gel, a 100-mL column would have a total capacity of 2.5 mmol. We typically load 2.5 μmol of protein, which corresponds to 5 μmol of divalent ion-binding sites. Thus, the ratio of chelator sites to protein sites is on the order of 500:1.

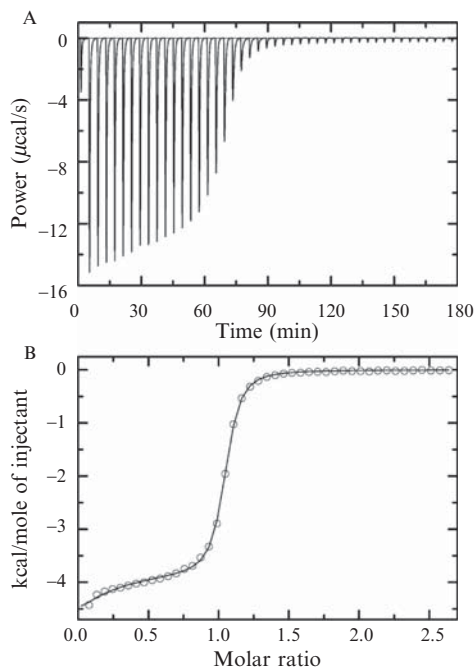


Figure 10.1 Titration of an aliquot of EDTA-derivatized agarose with Ca^{2+} . The raw and integrated ITC data are displayed in panels A and B, respectively.

To date, we have encountered two proteins for which the EDTA-derivatized agarose was unable to quantitatively remove the bound Ca^{2+} . Because K^+ does not compete with Ca^{2+} for the EF-hand sites in rat α -parvalbumin, the protein exhibits substantially higher affinity for Ca^{2+} when K^+ , rather than Na^+ , is the major solvent cation. The average Ca^{2+} -binding constant is $1.6 \times 10^9 \text{ M}^{-1}$ in 0.15 M KCl, 0.025 M Hepes, pH 7.4. After passage over EDTA agarose in this buffer, the protein retains roughly 0.5 equivalents of Ca^{2+} . In this case, the Ca^{2+} was removed in the presence of 0.15 M NaCl, 0.025 M Hepes, pH 7.4. The Na^+ was subsequently replaced by K^+ by repeated cycles of concentration and dilution, in an ultrafiltration cell, with (Ca^{2+} -free) K^+ -containing buffer.

The S55D/E59D variant of rat α -parvalbumin was the second protein that successfully resisted Ca^{2+} removal with EDTA-agarose. As discussed later in the chapter, this protein exhibits extremely high affinity for Ca^{2+} . The binding constant for Ca^{2+} at the engineered CD site exceeds 10^9 M^{-1} in Hepes-buffered saline and 10^{10} M^{-1} when K^+ replaces Na^+ as the major solvent cation. To remove residual Ca^{2+} in these cases, it was necessary to dialyze the protein (20 mL, at 1.5 mg/mL) against a large excess of EDTA

(1 L of NaCl/Hepes, containing 0.01 M EDTA), then dialyze away the bulk of the EDTA (against 1 L of buffer containing 0.001 M EDTA). Completion of EDTA removal and solvent cation exchange, if required, were completed by repeated cycles of concentration and dilution with the appropriate Ca^{2+} -free buffer in an ultrafiltration cell.

2.6. Binding parameters for competing chelators

Because the divalent ion-binding parameters extracted from titrations conducted in the presence of a competing chelator are dependent on the corresponding values for the chelator, accurate knowledge of the chelator binding behavior is important. The three competitive chelators that we routinely employ are EDTA, EGTA, and NTA. The relevant binding constants and enthalpies that we observe in 0.15 M NaCl, 0.025 M Hepes, pH 7.4, are listed in Table 10.2.

The Ca^{2+} -binding parameters for NTA were obtained by titrating a 1.0 mM, 0.5 mM, and 0.25 mM aliquots of the chelator with 20 mM or 5 mM Ca^{2+} . The resulting data were analyzed globally with a single-site model to extract the binding constant and binding enthalpy.

The Ca^{2+} -binding parameters for EGTA were determined by titrating EGTA with Ca^{2+} in the presence of 2.0 mM NTA. The equations used to treat these data are described subsequently. To determine the binding parameters associated with EDTA, the chelator was first titrated with Mg^{2+} . The binding constant and enthalpy for Mg^{2+} , $5.5 \times 10^5 \text{ M}^{-1}$ and 4.2 kcal/mol, respectively, can be readily estimated from direct titrations with Mg^{2+} . The corresponding values for Ca^{2+} were then obtained by titrating samples of EDTA with Ca^{2+} and with Ca^{2+} in the presence of Mg^{2+} . The equations used to treat these data are likewise described subsequently.

Clearly, the competitive chelators employed in the experiments should not bind to the protein. Although we have not encountered this problem to date in our analyses, interactions between EF-hand proteins and Ca^{2+} indicators have been reported. For example, Haiech *et al.* reported that EGTA binds to hake parvalbumin with an apparent dissociation constant of 35 mM at pH 7.55 (Haiech *et al.*, 1979). Moreover, whether such interactions occur may be strongly influenced by solution conditions. To determine whether protein-chelator interactions are occurring, the protein should be titrated with a solution of the chelator. Ideally, the observed heat fluxes should be identical to those accompanying addition of chelator to buffer alone. In practice, there may be small differences because the frictional heat generated by injecting the titrant into the stirred sample cell will differ slightly in the presence of the protein. However, the difference should be small, and the heat effects should be constant.

2.7. ITC data collection

We routinely employ a set of 10 titrations, listed in [Table 10.1](#), for the evaluation of parvalbumin/polcalcin Ca^{2+} and Mg^{2+} -binding parameters. The significance of the experiment number is described subsequently. Protein concentrations between 50 and 60 μM are optimal. If the concentrations in the experiments numbered 11 and 13 significantly exceed 60 μM , the amount of titrant in the buret will not be sufficient to reach the endpoint for the reaction. When the protein concentration is significantly less than 50 μM , the endpoint is reached rather earlier during the titration, with a resultant decrease in the informational content of the experiment.

Because overfilling the VP-ITC sample cell requires approximately 1.8 mL of solution, we typically prepare 2.0 mL samples in 5.0-mL plastic vials. To prepare sample 7, for example, 20 μL of a 100 mM Mg^{2+} solution is combined with 1.98 mL of (60 μM) protein solution. Sample 9 is prepared similarly by combining 20 μL of 1.0 M Mg^{2+} and 1.98 mL of protein. We find it convenient to prepare the entire sample series at the same time. The samples are then rapidly frozen in liquid nitrogen and stored at -20°C prior to use.

2.8. Data set preparation

Although one can probably imagine alternative strategies for supplying data to the fitting program, we combine the data from the 10 experiments described previously into a single composite data file. As currently configured, that file includes 13 columns. An excerpt from a typical file is displayed in [Fig. 10.2](#). These values correspond to the first three points from a titration with Ca^{2+} in the presence of EDTA. Clearly, the data file must include the injection heat (Col. 1) for each titration point and an estimated standard deviation (Col. 2), as well as the associated total ligand

Table 10.1 Titrations used to analyze parvalbumin and polcalcin divalent ion binding

Exp. no.	Protein concentration	Competing ion or chelator	Titrant	Injection volume	Number of injections
1	60 μM	—	1.0 mM Ca^{2+}	7 μL	42
2	30 μM	—	1.0 mM Ca^{2+}	5 μL	30
4	60 μM	—	2.0 mM Mg^{2+}	10 μL	29
7	60 μM	1.0 mM Mg^{2+}	1.0 mM Ca^{2+}	7 μL	42
8	60 μM	5.0 mM Mg^{2+}	1.0 mM Ca^{2+}	7 μL	42
9	60 μM	10.0 mM Mg^{2+}	1.0 mM Ca^{2+}	7 μL	42
11	60 μM	60 μM EDTA	1.0 mM Ca^{2+}	7 μL	42
13	60 μM	60 μM EGTA	1.0 mM Ca^{2+}	7 μL	42
15	60 μM	0.10 mM NTA	1.0 mM Ca^{2+}	7 μL	42
17	60 μM	EDTA	2.0 mM Mg^{2+}	7 μL	42

1	2	3	4	5	6	7
Inj. heat (μ cal)	Std. dev. (μ cal)	[Ca ²⁺] _{tot} (M)	[Mg ²⁺] _{tot} (M)	[protein] (M)	[EDTA] (M)	[EGTA] (M)
-7.46189	1.00E+07	1.5645E-06	0.0000	5.7800E-05	5.7918E-05	0.0000
-43.0542	0.350000	7.0228E-06	0.0000	5.7718E-05	5.7632E-05	0.0000
-41.3975	0.350000	1.2454E-05	0.0000	5.7432E-05	5.7350E-05	0.0000
8	9	10	11	12	13	
[NTA] (M)	Expt. no.	Inj. no.	Model no.	Inj. vol. (L)	Tot. vol. (L)	
0.0000	11	1	4	2.0000E-06	1.4193E-03	
0.0000	11	2	4	7.0000E-06	1.4123E-03	
0.0000	11	3	4	7.0000E-06	1.4263E-03	

Figure 10.2 Excerpt from a composite data file. The first three lines from the titration of a parvalbumin sample with Ca²⁺ in the presence of EDTA. Note that the standard deviation associated with the first titrant addition, the 2.0- μ L preinjection, is assigned a value of 10⁷, to ensure that it receives no weight in the least-squares analysis.

(Col. 3 for Ca²⁺, Col. 4 for Mg²⁺) and total protein concentrations (Col. 5). If there is a competing chelator, as in this case, that concentration must also be supplied (Cols. 6, 7, and 8 for EDTA, EGTA, and NTA, respectively).

The experiment number, listed in Col. 9, tells the fitting program which model is to be used to simulate the injection heat. Experiments 1–3 correspond to direct titrations with Ca²⁺, 4–6 to direct titrations with Mg²⁺, 7–10 to titrations with Ca²⁺ in the presence of Mg²⁺, 11–12 to titrations with Ca²⁺ in the presence of EDTA, 13–14 to titrations with Ca²⁺ in the presence of EGTA, 15–16 to titrations with Ca²⁺ in the presence of NTA, and 17–18 to titrations with Mg²⁺ in the presence of EDTA. The injection number for the particular experiment is listed in Col. 10. The model number appears in Col. 11. It will differ depending on whether independent-site (Scatchard) or general (Adair) fitting equations are to be employed for the least-squares minimization. Columns 12 and 13 provide the injection volume and total volume, respectively.

