Articles

Hierarchical Modeling of Phenolic Ligand Binding to 2Zn−Insulin Hexamers

Duane T. Birnbaum,[‡] Steven W. Dodd,[‡] Bo E. H. Saxberg,[§] Alexander D. Varshavsky,[§] and John M. Beals*,[‡]

Departments of Biopharmaceutical Development and Information Sciences, Eli Lilly and Company, Lilly Research Laboratories, Indianapolis, Indiana 46285

Received January 10, 1996; Revised Manuscript Received February 15, 1996[⊗]

ABSTRACT: Phenolic ligands, *e.g.*, phenol and *m*-cresol, bind to 2Zn(II)−insulin hexamers and induce a conformational change at the N-terminus of the B-chain for each monomer. The binding of these phenolic ligands to 2Zn(II)−insulin hexamers has been studied by isothermal titrating calorimetry (ITC). The binding isotherms were modeled and thermodynamic parameters were quantified using a novel, flexible algorithm that permitted the development of a hierarchical series of physical models. With the insulin hexamer represented as a dimer of trimers, the modeling demonstrated that ligand binding is highly cooperative in nature, both intra- and inter-trimer. The isotropic inter-trimer cooperativity was dominant and negative in every system studied, with initial binding constants typically an order of magnitude greater for the binding of ligands to the first trimer relative to the second. The inter-trimer cooperativity estimated from the modeling of solution calorimetry data is consistent with a $T_6 \leftrightarrow T_3R_3 \leftrightarrow R_6$ equilibrium first proposed from crystallographic investigations. Intra-trimer cooperativity was present only in the enthalpy coefficient space, not in the equilibrium coefficient space, and therefore, less of a factor. The order of binding affinity for the ligands studied is resorcinol ≫ phenol ≥ *m*-cresol as determined from their overall free energies of binding to the 2Zn(II)−insulin hexamer (−26.6, −23.4, and −23.4 kcal/mol, respectively) and their intrinsic binding constants (8780, 5040, and 3370 L/mol, respectively) at 14 °C. The temperature dependence of phenol binding to 2Zn(II)−insulin hexamer was modeled. Increasing temperature decreased the magnitude of both the intrinsic binding constant and the inter-trimer cooperativity. The second phase of the ITC binding profile was also found to be highly temperature dependent. At lower temperatures the second phase is endothermic but gradually decreases with increasing temperature and subsequently becomes exothermic. This effect is attributed to loss of water from the hydration shell of the insulin hexamer with increasing temperature and consequently reduces the entropic contributions to the T ↔ R transition in the phenol/2Zn(II)−insulin hexamer system.

Insulin is a two-chained molecule consisting of a 21 amino acid A-chain and a 30 amino acid B-chain linked by two disulfide bonds (Baker et al., 1988). In zinc-free solutions, insulin monomers self-associate to form dimers and higher aggregate species (Pekar & Frank, 1972). However, in the presence of various divalent metal ions (at 0.3 mol/monomer), insulin associates into a discrete hexameric structure (Goldman & Carpenter, 1974).

Commercial insulin preparations contain phenolic excipients (*e.g.*, phenol, *m*-cresol, and methylparaben) as antimicrobial agents. These phenolic species also bind to specific sites on 2Zn−insulin hexamers,[1] In_6Zn_2, causing a conformational change that increases the chemical stability of insulin in commercial preparations (Brange & Langkjær, 1992). X-ray crystallographic data have identified the location of six binding sites on the insulin hexamer and the nature of the conformational change that the binding of these phenolic ligands induces (Derewenda et al., 1989). The phenolic ligands are stabilized in a binding pocket between adjacent monomers by hydrogen bonds with the carbonyl oxygen of ^{A6}Cys and the amide proton of ^{A11}Cys and numerous van der Waals contacts. The binding of these ligands stabilizes a conformational change that occurs at the N-terminus of the B-chain in each insulin monomer, shifting the conformational equilibrium of residues B1 to B8 from an extended structure (T-state) to an α-helical structure (R-state), referred to as a T ↔ R transition.