Preparation of the data file requires some care. We integrate the data from separate experiments and perform baseline smoothing, if necessary, using the Origin-based software supplied with the VP-ITC. The topic of baseline smoothing deserves comment. Obviously, one manipulates raw data at his peril, as there is always the danger of reducing the information content of the data set (Johnson and Faunt, 1992). However, the MicroCal baseline algorithm frequently cuts a sawtooth pattern through the data, greatly exaggerating the variance. In these cases, we adjust the baseline so that it bisects the perceived average in the point-to-point noise.

Because there is generally insufficient protein sample available to preinse the cell, the protein sample is diluted by residual buffer in the sample cell at the time of loading. Consequently, the actual and calculated protein concentrations will differ slightly. This situation can compromise the quality of the

least-squares fit. Although it is possible to correct these inaccuracies during the least-squares analysis, as described subsequently, it is preferable to begin with the most accurate estimates for the protein concentration. In many of the titrations, the equivalence point is well-defined. These individual data sets are subjected to a preliminary analysis using the MicroCal software, the motivation being to correct minor errors in protein concentration. The sample concentration is varied until the fitted stoichiometries for both sites are 1.00.

The MicroCal conversion software produces a spreadsheet listing the relevant information associated with each injection of titrant. The first four columns list the injection heat (μcal), the injection volume (μL), the dilution-corrected ligand concentration (mM) in the cell following the injection, and the corresponding dilution-corrected protein concentration (mM). The four columns of data are exported as an ASCII file and subsequently used in the construction of the master data file. We employ a Fortran-based utility called “itcdatafile” for this task.

This program asks for the total number of titrations to be included in the composite data set. Then, for each experiment, it prompts for the name of the corresponding ASCII file, the experiment number, and a standard deviation. Except for the first point in the titration, which is always a 2.0- μL preinjection, we generally utilize a uniform standard deviation for each data point. This number was determined by repeatedly titrating a solution of EDTA with Ca^{2+} , calculating the standard deviation associated with each point in the titration, and averaging the values. For our calorimeter, this value turned out to be 0.35 μcal . The use of a uniform variance is not particularly realistic, as the signal-to-noise varies from experiment to experiment and even from replicate to replicate. However, it is approximately correct and permits the calculation of a reduced χ^2 value:

$$\chi_v^2 = \frac{1}{v} \sum \left(\frac{1}{\sigma_i^2} [\gamma(i) - \gamma(x_i)]^2 \right). \quad (10.1)$$

In Eq. (10.1), v is the number of degrees of freedom (the number of data points minus the number of variable parameters), σ_i is the standard deviation of the i th point, $\gamma(i)$ is the observed injection heat and $\gamma(x_i)$ is the calculated injection heat. The heat associated with the 2- μL preinjection is invariably low due to diffusion of titrant from the buret during the relatively lengthy thermal equilibration period. In order to effectively ignore this point during the least-squares minimization process, it is assigned a standard deviation of 10^7 .

If the experiment number corresponds to a titration in the presence of a fixed level of Mg^{2+} , the program asks for the initial concentration. Likewise, if a competing chelator is present, it prompts for the initial chelator concentration. When the information has been collected for all of the titrations, the program asks for an output file name and subsequently writes out the file.

2.9. Preparation of the parameter file

The parameter file contains the initial estimates of the Ca^{2+} - and Mg^{2+} -binding enthalpies and binding constants for the protein of interest. It also contains the corresponding binding parameters for the competing chelators. These are so-called global parameters, which apply to the entire data set. There are also some experiment-specific, or local, parameters. These include binding stoichiometries for the proteins and chelators in each of the titrations and baseline offsets for each experiment. The stoichiometry values are set to 1.0 by default. As the χ^2 minimum is approached, experiments for which the assumed protein concentration is incorrect become obvious—exhibiting large χ^2 contributions in the vicinity of the equivalence point. For these cases, the corresponding stoichiometry coefficients can be adjusted by trial and error to find the optimal values. The baseline offsets are all set to 0.0 by default. As the χ^2 minimum is approached, it may become apparent that one or more of the experiments requires a baseline correction. This is likewise done empirically, adjusting the value until no further improvement in χ^2 is observed.

We employ a utility called “itcparfile” to produce the parameter file. As currently configured, the resulting parameter file contains four columns. An excerpt from a file is displayed in Fig. 10.3. The first column is a shorthand label for the parameter. The second column is a text descriptor. The third column contains the actual parameter values, and the final column, consisting of 1’s and 0’s, indicates whether the corresponding parameter is to be varied (1) or not (0). The file contains 80 lines. Lines 1–4 present the binding enthalpies and binding constants for Ca^{2+} ; lines 5–8 present the corresponding values for Mg^{2+} . Lines 9 and 10, used only in the Adair model, correspond to the enthalpy change and equilibrium constant, respectively, for the formation of a mixed Ca^{2+} - Mg^{2+} complex. Lines 11–18 present the binding parameters for the competing chelators. Lines 19–38 contain stoichiometry values (n_1 , n_2) for experiments 1–10. Lines 39–62 contain stoichiometry values (n_1 , n_2 , n_C) for experiments 11–18. Lines 63–80 correspond to the baseline adjustments for experiments 1–18.

2.10. General comments on ITC model development

We herein offer a brief introduction to ITC model development. For simplicity, we consider a molecule M having a single-site for ligand X :



For relevance, we further assume that M is a small-molecule chelator (e.g., EDTA or EGTA) and that X is either Ca^{2+} or Mg^{2+} . The model development process begins with the equation for the cumulative heat of binding following the i th titrant addition:

1	2	3	4
H1	SITE 1 CA ENTHALPY	-4100.000	1
K1	SITE 1 CA-BINDING CONSTANT	22000000.000	1
H2	SITE 2 CA ENTHALPY	-3460.000	1
K2	SITE 2 CA-BINDING CONSTANT	1500000.000	1
HM1	SITE 1 MG ENTHALPY	3010.000	1
KM1	SITE 1 MG-BINDING CONSTANT	9200.000	1
HM2	SITE 2 MG ENTHALPY	4160.000	1
KM2	SITE 2 MG-BINDING CONSTANT	170.000	1
HCAMG	MIXED-SPECIES BINDING ENTHALPY	0.000	0
BCAMG	MIXED-SPECIES BINDING CONSTANT	0.000	0
HEDTA	EDTA CA-BINDING ENTHALPY	-6300.000	0
KEDTA	EDTA CA-BINDING CONSTANT	44000000.000	0
HMEDTA	EDTA MG-BINDING ENTHALPY	4400.000	0
KMEDTA	EDTA CA-BINDING CONSTANT	550000.000	0
HEGTA	EDTA CA-BINDING ENTHALPY	-7500.000	0
KEGTA	EGTA CA-BINDING CONSTANT	17000000.000	0
HNTA	NTA CA-BINDING ENTHALPY	-1900.000	0
KNTA	NTA CA-BINDING CONSTANT	11000.000	0
N1	EXPT 1 CA TITRATION	1.000	0
N2	EXPT 1 CA TITRATION	1.000	0
N1	EXPT 2 CA TITRATION	1.000	0
N2	EXPT 2 CA TITRATION	1.000	0
⋮	⋮	⋮	⋮
BL	BASELINE EXPT 15	0.000	0
BL	BASELINE EXPT 16	0.000	0
BL	BASELINE EXPT 17	0.000	0
BL	BASELINE EXPT 18	0.000	0

Figure 10.3 Excerpt from a typical parameter file. Binding enthalpies and binding constants for the protein and chelators occupy the first 18 lines. Binding stoichiometries for 18 potential experiments occupy the following 34 lines. Baseline adjustment parameters occupy the final 18 lines.

$$Q_i = \Delta H \alpha_i. \quad (10.3)$$

In this general equation, ΔH is the enthalpy of the reaction and α_i is the fraction of binding sites occupied on M after the i th addition. ΔH is commonly expressed in cal/mol, whereas Q_i is typically measured in μcal . Thus, for consistency of units, the right side of the equation must be multiplied by $10^6 V_o [M]_{t,i}$, where V_o represents the sample cell volume and $[M]_t$ is the total concentration of M :

$$Q_i = 10^6 V_o [M]_{t,i} \Delta H \alpha_i. \quad (10.4)$$

The factor of 10^6 serves to convert calories to microcalories, and the product $V_o [M]_t$ equals the total moles of M in the sample cell.

For a single-site system, the value of α_i is given by the binding isotherm:

$$\alpha_i = \frac{K x_i}{1 + K x_i}, \quad (10.5)$$

where K is the association constant and x_i is the free molar concentration of ligand after the i th injection, shorthand for $[X]_i$. The denominator in this equation corresponds to the binding partition function, or binding polynomial. It reflects the statistical weight of each of the ligated forms at ligand concentration x_i . For the single-site system, there are just two forms, apo- and bound.

Note that Eq. (10.5) requires knowledge of the free ligand concentration, whereas the independent variable in an ITC experiment is the total ligand concentration, $[X]_{t,i}$. However, given $[X]_{t,i}$, the total concentration of M , $[M]_{t,i}$, and an estimate for K , we can solve for the free ligand concentration. For a one-site system, the solution can be obtained either analytically or numerically.

In the former case, we can write an expression for the association constant K , taking advantage of the fact that $x_i = [X]_{t,i} - [MX]_i$ and $[M]_i = [M]_{t,i} - [MX]_i$:

$$K = \frac{[MX]_i}{[M]_i x_i} = \frac{[MX]_i}{\left([M]_{t,i} - [MX]_i\right) \left([X]_{t,i} - [MX]_i\right)}. \quad (10.6)$$

Clearing the denominator and collecting terms yields a quadratic equation that can be solved for $[MX]_i$, yielding:

$$[MX]_i = \frac{1}{2} \left(\left([M]_{t,i} + [X]_{t,i} + \frac{1}{K} \right) - \left(\left([M]_{t,i} + [X]_{t,i} + \frac{1}{K} \right)^2 - 4^* [M]_{t,i} [X]_{t,i} \right)^{1/2} \right). \quad (10.7)$$

The resulting value can then be used to calculate α_i because

$$\alpha_i = \frac{[MX]_i}{[M]_{t,i}}. \quad (10.8)$$

Given a value for α_i and an estimate for ΔH , Q_i can be calculated via Eq. (10.4).

Alternatively, the free ligand concentration can be estimated numerically. To do this, we define a function, $f(x_i)$, equal to the estimated free-ligand concentration (x_i) plus the estimated bound-ligand concentration minus the (known) total ligand concentration:

$$f(x_i) = x_i + [MX]_i - [X]_{t,i}. \quad (10.9)$$

We can substitute for $[MX]_i$, making use of Eqs. (10.5) and (10.8), to obtain an expression in terms of x_i and K :

$$f(x_i) = x_i + [M]_{t,i} \left(\frac{Kx_i}{1 + Kx_i} \right) - [X]_{t,i}. \quad (10.10)$$

If the initial estimate for x_i is correct, $f(x_i)$ will equal zero. If not, we try other values of x_i until the value is acceptably small, say $10^{-6}[X]_{t,i}$.