Spectroscopic (Wollmer et al., 1987) and crystallographic (Baker et al., 1988; Derewenda et al., 1989; Smith & Ciszak, 1994) data on ligand-bound, 2Zn−insulin and/or 2Co−insulin hexamers have identified three distinct hexameric species: T_6, T_3R_3, and R_6, where T_3R_3 is an intermediate structure with the ligated insulin molecules of one trimer in a R-state conformation and the unligated insulin molecules in the other trimer in a T-state conformation. The metal coordination state in the T_3 trimer is octahedral, with three

* Author to whom correspondence should be addressed.
[‡] Biopharmaceutical Development.
[§] Mathematical and Statistical Sciences.
[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1996.
[1] Abbreviations: ITC, isothermal titration calorimetry; LFSE, ligand field stabilization energy; CD, circular dichroism spectroscopy; MWCO, molecular weight cutoff; SSE, sum of squared error; T and R are used to designate the extended and α-helical conformations (respectively) of residues B1−B8 of the insulin monomer; 2Zn(II)−insulin hexamer and In_6Zn_2 may represent any conformation of the insulin hexamer. T_6, T_3R_3, and R_6 are used to designate the three crystallographically identified conformations of the insulin hexamer.

S0006-2960(96)00055-4 CCC: $12.00

nitrogen atoms from the symmetry related His^{B10} residues coordinating the metal atom. Water molecules complete the coordination sphere. In a R_3 trimer, the complex is pseudo-tetrahedral with the same three insulin ligands as in the T-state, plus a monovalent anion. Thus, it has been proposed that the insulin hexamer behaves as a dimer of trimers, and the binding process is represented by the following: $T_6 \leftrightarrow T_3R_3 \leftrightarrow R_6$ (Brader & Dunn, 1991).

Solution studies on 2Zn− and/or 2Co−insulin have qualitatively identified the presence of positive and negative homotropic cooperativity during the binding of these ligands (Brader & Dunn, 1991; Krüger et al., 1990; Roy et al., 1989). In 2Co−insulin hexamers, the intra-trimer ligand binding is thought to be modulated by positive cooperativity and the inter-trimer ligand binding modulated by negative cooperativity (Krüger et al., 1990). However, no positive intra-trimer cooperativity was observed in 2Zn−insulin hexamers using CD spectroscopy (Krüger et al., 1990).

Recently, two attempts have been made at modeling the $T \leftrightarrow R$ transition in insulin hexamers (Bloom et al., 1995, Jacoby et al., 1993). One group used a three-parameter model to reproduce binding curves generated by CD spectroscopy (Jacoby et al., 1993). The authors treated the $T_6 \leftrightarrow T_3R_3$ and $T_6 \leftrightarrow R_6$ equilibria separately, fitting each region of the binding curve to linearized third order polynomial equations. Each polynomial was a function of an intrinsic binding constant, equilibrium constants for the $T_6 \leftrightarrow T_3R_3$ and $T_6 \leftrightarrow R_6$ equilibria, free ligand concentration, and a factor related to ellipticity differences between the T_3R_3 intermediate and R_6 states to the T_6 state. The authors assumed that the free ligand concentration was equal to that of total ligand concentration. An alternate approach modeled binding curves measured by monitoring changes in the absorbance of Co^{+2} as it changed its coordination state from octahedral to tetrahedral with the addition of phenolic ligands[2] (Bloom et al., 1995). The change in metal ion coordination parallels the $T \leftrightarrow R$ conformational change in insulin. The four parameters extracted from the model yielded quantitative information for the $T_6 \leftrightarrow T_3R_3$ and $T_6 \leftrightarrow R_6$ equilibrium constants as well as relative binding affinities to each trimer for the specific ligand. Again, the authors assumed that free ligand concentration was equal to the total ligand concentration. Possible concerns with these models are the lack of accounting for intra-trimer cooperativity, and the restrictions of inter-trimer cooperativity to be completely isotropic in nature. A more serious concern in both models is the assumption that $[\text{ligand}]_{\text{free}} = [\text{ligand}]_{\text{total}}$. For a ligand such as phenol, with a relatively moderate affinity for the insulin hexamer, at least 15% of the total ligand present in solution is bound to the hexamer at saturation. At the early stages of the titration, the percent phenol bound will be even higher because the first trimer displays a much greater affinity for phenolic ligands relative to the second trimer, and the number of available sites is at its greatest.