The bisection method is an efficient way to proceed. We first identify a value for x_i that yields a positive value of $f(x_i)$. The total ligand concentration, $[X]_{t,i}$, is a logical choice. We refer to that value of x_i as *up* (for upper limit). The value of x_i is then decremented until the value of $f(x_i)$ is negative. That value of x_i is referred to as *lo* (for lower limit). Next, we calculate the average of *up* and *lo* (i.e., we bisect them) to obtain a better estimate for x_i . We then recalculate $f(x_i)$. If it is sufficiently small, we are done. Otherwise, we continue. If $f(x_i)$ is positive, we assign x_i to *up*; if $f(x_i)$ is negative, x_i becomes the new value for *lo*. The updated values for *up* and *lo* are once again averaged to yield a new value of x_i . As before, assuming $f(x_i)$ is still not acceptably small, if $f(x_i)$ is positive, x_i is assigned to *up*. Otherwise, it is assigned to *lo*. The bisection routine is continued until the criterion for acceptability is met.

Regardless of the model used to simulate the cumulative heat, the injection heat for the i th titrant addition is obtained by subtracting the cumulative heats associated with the i th and $(i - 1)$ th data points:

$$q_i = (Q_i - Q_{i-1}) + \left(\frac{dV}{V_o} \right) \left(\frac{Q_i + Q_{i-1}}{2} \right) + bl. \quad (10.11)$$

In both the VP-ITC and Nano ITC, the sample cell is overfilled, so that the observed reaction volume is constant. With this configuration, however, each addition of titrant displaces an equal volume of the reaction mixture from the cell. The second term in Eq. (10.5) corrects for the undetected heat of reaction associated with that displaced volume. The “*bl*” term, short for baseline, corresponds to the mixing heat associated with the titrant injection.

To summarize, ITC model development involves four basic steps: (1) Express the cumulative heat is expressed in terms of the binding enthalpies and fractional occupancies. (2) Express the fractional occupancies in terms of the binding constants and free ligand concentrations. (3) Solve for the free ligand concentration, employing the binding constants and total concentrations of M and X . (4) Calculate the injection heat from the difference between the cumulative heats for the i th and $(i - 1)$ th titrant additions. Implementation of this process is described for two cases: binding of Ca^{2+} to a small molecule chelator in the presence of a competing chelator and binding of Ca^{2+} to chelator in the presence of a competing metal ion, Mg^{2+} .

2.11. Binding in the presence of a competing chelator

Often, ligand-binding affinity precludes a determination of the binding constant by direct titration. For example, the Ca^{2+} affinity of EGTA is sufficiently high that direct determination of the binding constant is problematic. One solution to this problem involves titrating the molecule of interest in the presence of a competitor molecule having lower affinity for the ligand. For example, EGTA can be titrated with Ca^{2+} in the presence of NTA. In this case, the cumulative heat for the i th titrant addition includes contributions from the binding of Ca^{2+} to both EGTA and NTA, as expressed by the following equation:

$$Q_i = \Delta H_{EGTA} \alpha_{i,EGTA} + \Delta H_{NTA} \alpha_{i,NTA}, \quad (10.12)$$

where $\alpha_{i,EGTA}$ and $\alpha_{i,NTA}$ represent the fractions of the chelators coordinating Ca^{2+} after the i th addition. They are calculated as follows:

$$\alpha_{i,EGTA} = \frac{K_{EGTA}[\text{Ca}^{2+}]_i}{1 + K_{EGTA}[\text{Ca}^{2+}]_i}; \quad \alpha_{i,NTA} = \frac{K_{NTA}[\text{Ca}^{2+}]_i}{1 + K_{NTA}[\text{Ca}^{2+}]_i}. \quad (10.13)$$

Employing these expressions and making the units consistent, Q_i can then be written as

$$Q_i = 10^6 V_o \left([\text{EGTA}]_{t,i} \left(\frac{\Delta H_{EGTA} K_{EGTA} [\text{Ca}^{2+}]_i}{1 + K_{EGTA} [\text{Ca}^{2+}]_i} \right) + [\text{NTA}]_{t,i} \left(\frac{\Delta H_{NTA} K_{NTA} [\text{Ca}^{2+}]_i}{1 + K_{NTA} [\text{Ca}^{2+}]_i} \right) \right). \quad (10.14)$$

The next step is to solve for the free ligand concentration. In this case, the free Ca^{2+} concentration can be obtained numerically for this experiment by minimizing this equation:

$$f(x_i) = [\text{Ca}^{2+}]_i + [\text{EGTA}]_{t,i} \left(\frac{K_{EGTA} [\text{Ca}^{2+}]_i}{1 + K_{EGTA} [\text{Ca}^{2+}]_i} \right) + [\text{NTA}]_{t,i} \left(\frac{K_{NTA} [\text{Ca}^{2+}]_i}{1 + K_{NTA} [\text{Ca}^{2+}]_i} \right) - [\text{Ca}^{2+}]_t. \quad (10.15)$$

The injection heats are obtained by substituting the expressions for Q_i (and Q_{i-1}) from [Eq. \(10.14\)](#) into [Eq. \(10.11\)](#).

2.12. Binding in the presence of a competing metal ion

Another strategy for estimating the binding constant for a high-affinity ligand involves titrating the molecule of interest with ligand X in the presence of low-affinity ligand Y (Eatough, 1970; Hu and Eftink, 1994; Sigurskjold, Berland, and Svensson, 1994). For example, whereas the Ca^{2+} affinity of EDTA is too high to measure directly, the Mg^{2+} affinity is readily measured by direct titration. Thus, the Ca^{2+} association constant for EDTA can be determined by titrating the chelator with Ca^{2+} in the presence of a fixed level of Mg^{2+} . For this system, the cumulative heat for the i th titrant addition will include contributions from both the binding of Ca^{2+} and the displacement of Mg^{2+} :

$$Q_i = \Delta H_{\text{Ca}} \alpha_{i,\text{Ca}} + \Delta H_{\text{Mg}} \alpha_{i,\text{Mg}}, \quad (10.16)$$

where $\alpha_{i,\text{Ca}}$ or $\alpha_{i,\text{Mg}}$ represent the fraction of chelator coordinating Ca^{2+} or Mg^{2+} , respectively. These fractional occupancies are given by the following expressions:

$$\begin{aligned} \alpha_{i,\text{Ca}} &= \frac{K_{\text{Ca}}[\text{Ca}^{2+}]_i}{1 + K_{\text{Ca}}[\text{Ca}^{2+}]_i + K_{\text{Mg}}[\text{Mg}^{2+}]_i}; \\ \alpha_{i,\text{Mg}} &= \frac{K_{\text{Mg}}[\text{Mg}^{2+}]_i}{1 + K_{\text{Ca}}[\text{Ca}^{2+}]_i + K_{\text{Mg}}[\text{Mg}^{2+}]_i}. \end{aligned} \quad (10.17)$$

Note that the binding partition function (denominator) contains an additional term corresponding to the Mg^{2+} -bound form of the protein. Clearly, to have a discernible impact on the affinity for Ca^{2+} , the Mg^{2+} must be present at a concentration comparable to, or exceeding, its dissociation constant. Substituting the terms from Eq. (10.17) into Eq. (10.16) yields the following expression for Q_i :

$$\begin{aligned} Q_i = 10^6 V_o [\text{EGTA}]_{t,i} & \left(\frac{\Delta H_{\text{Ca}} K_{\text{Ca}} [\text{Ca}^{2+}]_i}{1 + K_{\text{Ca}} [\text{Ca}^{2+}]_i + K_{\text{Mg}} [\text{Mg}^{2+}]_i} \right. \\ & \left. + \frac{\Delta H_{\text{Mg}} K_{\text{Mg}} [\text{Mg}^{2+}]_i}{1 + K_{\text{Ca}} [\text{Ca}^{2+}]_i + K_{\text{Mg}} [\text{Mg}^{2+}]_i} \right). \end{aligned} \quad (10.18)$$

Application of this equation requires knowledge of both the free Ca^{2+} and Mg^{2+} concentrations. Sigurskjold has described an analytical solution for the free ligand concentration in this type of competitive ITC analysis (Sigurskjold, 2000). Alternatively, these values can be obtained by simultaneously minimizing these two equations:

$$f\left([Ca^{2+}]_i\right) = [Ca^{2+}]_i + [EGTA]_{t,i} \left(\frac{K_{Ca}[Ca^{2+}]_i}{1 + K_{Ca}[CaH^{2+}]_i + K_{Mg}[Mg^{2+}]_i} \right) - [Ca^{2+}]_{t,i}. \quad (10.19)$$

$$f\left([Mg^{2+}]_i\right) = [Mg^{2+}]_i + [EGTA]_{t,i} \left(\frac{K_{Mg}[Mg^{2+}]_i}{1 + K_{Ca}[Ca^{2+}]_i + K_{Mg}[Mg^{2+}]_i} \right) - [Mg^{2+}]_{t,i}. \quad (10.20)$$

A strategy for performing the simultaneous minimization is described in a subsequent section.

2.13. Least-squares minimization

If an appropriate mathematical model is available, the exercise of curve fitting reduces to a search for optimal parameter values. The least-squares strategy for parameter optimization assumes that the set of parameters that provides the best agreement with observation (i.e., the smallest variance) has the highest probability of being correct (Johnson and Faunt, 1992). The validity of this assumption is conditional. It requires that the fitting model be correct, that the data points be independent of one another, and that there be sufficient data points to adequately sample the experimental uncertainty. It further requires that the data be free of systematic error, that the experimental uncertainty can be described by a Gaussian distribution, and that the uncertainty can be attributed entirely to the dependent variables.

The term “least-squares minimization” actually encompasses a number of numerical strategies (Bevington and Robinson, 1992; Johnson and Faunt, 1992;). In so-called grid-search methods, after calculating an initial χ^2 value, one of the fitting parameters a_j is incremented by $\pm\Delta a_j$ (where the sign is chosen such that χ^2 decreases) and χ^2 is recalculated. This step is repeated until χ^2 no longer decreases and begins to increase. This process is carried out for each fitting parameter in turn. The entire procedure is repeated until no further significant reduction in χ^2 is achieved.