Microcalorimetry provides some distinct advantages over spectroscopic methods, such as CD, for measuring the binding of various ligands to proteins, and in particular the binding of phenolic ligands to the insulin hexamer. There are no optical restrictions on the metal ion used, and interference from the intense absorption of phenolic chromophores is not an issue as it is with CD. The high sensitivity of the method to changes in heat induced by the binding of ligands allows for greater sampling densities within a binding curve. Higher sampling densities are necessary for the complex modeling process used in this paper, especially in regions of the binding curve where significant changes in slope occur.

To estimate the various individual heats of binding, we constructed a physical model that represents all possible species formed by the binding of phenolic ligands to the insulin hexamer. In the following treatment, observations in the literature are used to condition the parametrization of the physical model, *i.e.*, there is a focus on representing the transition as the allosteric behavior of a dimer of trimers. The qualitative and quantitative nature of this cooperative binding process (positive, negative, symmetric, isotropic, anisotropic, etc.) will then be revealed by a systematic search through a hierarchical series of models and subsequent location of an "optimal" model and its resultant set of parameters. This approach is able to obtain quantitative information from a complex set of binding curves created by a system with multiple interdependent binding sites. The hierarchical nature of the modeling is the key to extracting subtleties such as cooperativity and allows for the direct comparison of parameters obtained from fitting different systems, regardless of the dimensionality of the model used.

Due to the wealth of structural information available, hexameric insulin provides an excellent system in which to test our approach. The complexity created by the large conformational change upon binding of phenolic ligands, its cooperative nature, and multiple binding sites is a unique challenge. With the successful treatment of the phenolic ligand/insulin hexamer binding system, the model may then be applied to other protein-binding systems where less structural information is known.

MATERIALS AND METHODS

Biosynthetic human zinc−insulin crystals (BHI), distilled *m*-cresol (99%), and liquefied distilled phenol (89%) were obtained through Eli Lilly and Co. (Indianapolis, IN) and were of the highest pharmaceutical grade. Molecular biology grade Tris buffer, resorcinol (99%), sodium chloride (99%), and zinc(II) chloride (98%) were used as purchased. All solutions were prepared with water obtained from the Milli-Q Plus water purification system (Millipore Co., Bedford, MA). Trace metal analysis revealed the presence of 17 metals at concentrations of <0.1−0.001 ppm. Sodium concentrations were <1.0 ppm.

Isothermal Titration Calorimetry. Human zinc−insulin crystals were dissolved in water to yield approximately 10 mL of stock solution at 35−40 mg of BHI/mL and subsequently adjusted to pH 7.8 with a dilute NaOH solution. The solution was dialyzed (Spectra/Por*7 regenerated cellulose dialysis membrane; nominal MWCO 2000) against 4 L of ultrapure water for 5−6 h. Insulin concentrations were determined from its absorption at 276 nm, $\epsilon_{276} = 6100$ M^{-1} cm^{-1} (Frank & Veros, 1968). An aliquot of the insulin buffer was diluted with water; then aliquots of Tris chloride

[2] Cobalt(II) was used because zinc(II) does not absorb UV−visible radiation. Cobalt(II) also has a preference for an octahedral coordination state relative to tetrahedral, which is thought to modulate the degree of intra-trimer cooperativity (Krüger et al., 1990). Therefore, studies with Co(II)−insulin complexes are not directly transferable to pharmaceutically relevant formulations.