Although reliable, grid-search algorithms converge slowly, especially when the parameters are highly correlated. In the more efficient gradient-search or steepest-descent methods, all of the fitting parameters a_j are adjusted simultaneously. The relative magnitudes of the adjustments are chosen so that the movement in parameter space is in the direction of maximum reduction in χ^2 (i.e., along the χ^2 gradient ($\nabla\chi^2$)). The gradient

$\nabla\chi^2$ is a vector pointing in the direction along which χ^2 changes most rapidly:

$$\nabla\chi^2 = \sum_{j=1}^n \left[\frac{\partial\chi^2}{\partial a_j} \hat{a}_j \right], \quad (10.21)$$

where $\hat{a}_j$ is a unit vector in the direction of the a_j coordinate axis. The partial derivatives of χ^2 with respect to each of the parameters are calculated numerically:

$$\frac{\partial\chi^2}{\partial a_j} = \frac{\chi^2(a_j + \delta a_j) - \chi^2(a_j)}{\delta a_j}, \quad (10.22)$$

where δa_j is some fraction of the parameter value. The sign of each δa_j increment is chosen so that χ^2 decreases.

So-called expansion methods offer yet another approach to minimizing χ^2 . In this approach, the fitting function $y(x)$ is approximated by means of a truncated Taylor series about the point a'_j :

$$y(x) = y'(x) + \sum_{j=1}^m \left[\frac{\partial y'(x)}{\partial a_j} \delta a_j \right]. \quad (10.23)$$

In Eq. (10.23), $y'(x)$ is the value of the fitting function when the parameters have starting values a'_j and the derivatives are evaluated at those values. Then, χ^2 can be expressed explicitly as a function of the parameter increments δa_j :

$$\chi^2 = \sum \left(\frac{1}{\sigma_i^2} \left\{ y_i - y'(x_i) - \sum_{j=1}^m \left[\frac{\partial y'(x_i)}{\partial a_j} \delta a_j \right] \right\}^2 \right). \quad (10.24)$$

Then χ^2 is minimized with respect to each of the parameter increments δa_k to obtain the optimal values of the increments:

$$\frac{\partial\chi^2}{\partial\delta a_k} = -2 \sum \left(\frac{1}{\sigma_i^2} \left\{ y_i - y'(x) - \sum_{j=1}^m \left[\frac{\partial y'(x)}{\partial a_j} \delta a_j \right] \right\} \frac{\partial y'(x_i)}{\partial a_j} \right) = 0. \quad (10.25)$$

The resulting m simultaneous equations are solved to obtain the optimal parameter increments, which are then used to update the starting parameter estimates. This process is carried out iteratively until no further significant reduction in χ^2 is obtained. The advantage of this approach is that it converges very rapidly in the neighborhood of a minimum.

One of the most effective least-squares fitting strategies, the Levenberg-Marquardt algorithm, combines the most desirable features of the steepest-descent and expansion methods. When χ^2 is far from a minimum, the Marquardt algorithm behaves like a gradient or steepest-descent search. However, in the neighborhood of a minimum, it behaves more like the expansion method.

Ideally, least-squares minimization should begin with good initial parameter estimates. Often, the binding parameters are known for a closely related protein, an ortholog or site-specific variant. In our experience, however, even with reasonably good starting parameter estimates, the Levenberg-Marquardt algorithm often fails to converge to a global minimum. This failure might reflect the ruggedness of the χ^2 hypersurface resulting from the variation of eight parameters—four enthalpies and four binding constants—and the presence of multiple local minima. Whatever the reason, our least-squares minimization strategy relies heavily on Monte Carlo simulation.

The program “itcmonte” prompts for the data file and parameter file names, the number of parameter sets desired (typically 100), and a value for “delta.” Initially, delta should be set to 0.20. The program generates random numbers between -1.0 and $+1.0$. Each of the variable parameters is multiplied by a random number and the value for delta. This product is then added to the current value of the parameter. After this process has been carried out for each of the adjustable parameters, the resulting modified parameter set is used to calculate simulated injection heats for each of the data points in the composite dataset. If the resulting χ^2 value is less than the starting value, the updated parameter values of the parameters are written to a file and used as the starting point for another round of modification. Otherwise, they are discarded.

This procedure is repeated until the specified number of parameter sets has been collected. Upon completion, the program displays the χ^2 minimum, and the associated parameter values, and writes the latter to a file (“bestpars”). It also performs a simulation with the optimal parameter values and writes an ASCII file called “monteplot” containing the experimental values, the calculated values, and residuals for each experiment. This file can be imported into a plotting program to evaluate the quality of the fit.

The program then asks whether you wish to continue with additional simulations and, if so, the number of parameter sets desired and the value of delta to be employed. As the χ^2 limit decreases, the value of delta should likewise be decreased. As one approaches the minimum in the error space, and the parameter values approach their optimal values, smaller corrections to the parameter values should be used.

In the absence of systematic errors and an accurate estimate of the standard deviation for each point, the reduced χ^2 value for a perfect least-squares fit would be 1.0. In practice, we obtain values between 2 and 5. Occasionally, the Monte Carlo simulations will fail to yield a χ^2 of this magnitude. This situation might reflect a problem in the dataset, perhaps mistyping protein or chelator

concentrations or inadvertently entering concentrations in the wrong units (e.g., millimolar rather than molar). Alternatively, if the initial estimates were extremely poor, the simulations may simply have wandered into the wrong sector of parameter space. If the Monte Carlo simulations stall at a high χ^2 value, it is recommended that the data and simulated fit be examined.

Nonconvergence of the χ^2 values can also be a signal that the fitting model is incorrect. This situation was encountered during the initial analysis of the polycalcin Phl p 7, described subsequently. Initially, we tried the two-site independent-site model that we typically employ for analysis of parvalbumin binding data. When it failed to yield an acceptable χ^2 value, we modified the data set to use the Adair fitting functions, which readily yielded a solution.

Once the Monte Carlo simulations have identified a set of parameter values that yield a reduced χ^2 value in the acceptable range, a round of least-squares minimization is performed. We utilize a standard Levenberg-Marquardt algorithm, adapted from the CURFIT routine in the original Bevington monograph (Bevington, 1969), for minimization. The program called “itcmarmqmin” prompts for the names of the data set and the parameter file. At the conclusion of the minimization, it writes a file called “marqplot”, comparable to “monteplot”, which can be imported into a plotting program for visualization.

2.14. Error analysis

The least-squares analysis is not concluded with the acquisition of a set of parameter values. The importance of knowing the uncertainties associated with the estimated binding parameters cannot be overstated (Johnson and Faunt, 1992). Tellinghuisen has discussed this issue in the context of analyzing ITC data (Tellinghuisen, 2004). In our work, we have conducted statistical error analyses by Monte Carlo analysis, simulated data sets, and confidence intervals. In the Monte Carlo analysis, the optimized parameter values are modified with the aid of a random-number generator. A simulation is conducted with the resulting parameter set, and the χ^2 value is recorded. This procedure is repeated perhaps 5,000 to 10,000 times. At the end, the F-statistic is used to estimate the χ^2 value associated with, say, the 95% confidence interval. The range of parameter values observed at this limit is accepted as the uncertainty associated with a given parameter.

Alternatively, synthetic data sets can be used to estimate parameter uncertainty. A set of simulated data is generated with the optimized parameter values using the program “syndat.” Random noise is introduced into this data set by “Gaussian data smearing.” This procedure is repeated until 525 simulated datasets have been collected, which is equivalent, in principle, to having repeated the actual experimental analysis 525 times. Each data set in this ensemble is then subjected to Marquardt-Levenberg minimization. The distribution of parameter values can be used as an estimate of the parameter uncertainty.

Our preferred method of error analysis involves generation of confidence intervals for each parameter. Following identification of the optimal parameter values and the lowest chi-square value, $\chi^2(\min)$, confidence intervals are obtained by incrementing the parameter of interest, fixing its value, and repeating the least-squares minimization, allowing the remaining parameters to vary. This procedure is repeated until the chi-square value, $\chi^2(\text{par})$, exceeds a specified threshold, given by

$$\chi^2(\text{par}) = \left(1 + \frac{p}{\nu} F(p, \nu, P)\right) \left(\chi^2(\min)\right). \quad (10.26)$$

In this equation, p is the number of fitting parameters, ν the degrees of freedom, P the probability that the increase in χ^2 could be the result of random errors (e.g., 0.10 for the 90% confidence limit), and $F(p, \nu, P)$ the corresponding F statistic. After the upper limit has been determined, the parameter value is decremented from its optimal value until $\chi^2(\text{par})$ again exceeded the specified threshold.

The program "itcconfint" is used for this procedure. At the conclusion of the analysis, the confidence intervals are written to a file called "parlimits." In addition to defining confidence intervals for each of the parameters, the program will occasionally reveal a combination of parameters that produces a (slightly) improved χ^2 value. If this happens, the improved parameter estimates are written to file called "newoptpars."

3. ILLUSTRATIVE GLOBAL ITC ANALYSES OF DIVALENT ION BINDING

3.1. The independent two-site model

For all of the parvalbumin analyses that we have performed, it has been possible to employ an independent two-site model, sometimes called a Scatchard model. For a protein with two ligand-binding sites, the cumulative heat after the i th titrant addition can be expressed as

$$Q_i = \Delta H_1 \alpha_{1,i} + \Delta H_2 \alpha_{2,i}, \quad (10.27)$$

where ΔH_1 and ΔH_2 are the binding enthalpies associated with sites 1 and 2 and $\alpha_{1,i}$ and $\alpha_{2,i}$ are the corresponding fractional occupancies. The extent of binding after the i th addition, X_i , for a two-site system is given by:

$$\bar{X}_i = \left(\frac{n_1 k_1 x_i}{1 + k_1 x_i} + \frac{n_2 k_2 x_i}{1 + k_2 x_i} \right). \quad (10.28)$$

In this equation, the first term, $n_1 k_1 x_i / (1 + k_1 x_i)$, describes the fractional occupancy of site 1, $\alpha_{1,i}$, and the second term describes the fractional occupancy of site 2, $\alpha_{2,i}$. The n_1 and n_2 values allow for nonunity binding-site stoichiometries, which could result from minor errors in the protein concentration. The denominator in each term is the binding polynomial for the particular site. The equation for the free ligand concentration is given by

$$x_i = [X]_{t,i} - [M]_{t,i} \left(\frac{n_1 k_1 x_i}{1 + k_1 x_i} + \frac{n_2 k_2 x_i}{1 + k_2 x_i} \right). \quad (10.29)$$