(final concentration 25 mM) and Zn^{2+} were added such that the metal/insulin mole ratio was 0.33. The anion, Cl^-, was added as NaCl to insulin and phenolic ligand solutions prior to the titration to maintain a constant level at 50 mM. The pH was adjusted to 8.0 using dilute base and stirred for one half-hour. The solution was filtered through a 0.22-μm (Whatman Anotop 10+) filter prior to filling the calorimeter.

Stock solutions of phenolic ligand between 10 and 250 mM were prepared by dissolving the material in Milli-Q H_2O. Tris chloride buffer and NaCl were added such that their concentrations and the solution pH matched that of the insulin sample solution (25 and 50 mM, respectively; pH = 8.0). Two different calorimeters were used, but only one for any given data set, as defined by the preservative and temperature conditions. The solution was drawn into a 100-μL or 250-μL injection/stirrer syringe without entrapping air bubbles and inserted into the MicroCal Omega or MCS Reaction Cell (Northampton, MA).

The MicroCal Omega or the MCS Reaction Cell was equilibrated and calibrated at the desired temperature. Six titrations were conducted at three insulin concentrations, generating a total of 18 binding curves. The titration protocol for half of the experiments consisted of 50 injections of variable size (1−15 μL) at 5-min intervals (long curves). The smaller injection sizes allowed for higher sampling density in regions of greatest heat change. To further increase the sampling density at the critical, initial stage of the binding curve, the other half of the experiments were performed with fewer injections (21) and with the phenolic ligand solution at one-tenth the concentration (short curves). The last 5−6 injections of the short curves overlapped (with respect to ligand concentration) the first few injections of the long curves. Short and long curves were then linked into one profile. A filter time of 2 s was used for each data point recorded by the calorimeter. Because the MicroCal Omega/MCS calorimeter uses a constant volume cell (1.25 mL for the Omega calorimeter and 1.30 mL for the MCS calorimeter), each injection of ligand solution displaces an equal volume of sample out of the cell. This displacement offsets the macromolecule concentration, ligand concentration, and the measured heat. The contribution of the displaced volume to the measured heat will depend on the mixing time and kinetics of the reaction. Estimated corrections for these heats were, in all cases, negligible (typically <0.1%, compared to a standard experimental error of approximately 5−10%). The resulting data, obtained as heat per injection, are integrated to yield the binding isotherm of total heat versus ligand concentration. The manufacturer's software corrects the concentration of ligand and macromolecule for displaced volume and outputs the corrected data in a spreadsheet (MicroCal, 1993). The corrected ligand concentrations along with the heat per mole of injectant were used in the fitting procedure to estimate thermodynamic parameters of ligand binding. The macromolecule concentration was held constant at its initial value during this fitting routine to ensure proper limiting behavior in the integral curves (described later). As a check, the fitting procedure for parameter estimation was also run with the macromolecule concentration held constant at its final value after saturation (≤10% of initial). The resulting fits for the two values of macromolecule concentrations (initial and final) yielded identical equilibrium coefficients while enthalpic parameters varied by less than 5%. Therefore, constraining the variation occurring in macromolecule concentration did not have a significant effect on the parameter estimations. Heats of dilution were obtained by fitting the last 10−15 injections past saturation in the binding isotherm to a straight line. The estimated heats of dilution (equation for the straight line) were subtracted from the raw data. Heats of dilution were measured for the short curves by titrating the same ligand solution into buffer. In all cases, heats of dilution for the short curves were found to be negligible. The excellent overlap achieved (with respect to heat per mole of injectant) between the beginning of the long curves and end of the short curves justifies the method used for subtracting heats of dilution from the long curves.