Multiplying both sides of the equation by $(1 + k_1 x_i)$ and $(1 + k_2 x_i)$ yields a cubic equation that can be solved exactly for x_i , employing the following system of equations:

$$A = k_1 k_2. \quad (10.30)$$

$$B = k_1 + k_2 + k_1 k_2 \left([M]_{t,i} (n_1 + n_2) - [X]_{t,i} \right). \quad (10.31)$$

$$C = 1 + [M]_{t,i} (n_1 k_1 + n_2 k_2) - [L]_{t,i} (k_1 + k_2). \quad (10.32)$$

$$D = -[X]_{t,i}. \quad (10.33)$$

$$F = \frac{1}{3} \left(\frac{3C}{A} - \frac{B^2}{A^2} \right). \quad (10.34)$$

$$G = \frac{1}{27} \left(\frac{2B^3}{A^3} - \frac{9BC}{A^2} + \frac{27D}{A} \right). \quad (10.35)$$

$$H = \frac{G^2}{4} + \frac{F^3}{27}. \quad (10.36)$$

$$I = \left(\frac{G^2}{4} - H \right)^{0.5}. \quad (10.37)$$

$$J = I^{1/3}. \quad (10.38)$$

$$K = \cos^{-1} \left(\frac{-G}{2I} \right). \quad (10.39)$$

$$x_i = 2J \cos \left(\frac{K}{3} \right) - \frac{B}{3A}. \quad (10.40)$$

Alternatively, the value of x_i can be obtained by bisection, as described earlier, minimizing the following function:

$$f(x_i) = x_i + [M]_{t,i} \left(\frac{n_1 k_1 x_i}{1 + k_1 x_i} + \frac{n_2 k_2 x_i}{1 + k_2 x_i} \right) - [X]_{t,i}. \quad (10.41)$$

Once the free-ligand concentration has been obtained, the cumulative heat can be calculated with the expression:

$$Q_i = 10^6 V_o [M]_{t,i} \left(\frac{\Delta H_1 n_1 K_1 x_i}{1 + K_1 x_i} + \frac{\Delta H_2 n_2 K_2 x_i}{1 + K_2 x_i} \right), \quad (10.42)$$

which may be substituted into [Eq. \(10.11\)](#) to calculate the injection heat. These equations can be used to model data obtained by direct titration of the protein with either Ca^{2+} or Mg^{2+} .

3.2. Competing chelator

If the titration is carried out in the presence of a competing chelator (e.g., titration of the protein with Ca^{2+} in the presence of EDTA), the cumulative heat includes contributions from binding of the metal ion to both the protein and the chelator:

$$Q_i = \Delta H_1 \alpha_{1,i} + \Delta H_2 \alpha_{2,i} + \Delta H_C \alpha_{C,i} \quad (10.43)$$

where ΔH_C and $\alpha_{C,i}$ are the binding enthalpy and fractional occupancy of the chelator. Substituting the relevant fractional occupancies for the protein and chelator binding sites (given by [Eqs. \(10.23\) and \(10.5\)](#), respectively) yields:

$$Q_i = 10^6 V_o \left([M]_{t,i} \left(\frac{\Delta H_1 n_1 k_1 x_i}{1 + k_1 x_i} + \frac{\Delta H_2 n_2 k_2 x_i}{1 + k_2 x_i} \right) + [C]_{t,i} \left(\frac{\Delta H_C n_C k_C x_i}{1 + k_C x_i} \right) \right) \quad (10.44)$$

In this equation, $[C]_{t,i}$ represents the total chelator concentration, and the expression in parentheses immediately following describes the binding of ligand to the chelator.

To apply Eq. (10.44), it is of course necessary to solve for the free ligand concentration. The relevant equation to be minimized has the following general form:

$$f(x_i) = x_i + [M]_{t,i} \left(\frac{n_1 k_1 x_i}{1 + k_1 x_i} + \frac{n_2 k_2 x_i}{1 + k_2 x_i} \right) + [C]_{t,i} \left(\frac{n_C k_C x_i}{1 + k_C x_i} \right) - [X]_{t,i} \quad (10.45)$$

3.3. Competing metal ion

When a protein is titrated with Ca^{2+} in the presence of Mg^{2+} , the cumulative heat expression must include the contributions from the binding of both ions:

$$Q_i = \Delta H_{1\text{Ca}} \alpha_{1\text{Ca},i} + \Delta H_{2\text{Ca}} \alpha_{2\text{Ca},i} + \Delta H_{1\text{Mg}} \alpha_{1\text{Mg},i} + \Delta H_{2\text{Mg}} \alpha_{2\text{Mg},i}, \quad (10.46)$$

where, for example, $\Delta H_{1\text{Ca}}$ and $\alpha_{1\text{Ca},i}$ represent the Ca^{2+} -binding enthalpy for site 1 and the fractional occupancy of that site by Ca^{2+} , respectively.

The extent of Ca^{2+} binding in the presence of a fixed concentration of Mg^{2+} is described by this equation:

$$\begin{aligned} \bar{X}_i = & \frac{n_1 k_{1\text{Ca}} [\text{Ca}^{2+}]_i}{1 + k_{1\text{Ca}} [\text{Ca}^{2+}]_i + k_{1\text{Mg}} [\text{Mg}^{2+}]_i} \\ & + \frac{n_2 k_{2\text{Ca}} [\text{Ca}^{2+}]_i}{1 + k_{2\text{Ca}} [\text{Ca}^{2+}]_i + k_{2\text{Mg}} [\text{Mg}^{2+}]_i}, \end{aligned} \quad (10.47)$$

where $k_{1\text{Ca}}$ and $k_{2\text{Ca}}$ represent the Ca^{2+} -binding constants for sites 1 and 2 and $k_{1\text{Mg}}$ and $k_{2\text{Mg}}$ represent the corresponding Mg^{2+} -binding constants. The two terms on the right-hand side of Eq. (10.47) correspond to $\alpha_{1\text{Ca},i}$ and $\alpha_{2\text{Ca},i}$ in Eq. (10.46). An analogous equation can be written for Mg^{2+} :

$$\begin{aligned} \bar{Y}_i = & \frac{n_1 k_{1\text{Mg}} [\text{Mg}^{2+}]_i}{1 + k_{1\text{Ca}} [\text{Ca}^{2+}]_i + k_{1\text{Mg}} [\text{Mg}^{2+}]_i} \\ & + \frac{n_2 k_{2\text{Mg}} [\text{Mg}^{2+}]_i}{1 + k_{2\text{Ca}} [\text{Ca}^{2+}]_i + k_{2\text{Mg}} [\text{Mg}^{2+}]_i}, \end{aligned} \quad (10.48)$$

where the two terms on the right side of the equation correspond to $\alpha_{1\text{Mg},i}$ and $\alpha_{2\text{Mg},i}$ in Eq. (10.46).

Substituting these fractional occupancies into Eq. (10.46) yields the following expression for the cumulative heat of binding:

$$Q_i = n_1 \left(\frac{\Delta H_1 k_1 [Ca^{2+}] + \Delta H_{1Mg} k_{1Mg} [Mg^{2+}]}{1 + k_1 [Ca^{2+}] + k_{1Mg} [Mg^{2+}]} \right) + n_2 \left(\frac{\Delta H_2 k_2 [Ca^{2+}] + \Delta H_{2Mg} k_{2Mg} [Mg^{2+}]}{1 + k_2 [Ca^{2+}] + k_{2Mg} [Mg^{2+}]} \right). \quad (10.49)$$

Application of Eq. (10.49) requires knowledge of both the free Ca^{2+} and free Mg^{2+} concentrations. We begin by solving for $[Ca^{2+}]_i$ in the absence of Mg^{2+} and for $[Mg^{2+}]_i$ in the absence of Ca^{2+} . These estimates, which can be obtained analytically using Eqs. (10.30–40), constitute lower limits because the presence of the competing ion will displace some of the primary ligand from the protein, and vice versa. However, they can be used as the starting point in a nested bisection routine to obtain the true $[Ca^{2+}]_i$ and $[Mg^{2+}]_i$ values. The two equations that must be simultaneously minimized are:

$$f([Ca^{2+}]_i) = [Ca^{2+}]_i + [M]_{t,i} \left(\frac{n_1 k_{1Ca} [Ca^{2+}]_i}{1 + k_{1Ca} [Ca^{2+}]_i + k_{1Mg} [Mg^{2+}]_i} + \frac{n_2 k_{2Ca} [Ca^{2+}]_i}{1 + k_{2Ca} [Ca^{2+}]_i + k_{2Mg} [Mg^{2+}]_i} \right) - [Ca^{2+}]_{t,i}. \quad (10.50)$$

$$f([Mg^{2+}]_i) = [Mg^{2+}]_i + [M]_{t,i} \left(\frac{n_1 k_{1Mg} [Mg^{2+}]_i}{1 + k_{1Ca} [Ca^{2+}]_i + k_{1Mg} [Mg^{2+}]_i} + \frac{n_2 k_{2Mg} [Mg^{2+}]_i}{1 + k_{2Ca} [Ca^{2+}]_i + k_{2Mg} [Mg^{2+}]_i} \right) - [Mg^{2+}]_{t,i}. \quad (10.51)$$

Typically, we begin with Eq. (10.50), using our preliminary estimates for x_i and y_i . The resulting improved value of x_i is used in Eq. (10.51) to obtain an improved value of y_i . That improved value for y_i is then recycled back into Eq. (10.50) to further refine our estimate of x_i . The process continues until the changes in x_i and y_i become acceptably small.