Hierarchical Modeling. The insulin hexamer is configured as an association of two trimers, with a twofold rotational axis of symmetry between trimers, and a threefold longitudinal rotational axis of symmetry bisecting the twofold axis (Baker et al., 1988). As shown in Figure 1, there are nine possible states for such a configuration with up to six ligands bound. The calorimetrically determined heat of binding, ΔQ, is calculated from the product of an average enthalpy change due to each species, ΔH, the individual equilibrium binding coefficients, K, the insulin monomeric concentration, $[\mathrm{Ins}]_{tot}$, and the reaction volume, V_0 (eq 1). The subscripts on K and ΔH, in eqs 1 and 2, correspond to the individual species identified in Figure 1. The enthalpy change associated with the binding of each ligand represents the simultaneous sums of all possible ligated species present in solution. There is no preference given to a specific binding pathway through the various species represented in the equilibria. The resulting parameters obtained from the best fit to the data will yield information on intrinsic and cooperative processes involved in the transition and will therefore highlight a preferred path. The equations below are similar to those used by Schön and Freire in their treatment of the binding of cholera toxin to the oligiosaccharide portion of its cell surface receptor, ganglioside G_{M1} (Schön & Freire, 1989).

$$\begin{aligned}\Delta Q = V_o([\mathrm{Ins}]_{tot}/6)[\{&\Delta H_{10}K_{10}[\mathrm{L}] + \\ &(\Delta H_{20}K_{20}K_{10} + \Delta H_{11}K_{11}K_{10})[\mathrm{L}]^2 + \\ &(\Delta H_{30}K_{30}K_{20}K_{10} + \Delta H_{21}K_{21}K_{20}K_{10})[\mathrm{L}]^3 + \\ &(\Delta H_{31}K_{31}K_{30}K_{20}K_{10} + \Delta H_{22}K_{22}K_{21}K_{20}K_{10})[\mathrm{L}]^4 + \\ &(\Delta H_{32}K_{32}K_{31}K_{30}K_{20}K_{10})[\mathrm{L}]^5 + \\ &(\Delta H_{33}K_{33}K_{32}K_{31}K_{30}K_{20}K_{10})[\mathrm{L}]^6\}/ \\ \{1 + K_{10}[\mathrm{L}] + (K_{20}K_{10} + K_{11}K_{10})[\mathrm{L}]^2 + (K_{30}K_{20}K_{10} &+ \\ K_{21}K_{20}K_{10})[\mathrm{L}]^3 + (K_{31}K_{30}K_{20}K_{10} + K_{22}K_{21}K_{20}K_{10})[\mathrm{L}]^4 &+ \\ (K_{32}K_{31}K_{30}K_{20}K_{10})[\mathrm{L}]^5 + (K_{33}K_{32}K_{31}K_{30}K_{20}K_{10})[\mathrm{L}]^6\}\}&\end{aligned} \quad (1)$$

Equation 1 defines the heat of binding, ΔQ, as a function of the free ligand concentration, [L]. However, only the total concentration of ligand in the solution is known experimentally, and therefore, a second equation must be used to define the relationship between free bound, and total ligand concentrations. This relationship is given as follows:

$$[L]_{tot} = [L] + ([Ins]_{tot}/6)\{\{K_{10}[L] + 2(K_{20}K_{10} + K_{11}K_{10})[L]^2 + 3(K_{30}K_{20}K_{10} + K_{21}K_{20}K_{10})[L]^3 + 4(K_{31}K_{30}K_{20}K_{10} + K_{22}K_{21}K_{20}K_{10})[L]^4 + 5(K_{32}K_{31}K_{30}K_{20}K_{10})[L]^5 + 6(K_{33}K_{32}K_{31}K_{30}K_{20}K_{10})[L]^6\}/\{1 + K_{10}[L] + (K_{20}K_{10} + K_{11}K_{10})[L]^2 + (K_{30}K_{20}K_{10} + K_{21}K_{20}K_{10})[L]^3 + (K_{31}K_{30}K_{20}K_{10} + K_{22}K_{21}K_{20}K_{10})[L]^4 + (K_{32}K_{31}K_{30}K_{20}K_{10})[L]^5 + (K_{33}K_{32}K_{31}K_{30}K_{20}K_{10})[L]^6\}\} \quad (2)$$

The computation of the free ligand concentration makes this problem intractable by traditional least-squares or gradient descent-type curve-fitting techniques. The intractable nature of the problem is realized when it becomes necessary to solve an implicit variational function that describes the free ligand concentration for each experimental curve relating heat measured to total ligand concentration, while simultaneously attempting to adjust the various equilibrium constants and heats of binding for the best fit to the experimental data.