3.4. Analysis of the divalent ion binding by the S55D/E59D variant of rat α -parvalbumin

The CD site of the S55D/E59D variant is an example of a “pentacarboxylate” ligand array. Five of the six metal-ion coordination positions are occupied by side-chain carboxylates—aspartate at $+x$, $+y$, $+z$, and $-x$, glutamate at $-z$. Although this ligand constellation displays reduced metal

ion affinity, relative to a four-carboxylate site, in a model-peptide system (Marsden, Hodges, and Sykes, 1988), heightened Ca^{2+} affinity is observed when it is introduced into either parvalbumin binding site (Henzl, Hapak, and Goodpasture, 1996; Henzl, Agah, and Larson, 2004; Lee *et al.*, 2004). As mentioned earlier, the S55D/E59D variant of rat α -parvalbumin exhibits very high affinity for Ca^{2+} . A representative titration of α 55/59 with Ca^{2+} in 0.15 M KCl, 0.025 M Hepes-KOH, pH 7.4, is displayed in Fig. 10.4A. Note that the injection heats return abruptly to baseline values with the addition of two molar equivalents of Ca^{2+} , indicative of a high-affinity

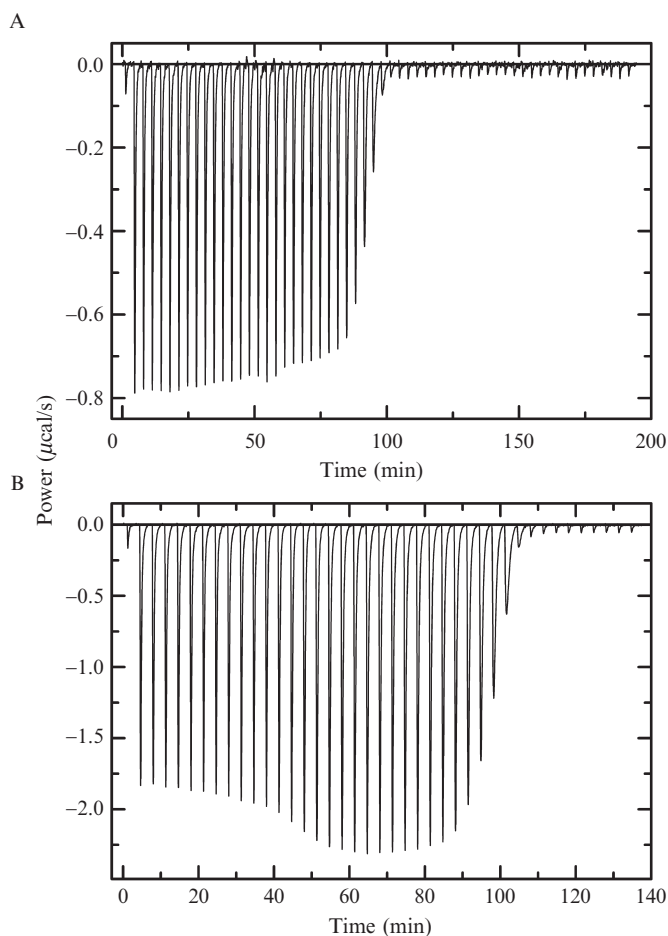


Figure 10.4 Representative raw data for the titration of rat α S55D/E59D with Ca^{2+} (A) and with Ca^{2+} in the presence of EGTA (B).

interaction. Fig. 10.4B presents raw data for the titration with Ca^{2+} in the presence of EGTA. Under our experimental conditions, the enthalpy change associated with the binding of Ca^{2+} to EGTA is substantially greater than that associated with either EF-hand motif in the parvalbumin. Thus, the appearance of the thermogram suggests that Ca^{2+} binding occurs in a highly ordered manner, with occupancy of the protein binding sites preceding that of the chelator, an indication that the Ca^{2+} affinity of the protein greatly exceeds that of EGTA.

Samples of α 55/59 were subjected to a battery of titrations comparable to that outlined in Table 10.1. Because the Ca^{2+} affinity is extremely high, the titrations in the presence of NTA were omitted. Instead, the protein was titrated with Ca^{2+} in the presence of two different levels of EDTA and EGTA. Integrated data for the titrations are displayed in Fig. 10.5. Analysis

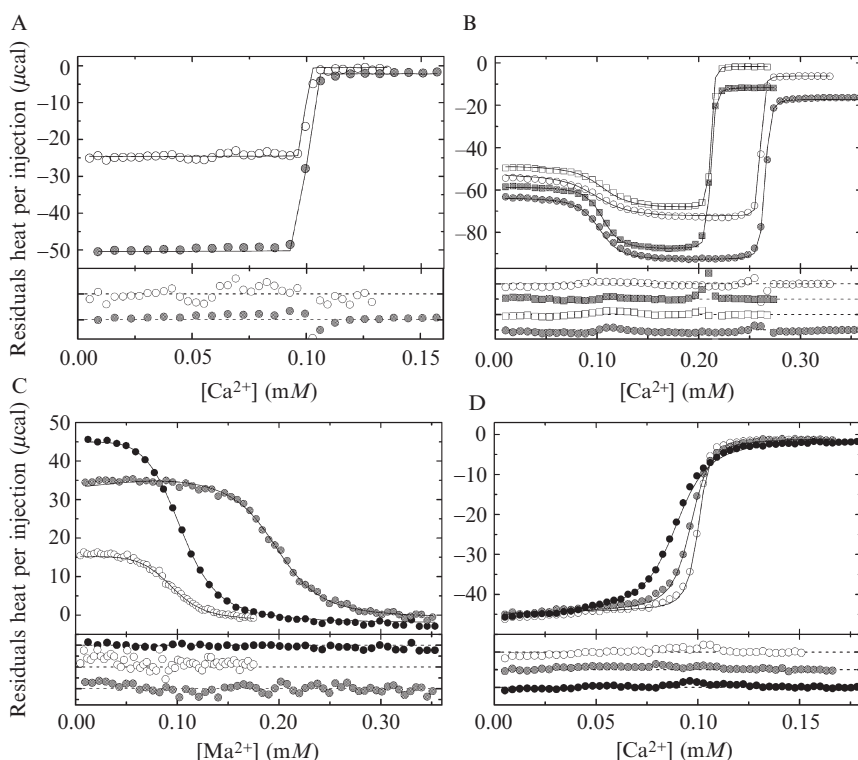


Figure 10.5 Integrated ITC data for the titration of rat S55D/E59D with Ca^{2+} (A), with Ca^{2+} in the presence of EDTA (B, white symbols) and EGTA (B, grey symbols), with Mg^{2+} (C, white and grey), with Mg^{2+} in the presence of EDTA (C, black), and with Ca^{2+} in the presence of Mg^{2+} (D) at 1 mM (white), 3 mM (grey), and 10 mM (black) Mg^{2+} . The solid lines reflect the best least-squares fit to the data.

of these data with the independent-site models described above yielded the optimal least-squares fit indicated by the solid lines in the figure. The optimal parameter values and associated uncertainties are listed in Table 10.2.

The first Ca^{2+} -binding constant, presumably corresponding to occupancy of the mutated CD binding site, approaches $2 \times 10^{10} \text{ M}^{-1}$. Clearly, the global analysis strategy is applicable to systems exhibiting very high divalent ion affinity. However, we have also applied it to systems having compromised Ca^{2+} affinity. For example, the combined D94S and G98E mutations in rat α -parvalbumin produce a CD-site ligand array in the environment of the EF site. The divalent ion affinity of the resulting “pseudo-CD site” is seriously attenuated, however, with the association constant down from $1.2 \times 10^8 \text{ M}^{-1}$ for the wild-type EF site to $5.4 \times 10^5 \text{ M}^{-1}$. Nevertheless, the global fitting method works equally well for this protein (Tanner *et al.*, 2005).

3.5. Analysis of positively cooperative divalent ion binding

Many ligand-binding phenomena exhibit positive cooperativity, the binding of oxygen by hemoglobin being the archetypal example. In these systems, the initial binding event provokes a conformational change that facilitates subsequent binding events. The Scatchard model, predicated on the apparent absence of binding-site interactions, fails for these systems. Instead, a more general binding equation is required.

For the two-site systems treated here, the extent of binding is described by the following equation:

$$\bar{X} = n \left(\frac{K_1 x_i + 2K_2 K_1 x_i^2}{1 + K_1 x_i + K_2 K_1 x_i^2} \right), \quad (10.52)$$

where the factor n allows for nonunity binding stoichiometries. The binding constants in Eq. (10.43), represented by capital letters, are stepwise macroscopic binding constants. The term “macroscopic” is applied because they describe the formation of the three macrostates of the protein: apo- (M), singly bound (MX), and doubly bound (MX_2). The term “stepwise” emphasizes their relevance to the individual ligation steps (i.e., $M \rightarrow MX \rightarrow MX_2$). Alternatively, the extent of binding can be described in terms of the overall macroscopic binding constants, β_1 and β_2 :

$$\bar{X} = n \left(\frac{\beta_1 x_i + 2\beta_2 x_i^2}{1 + \beta_1 x_i + \beta_2 x_i^2} \right). \quad (10.53)$$

The overall binding constants pertain to the overall conversion from the apo-protein to a given binding state (i.e., β_1 governs the $M \rightarrow MX$ reaction; β_2 governs the $M \rightarrow MX_2$ reaction).

Table 10.2 Divalent ion-binding parameters^a

	EDTA	EGTA	NTA	Rat α S55D/E59D ^b	PhI p 7 ^c
$\Delta H_{1,\text{Ca}}$	−6.3 (−6.4, −5.9)	−7.50 (−7.7, −7.3)	−1.9 (−2.1, −1.7)	−4.6 (−4.8, −4.4)	$\Delta H_{1,\text{Ca}}$ −1.9 (−2.0, −1.6)
$k_{1,\text{Ca}}$	4.4×10^7 (4.3, 4.5)	1.7×10^7 (1.6, 1.8)	1.1×10^4 (0.9, 1.2)	1.8×10^{10} (1.4, 2.2)	$K_{1,\text{Ca}}$ 1.7×10^6 (1.5, 2.0)
$\Delta H_{2,\text{Ca}}$	—	—	—	−4.6 (−4.8, −4.4)	$\Delta H_{2,\text{Ca}}$ −4.0 (−4.1, −3.7)
$k_{2,\text{Ca}}$	—	—	—	2.0×10^9 (1.6, 2.4)	$K_{2,\text{Ca}}$ 8.1×10^6 (6.8, 9.4)
$\Delta H_{1,\text{Mg}}$	4.3 (4.2, 4.5)	—	—	3.5 (3.4, 3.6)	$\Delta H_{1,\text{Mg}}$ 2.8 (2.6, 2.9)
$k_{1,\text{Mg}}$	5.5×10^5	—	—	2.7×10^6 (2.1, 3.3)	$K_{1,\text{Mg}}$ 2.8×10^4 (2.4, 3.1)
$\Delta H_{2,\text{Mg}}$	—	—	—	3.9 (3.8, 4.0)	$\Delta H_{2,\text{Mg}}$ 5.7 (5.2, 6.1)
$k_{2,\text{Mg}}$	—	—	—		$K_{2,\text{Mg}}$ 170 M^{-1} (147, 197)
—	—	—	—		$\beta_{\text{Ca,Mg}}$ 1.08×10^9 (0.49, 1.55)

^a Enthalpies are expressed in units of kcal/mol, binding constants in units of M^{-1} . Numbers in parentheses below the parameter values reflect the 95% confidence intervals.

^b Values originally reported in [Lee *et al.* \(2004\)](#).

^c Values originally reported in [Henzl *et al.* \(2008\)](#).

In terms of the stepwise constants, the function that must be minimized to obtain the free ligand concentration is given by:

$$f(x_i) = x_i + [M]_{t,i} n \left(\frac{K_1 x_i + 2K_2 K_1 x_i^2}{1 + K_1 x_i + K_2 K_1 x_i^2} \right) - [X]_{t,i}. \quad (10.54)$$

where n_C is included to allow for a nonunity binding stoichiometry for the chelator.