We have developed a novel, flexible algorithm to solve this irregular problem of optimizing multidimensional nonlinear parametrization with a variational implicit function solution. This is done through a 3-step iterative process. First, the free ligand concentration curves are estimated with a polynomial extrapolation that includes the dynamic adjustment of the optimization interval necessary to guarantee continuity in the resulting curves. Second, the enthalpies of binding are then estimated by a least-squares fit to the observed data. Finally, a global stochastic optimization program, which permits various constraints to be applied to the search space, is used for the nonlinear fit of equilibrium coefficients. The implicit function of free ligand concentration as a function of total ligand and total insulin concentrations must be solved to determine the nonlinear fit parameters describing equilibrium coefficients, heats of binding, and cooperativity. The algorithm permits dynamic adjustment of weighting functions so it can focus on resolving regions of poor convergence. The complete details of this algorithm and the mathematical methods are described in a forthcoming paper (Varshavsky et al., 1996).

The general equations for the 2Zn−insulin hexamer ligand binding system have nine equilibrium coefficients and nine heats of binding. A hierarchy of physical models can be defined that permits a progressive simplification of the underlying physical processes as the dimensionality is reduced. Either derivatives (plot of heat per injection vs total ligand concentration; corrected for heats of dilution and variable injection volumes), their corresponding integrals, or any weighted combination of the two can be used in the fitting algorithm. As an example, a possible total of 18 curves (three replicates at three different insulin concentrations producing nine derivatives and nine integrals) can be fit simultaneously in order to calculate the parameters that best reproduce the data. For arguments cited in Appendix A, we have chosen to present the optimized parameters resulting from the fits to integral data only.

An initial simplification assumes that the ligand binding cooperativity, if any, existing between the two insulin trimers

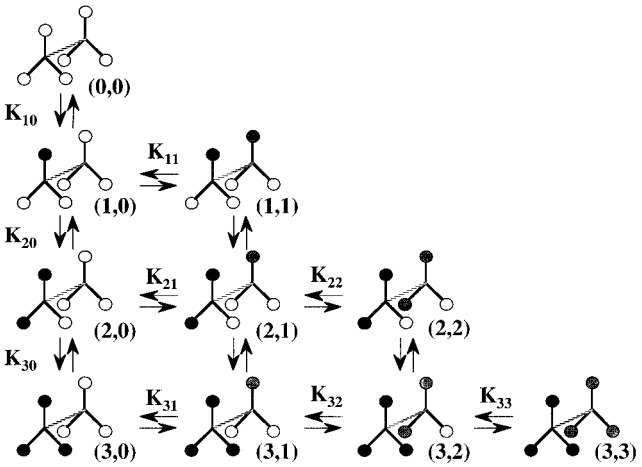

FIGURE 1: Schematic representation of the equilibrium binding processes in the 2Zn−insulin system. The insulin hexamer is depicted as a dimer of trimers. Unligated and ligated insulin molecules are represented by empty and filled circles, respectively. Inter-trimer binding occurs across any row and intra-trimer binding down any column. Subscripts on equilibrium constants, K, refer to the number of ligated insulin molecules within a specific trimer.

is independent of the cooperativity that might exist within each trimer. Technically, this is an assumption regarding linearity of superposition of binding between intra-trimer and inter-trimer sites. This assumption allows the resulting six degrees of freedom to be defined as one equilibrium coefficient, which defines the intrinsic equilibrium in the absence of cooperativity, and five correction factors that account for the various possible cooperativities (Table 1). A similar transformation is made involving the heats of binding. These intrinsic and cooperative factors are then used to map the original parameters in eqs 1 and 2 (Table 2).

As a consequence of this redefinition, a hierarchy of physical models can be defined (Table 3a). A much larger set of models can be constructed by the partial reduction of a model in either the equilibrium or enthalpy coefficient space (*e.g.*, reductions about Model 5 are shown in Table 3b). The approach is to begin fitting the data with the highest dimensional model (Model 6), then continue with models of lower dimensionality by exploring reductions made in the equilibrium and enthalpy coefficient space. With this approach, a model can be identified that provides as good a fit to the data as higher dimensional models; however, any further reductions in the parameter space significantly increase the error in the calculated ITC profile as identified by an increase in the sum of squares error. We then use the parameters from this minimum dimensional model to interpret our data.

Three constraints were used to help restrict the space in this high dimensionality problem. First, multiple data sets at three different macromolecule concentrations (three concentrations × three runs) were analyzed simultaneously. This permits the parameter estimation to account for the quality of the fit as the initial macromolecule concentration varies, *i.e.*, 0.6, 2.0, and 5.0 mM monomeric insulin, which impacts both eqs 1 and 2. Second, at saturation, the amount of phenolic ligand bound must equal the macromolecule concentration. Therefore, the amount of bound ligand divided by the amount of insulin monomer present in solution (hereafter referred to as γ) should approach unity as the reaction proceeds to saturation. Third, the flat asymptotic

Table 1: Parameters Used To Describe the Inter- and Intra-Trimer Cooperativity in the Binding of Phenolic Ligands to the 2Zn−Insulin Hexamer

parameters	description
equilibrium	
K_0	intrinsic equilibrium coefficient, in the absence of cooperativity
k_1	factor correcting K_0 for the presence of 1 intra-trimer ligand
k_2	factor correcting K_0 for the presence of 2 intra-trimer ligands
J_1	factor correcting K_0 for the presence of 1 inter-trimer ligand
J_2	factor correcting K_0 for the presence of 2 inter-trimer ligands
J_3	factor correcting K_0 for the presence of 3 inter-trimer ligands
enthalpic	
ΔH_0	intrinsic heat of binding, in the absence of cooperativity
∂g_1	incremental cooperative heat due to the presence of 1 intra-trimer ligand
∂g_2	incremental cooperative heat due to the presence of 2 intra-trimer ligands
∂h_1	incremental cooperative heat due to the presence of 1 inter-trimer ligand
∂h_2	incremental cooperative heat due to the presence of 2 inter-trimer ligands
∂h_3	incremental cooperative heat due to the presence of 3 inter-trimer ligands

Table 2: Mapping the Nine Individual Equilibrium Constants and Individual Heats of Binding from their Intrinsic and Correction Parameters

equilibrium constants[a]	heats of binding[b]
$K_{10} = K_0$	$\Delta H_{10} = \Delta H_0$
$K_{20} = K_0k_1$	$\Delta H_{20} = 2\Delta H_0 + \partial g_1$
$K_{30} = K_0k_2$	$\Delta H_{30} = 3\Delta H_0 + \partial g_1 + \partial g_2$
$K_{11} = K_0J_1$	$\Delta H_{11} = 2\Delta H_0 + \partial h_1$
$K_{21} = K_0J_2$	$\Delta H_{21} = 3\Delta H_0 + \partial g_1 + \partial h_2$
$K_{22} = K_0k_1J_2$	$\Delta H_{22} = 4\Delta H_0 + 2\partial g_1 + 2\partial h_2$
$K_{31} = K_0J_3$	$\Delta H_{31} = 4\Delta H_0 + \partial g_1 + \partial g_2 + \partial h_3$
$K_{32} = K_0k_1J_3$	$\Delta H_{32} = 5\Delta H_0 + 2\partial g_1 + \partial g_2 + 2\partial h_3$
$K_{33} = K_0k_2J_3$	$\Delta H_{33} = 6\Delta H_0 + 2\partial g_1 + 2\partial g_2 + 3\partial h_3$

[a] The equilibrium constants represent the equilibria between two ligated species as shown in Figure 1. [b] The heats of binding represent the overall enthalpy change for the given species starting from the unligated (T_6) hexamer.

behavior of the integrals at saturation was used to estimate the overall heat of the reaction at saturation ($\Delta H_{33} = \Delta H_{net}$). From eq 1, it can be seen that, in order to ensure proper limiting behavior in the integrals, the macromolecule concentration should be held constant. Constraints on the model by γ and ΔH_{net} were very powerful in restricting the equilibrium and enthalpy coefficient space, respectively.