The corresponding expression for the cumulative heat of binding is:

$$Q_i = 10^6 V_o [M]_{t,i} n \left(\frac{\Delta H_1 K_1 x_i + (\Delta H_1 + \Delta H_2) K_2 K_1 x_i^2}{1 + K_1 x_i + K_2 K_1 x_i^2} \right). \quad (10.55)$$

If the titration is performed in the presence of a competing chelator, the function to be minimized becomes

$$f(x_i) = x_i + [M]_{t,i} n \left(\frac{K_1 x_i + 2K_2 K_1 x_i^2}{1 + K_1 x_i + K_2 K_1 x_i^2} \right) + [C]_{t,i} \left(\frac{n_C K_C x_i}{1 + K_C x_i} \right) - [X]_{t,i}, \quad (10.56)$$

where n_C is included to allow for a nonunity binding stoichiometry for the chelator.

The corresponding expression for the cumulative heat of binding is then given by

$$Q_i = 10^6 V_o \left([M]_{t,i} n \left(\frac{\Delta H_1 K_1 x_i + (\Delta H_1 + \Delta H_2) K_2 K_1 x_i^2}{1 + K_1 x_i + K_2 K_1 x_i^2} \right) + [C]_{t,i} \left(\frac{\Delta H_c n_C K_C x_i}{1 + K_C x_i} \right) \right). \quad (10.57)$$

If the protein is titrated with ligand X in the presence of competing ion Y , then, as described earlier, we are faced with the task of obtaining self-consistent estimates for both x_i and y_i . The situation is somewhat more complicated, however, owing to the possibility of formation of a mixed MXY species. The partition function, or binding polynomial, for this case must include the contributions from M , MX , MX_2 , MY , MY_2 , and MXY . The extent of binding of ligand X for this case becomes:

$$\bar{X} = n \left(\frac{K_1 x_i + 2K_2 K_1 x_i^2 + \beta_{xy} x_i y_i}{1 + K_1 x_i + K_2 K_1 x_i^2 + \beta_{xy} x_i y_i + K_{1y} y_i + K_{2y} K_{1y} y_i^2} \right). \quad (10.58)$$

The corresponding expression for Y is

$$\bar{Y} = n \left(\frac{K_{1y}y_i + 2K_{2y}K_{1y}y_i^2 + \beta_{xy}x_iy_i}{1 + K_1x_i + K_2K_1x_i^2 + \beta_{xy}x_iy_i + K_{1y}y_i + K_{2y}K_{1y}y_i^2} \right). \quad (10.59)$$

Thus, the following two equations must be simultaneously minimized:

$$f(x_i) = x_i + [M]_{t,i}n \left(\frac{K_1x_i + 2K_2K_1x_i^2 + \beta_{xy}x_iy_i}{1 + K_1x_i + K_2K_1x_i^2 + \beta_{xy}x_iy_i + K_{1y}y_i + K_{2y}K_{1y}y_i^2} \right) - [X]_{t,i}. \quad (10.60)$$

$$f(y_i) = y_i + [M]_{t,i}n \left(\frac{K_{1y}y_i + 2K_{2y}K_{1y}y_i^2 + \beta_{xy}x_iy_i}{1 + K_1x_i + K_2K_1x_i^2 + \beta_{xy}x_iy_i + K_{1y}y_i + K_{2y}K_{1y}y_i^2} \right) - [Y]_{t,i}. \quad (10.61)$$

The expression for the cumulative heat must include the contributions from binding of both ions:

$$Q_i = \Delta H_1\alpha_{1,i} + \Delta H_{1y}\alpha_{1y,i} + \Delta H_2\alpha_{2,i} + \Delta H_{2y}\alpha_{2y,i} + \Delta H_{xy}\alpha_{xy,i}, \quad (10.62)$$

where ΔH_{xy} is the enthalpy associated with formation of the mixed complex and $\alpha_{xy,i}$ is the fraction of the macromolecule present as the mixed complex.

3.6. Modeling divalent ion binding by Phl p 7

Phl p 7 belongs to the family of proteins called polcalcins, small pollen-specific proteins containing two EF-hand motifs (Tinghino *et al.*, 2002; Valenta *et al.*, 1998;). Although their functions are presently unknown, they are believed to function in pollen-tube elongation. Interestingly, they are potent allergens.

Phl p 7 is expressed naturally in timothy grass, *Phleum pratense*. The cDNA sequence, which was expressed in *E. coli*, codes for 77 amino acids plus the initiation codon. Although the protein crystallizes as an unusual domain-swapped dimer at pH 3.4 (Verdino *et al.*, 2003), sedimentation analysis at neutral pH shows the protein to be exclusively monomeric in both apo- and Ca^{2+} -loaded states (Henzl *et al.*, 2008). Our interest in the protein grew from the observation that EF-hand 2 in Phl p 7 has a pentacarboxylate ligand array identical to that in the S55D or G98D variants

of rat β -parvalbumin (i.e., aspartyl carboxylates at $+\alpha$, $+\gamma$, $+\zeta$, and $-\alpha$; main-chain carbonyl at $-\gamma$; and glutamate at $-\zeta$). As discussed earlier, this ligand array imparts heightened divalent ion affinity on the parvalbumin CD and EF sites. It was therefore of interest to see how it performed in an entirely different protein background.

Samples of Phl p 7 were subjected to a battery of titrations similar to that employed for the S55D/E59D variant of rat α -parvalbumin, but including titrations in the presence of NTA. Raw data for the titration with Ca^{2+} and with Ca^{2+} in the presence of EGTA are displayed in Fig. 10.6. The difference in the appearance between the raw data for this protein and that of the S55D/E59D variant offers some insight into the capacity of this method for extracting binding parameters from very diverse systems. Integrated data for the entire suite of titrations are presented in Fig. 10.7. The independent two-site model described above failed to produce a

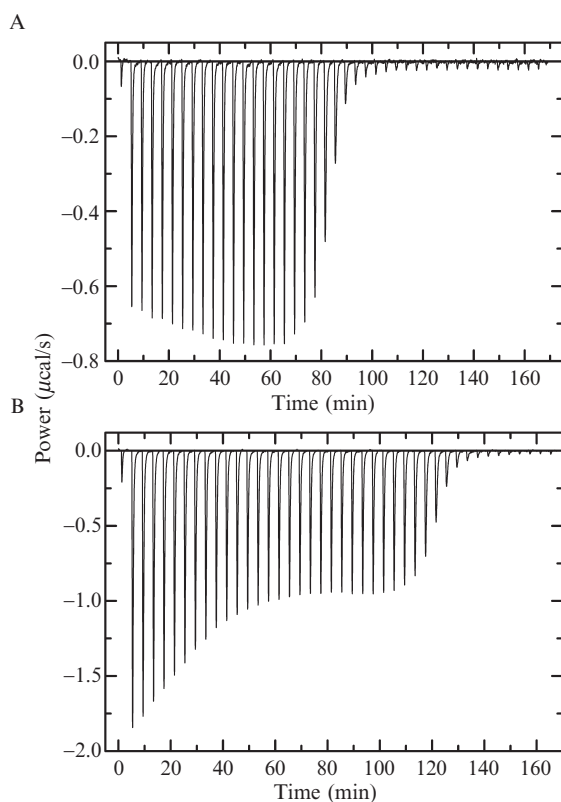


Figure 10.6 Representative raw data for the titration of the polycalcin Phl p 7 with Ca^{2+} (A) and with Ca^{2+} in the presence of EGTA (B).

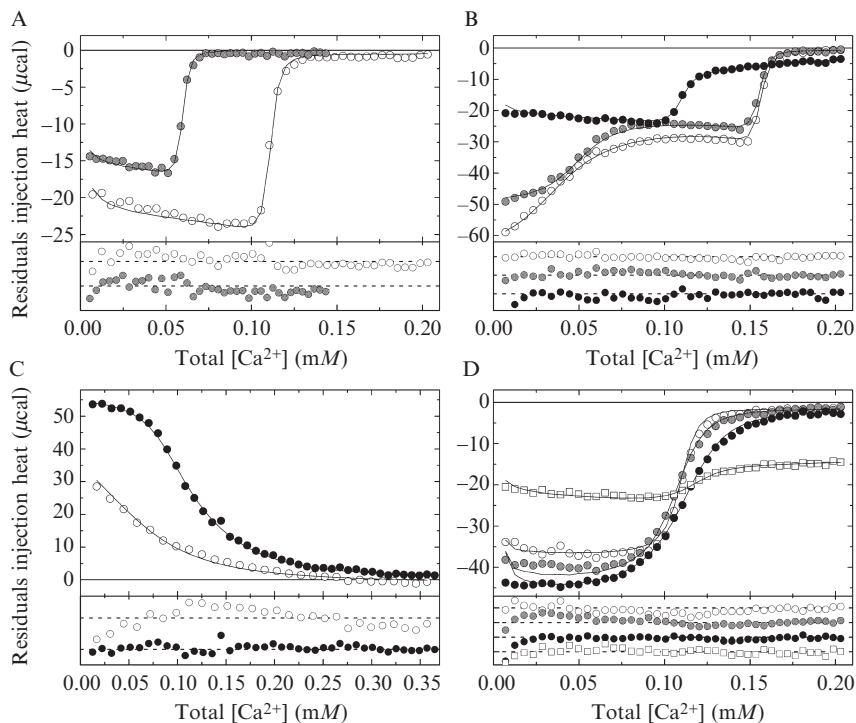


Figure 10.7 Integrated ITC data for the titration of Phl p 7 with Ca^{2+} (A, white and grey), with Ca^{2+} in the presence of EGTA (B, white), EDTA (B, grey), and 0.1 mM NTA (B, black), with Mg^{2+} (C, white), with Mg^{2+} in the presence of EDTA (C, black), with Ca^{2+} in the presence of Mg^{2+} at 1.0 mM (D, white), 5.0 mM (D, grey), and 10 mM (D, black), and with Ca^{2+} in the presence of 1.0 mM NTA (D, white rectangles). The solid lines represent the optimal least-squares fit.

satisfactory fit to the data. In retrospect, this result was not surprising. Because the polcalcins are proposed to function as Ca^{2+} -dependent regulatory proteins, it was quite plausible that Ca^{2+} binding would be accompanied by a substantial conformational change, which might well be manifested in positively cooperative metal-ion binding. Because positive macroscopic cooperativity cannot be described by a Scatchard-type model, it was necessary to employ a general Adair model.

The optimal fit obtained with this system of equations is indicated by the solid lines in Fig. 10.7. The corresponding parameter values are listed in Table 10.2. The results yielded substantial insight into this protein. The binding of Ca^{2+} was decidedly cooperative, with K_1 and K_2 values of 1.7×10^6 and 8.0×10^6 , respectively. These values yield a ratio of $K_2:K_1$ of 4.7. For equivalent binding sites, in the absence of cooperativity, this ratio would be 0.25. Thus, the minimum coupling free energy for Ca^{2+} binding

to Phl p 7 is $-RT \ln(4.7/0.25) = -1.7$ kcal/mol. In fact, the coupling energy could be substantially larger because the two sites EF-sites are decidedly nonequivalent. Whereas site 1 has just three anionic ligands, carboxylates occupy five of the six coordination sites in site 2. As discussed previously, in the rat α - and β -parvalbumin isoforms, the pentacarboxylate ligand arrays exhibit heightened divalent ion affinity. It is therefore likely that the initial Ca^{2+} -binding event occurs at site 2.

The positively cooperative Ca^{2+} -binding behavior can be interpreted in terms of a classic allosteric model, in which the protein is capable of occupying two distinct conformations having different affinities for ligand. In this case, the molecule resides largely in the low-affinity *T* state in the absence of ligand. Binding of the first ligand is discouraged because the preferred form for binding is present at very low concentrations and because the net binding free energy includes a penalty for the conformational change to the high-affinity *R* state. By contrast, binding of the second ligand occurs more readily because the protein resides almost exclusively in the *R* state. Thus, the net affinity associated with binding the second ligand can exceed that for the first, even though the second site has intrinsically lower affinity for the ligand.

Spectroscopic data indicate that, indeed, Ca^{2+} binding is accompanied by a substantial conformational change. Far-UV circular dichroism measurements reveal a significant increase in molar ellipticity in the presence of Ca^{2+} . Moreover, ANS fluorescence increases markedly in the presence of Ca^{2+} -bound Phl p 7, whereas it is virtually unchanged in the presence of the apo-protein.

In contrast to Ca^{2+} binding, the binding of Mg^{2+} to Phl p 7 is highly sequential. The two association constants— 2.78×10^4 and 170 M^{-1} —differ by more than two orders of magnitude.

Consistent with the apparent absence of cooperativity, the binding of Mg^{2+} does not appear to provoke a significant conformational change. The CD spectrum is unchanged in the presence of Mg^{2+} , and ANS fluorescence does not increase upon addition of the Mg^{2+} -bound Phl p 7.

Typically, the ratio of Ca^{2+} - to Mg^{2+} -binding constants for an EF-hand site is on the order of $1\text{--}2 \times 10^4$. In Phl p 7, however, the ratio of binding constants ($K_1: K_{1M}$) associated with site 2 is a mere 60. Two factors conspire to produce this anomalously low value. First, whereas Ca^{2+} -binding at site 2 is linked to an energetically costly conformational change, the Mg^{2+} -binding event evidently is not. Thus, whereas K_1 is substantially smaller than the intrinsic binding constant, K_{1M} approaches the intrinsic value. Secondly, the pentacarboxylate ligand array confers exceptionally high intrinsic Mg^{2+} affinity on site 2. By contrast, the ratio of Ca^{2+} and Mg^{2+} binding constants associated with site 1 is atypically large, at 4.7×10^4 . Presumably, this elevated ratio reflects the enhancement of site 1 Ca^{2+} affinity, but not Mg^{2+} affinity, by the positively cooperative interaction with site 2.

4. CONCLUSION

Global analysis of direct and competitive ITC experiments offers a robust alternative to radiochemical or fluorescence-based methods for measuring divalent ion affinity in EF-hand proteins. The approach can be used to model either cooperative or noncooperative binding data. It has been used to estimate Ca^{2+} -binding constants exceeding 10^{10} M^{-1} , as well as Mg^{2+} -binding constants on the order of 10^2 M^{-1} . The global ITC strategy has the advantage over other methods that it furnishes estimates of the binding enthalpies for Ca^{2+} and Mg^{2+} directly, thereby providing a more complete thermodynamic description of divalent ion binding.

ACKNOWLEDGMENT

This work was supported by NSF award MCB0543476 (to M.T.H.).

REFERENCES

- Beechem, J. M. (1992). Global analysis of biochemical and biophysical data. *Methods Enzymol.* **210**, 37–54.
- Bevington, P. R. (1969). *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York.
- Bevington, P. R., and Robinson, D. K. (1992). *Data Reduction and Error Analysis for the Physical Sciences*, 2nd ed. WCB/McGraw-Hill, Boston.
- Celio, M. R., Pauls, T., and Schwaller, B. (1996). *Guidebook to the Calcium-Binding Proteins*, Oxford University Press, Oxford.
- Eatough, D. (1970). Calorimetric determination of equilibrium constants for very stable metal-ligand complexes. *Anal. Chem.* **42**, 635–639.
- Freyer, M. W., and Lewis, E. A. (2008). Isothermal titration calorimetry: Experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods Cell Biol.* **84**, 79–113.
- Fukada, H., and Takahashi, K. (1998). Enthalpy and heat capacity changes for the proton dissociation of various buffer components in 0.1 M potassium chloride. *Proteins* **33**, 159–166.
- Goldberg, R. N., Kishore, N., and Lennen, R. M. (2002). Thermodynamic quantities for the ionization reactions of buffers. *J. Phys. Chem. Ref. Data* **31**, 231–370.
- Haiech, J., Derancourt, J., Pechere, J. F., and Demaille, J. G. (1979). Magnesium and calcium binding to parvalbumins: Evidence for differences between parvalbumins and an explanation of their relaxing function. *Biochemistry* **18**, 2752–2758.
- Heizmann, C. W., and Kagi, U. (1989). Structure and function of parvalbumin. *Adv. Exper. Med. Biol.* **255**, 215–222.
- Henzl, M. T., Agah, S., and Larson, J. D. (2004). Rat α - and β -parvalbumins: Comparison of their pentacarboxylate and site-interconversion variants. *Biochemistry* **43**, 9307–9319.
- Henzl, M. T., Davis, M. E., and Tan, A. (2008). Divalent ion binding properties of the timothy grass allergen, Phl p 7. *Biochemistry* **47**, 7846–7856.

- Henzl, M. T., Hapak, R. C., and Goodpasture, E. A. (1996). Introduction of a fifth carboxylate ligand heightens the affinity of the oncomodulin CD and EF sites for Ca^{2+} . *Biochemistry* **35**, 5856–5869.
- Hu, D. D., and Eftink, M. R. (1994). Thermodynamic studies of the interaction of trp aporepressor with tryptophan analogs. *Biophys. Chem.* **49**, 233–239.
- Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* **36**, 575–588.
- Johnson, M. L., and Faunt, L. M. (1992). Parameter Estimation by Least-Squares Methods. *Methods Enzymol.* **210**, 1–37.
- Kawasaki, H., and Kretsinger, R. H. (1995). Calcium-binding proteins 1: EF-hands. *Protein Profile* **2**, 297–490.
- Kretsinger, R. H. (1980). Structure and evolution of calcium-modulated proteins. *CRC critical reviews in biochemistry* **8**, 119–174.
- Kretsinger, R. H., and Nockolds, C. E. (1973). Carp muscle calcium-binding protein. II. Structure determination and general description. *J. Biol. Chem.* **248**, 3313–3326.
- Ladbury, J. E., and Chowdhry, B. Z. (1996). Sensing the heat: The application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. *Chem. Biol.* **3**, 791–801.
- Leavitt, S., and Freire, E. (2001). Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol.* **11**, 560–566.
- Lee, Y. H., Tanner, J. J., Larson, J. D., and Henzl, M. T. (2004). Crystal structure of a high-affinity variant of rat α -parvalbumin. *Biochemistry* **43**, 10008–10017.
- Lewis, E. A., and Murphy, K. P. (2005). Isothermal titration calorimetry. *Methods Mol. Biol.* **305**, 1–16.
- Marsden, B. J., Hodges, R. S., and Sykes, B. D. (1988). ^1H NMR studies of synthetic peptide analogues of calcium-binding site III of rabbit skeletal troponin C: Effect on the lanthanum affinity of the interchange of aspartic acid and asparagine residues at the metal ion coordinating positions. *Biochemistry* **27**, 4198–4206.
- Pauls, T. L., Cox, J. A., and Berchtold, M. W. (1996). The Ca^{2+} -binding proteins parvalbumin and oncomodulin and their genes: New structural and functional findings. *Biochim. Biophys. Acta* **1306**, 39–54.
- Pepper, D. S. (1998). Some Alternative Coupling Chemistries for Affinity Chromatography, in *Protein Protocols on CD-ROM*, (Walker, ed.), Humana Press, Totowa, N.J.
- Philo, J. S. (1997). An improved function for fitting sedimentation velocity data for low-molecular-weight solutes. *Biophys. J.* **72**, 435–444.
- Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619.
- Sigurskjold, B. W. (2000). Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal. Biochem.* **277**, 260–266.
- Sigurskjold, B. W., Berland, C. R., and Svensson, B. (1994). Thermodynamics of inhibitor binding to the catalytic site of glucoamylase from *Aspergillus niger* determined by displacement titration calorimetry. *Biochemistry* **33**, 10191–10199.
- Strynadka, N. C. J., and James, M. N. G. (1989). Crystal structures of the helix-loop-helix calcium-binding proteins. *Ann. Rev. Biochem.* **58**, 951–998.
- Tanner, J. J., Agah, S., Lee, Y. H., and Henzl, M. T. (2005). Crystal structure of the D94S/G98E variant of rat α -parvalbumin. An explanation for the reduced divalent ion affinity. *Biochemistry* **44**, 10966–10976.
- Tellinghuisen, J. (2004). Statistical error in isothermal titration calorimetry. *Methods Enzymol.* **383**, 245–282.
- Tinghino, R., Twardosz, A., Barletta, B., Puggioni, E. M., Iacovacci, P., Butteroni, C., Afferni, C., Mari, A., Hayek, B., Di, F. G., Focke, M., Westritschnig, K., et al. (2002).

- Molecular, structural, and immunologic relationships between different families of recombinant calcium-binding pollen allergens. *J. Allergy & Clin. Immunol.* **109**, 314–320.
- Valenta, R., Hayek, B., Seiberler, S., Bugajska-Schretter, A., Niederberger, V., Twardosz, A., Natter, S., Vangelista, L., Pastore, A., Spitzauer, S., and Kraft, D. (1998). Calcium-binding allergens: From plants to man. [Review] [53 refs]. *Intl. Arch. Allergy & Immunol.* **117**, 160–166.
- Verdino, P., Westritschnig, K., Valenta, R., and Keller, W. (2003). Three-dimensional structure of the panallergen Phl p 7. *Intl. Arch. Allergy & Immunol.* **130**, 10–11.