The estimation of thermodynamic parameters in the model was done by minimizing (optimizing) a residual function fitting the model to experimental data with the three-step process described earlier. The integral data are used in determining the sensitivity of the estimates. Sensitivity analysis was used to help estimate the local curvature (constraint) of the residual function associated with each parameter in a given optimal minimum and, therefore, to provide an additional method for assessing the quality of the minimum obtained. The algorithm for completing this analysis is briefly described in Appendix B. The output is reported as the maximum change allowed in each parameter that maintains the calculated profile within the standard experimental error. Therefore, the greater the change allowed in the parameter by this constraint, the flatter the residual function is about this minimum along the dimension corresponding to the varied parameter. Large variability in a parameter indicates either (a) reductions in the dimensionality of the model used to fit the data can be achieved, (b) considerable error exists in the measurement of the experimental points, and/or (c) the minimized parameter space is not highly constrained. It must be emphasized that this analysis concerns only a limited region around the parameters constituting the optimal minimum and does not specify if this minimum is local or global or indicate the quality of the fit obtained. Furthermore, because only three experimental curves are obtained per insulin concentration, the confidence level in our estimated standard deviations is relatively low. However, the analysis partially establishes the minimum dimensional model needed to fit the data. The full details of the method and algorithm used in this analysis will be described in a forthcoming paper (personal communication).

Due to the complexity of the model, the method, and algorithm, the fitting was done on a Cray II Super computer to reduce computation time. The time taken to fit the experimental curves generated for each system ranged from 2 to 5 min.

RESULTS

Binding Profiles. A typical titration profile for phenol into In_6Zn_2 is shown in Figure 2. For the phenol/In_6Zn_2 system at 25 °C, the first several injections produce exothermic responses while the latter injections produce endothermic responses that diminish as saturation is approached. This profile will be referred to hereafter as an exothermic/endothermic response. This biphasic behavior in the binding isotherms is sensitive to temperature, *i.e.*, the second endothermic phase diminishing as the temperature is increased (Figure 3). Similar exothermic/endothermic response profiles were observed for all the phenol titrations in the temperature range of 14−32 °C, all the *m*-cresol titrations, and the resorcinol titration at 14 °C. Examples of derivatives and integrals resulting from the titration of phenol into In_6Zn_2 at 25 °C, as a function of insulin concentration, are shown in Figure 4. In the In_6Zn_2 system with a chloride concentration of 50 mM and temperature of 14 °C, *m*-cresol and phenol saturated the hexamer at approximately 25 mM, whereas resorcinol saturates at approximately 10 mM.

Modeling. A systematic evaluation of the family of models (Table 3), starting with Model 6, led to the identification of Model 5b as the lowest dimensional model necessary to give the highest quality fit to all of the data. The sum of squared errors calculated from both integral and derivative profiles increases with other models of dimensionality less than or equal to Model 5b (Figure 5, Table 3). Models of dimensionality greater than Model 5b could not improve the fit (Figure 5, Table 3). In addition, multiple minima were identified with these higher-dimensional, over-

Table 3

model	description	equilibrium parameters	enthalpy parameters
(a) Basic Set of Models That May Be Used To Analyze Binding Isotherms			
1 (2 parameters)	all sites equivalent and independent	K_0, ($k_1 = k_2 = 1$); $J_1 = J_2 = J_3 = 1$	ΔH_0, ($\partial g_1 = \partial g_2 = 0$); $\partial h_1 = \partial h_2 = \partial h_3 = 0$
2 (4 parameters)	isotropic inter-trimer cooperativity; (no intra-trimer cooperativity)	K_0, ($k_1 = k_2 = 1$); $J_1 = J_2 = J_3 = J$	ΔH_0, ($\partial g_1 = \partial g_2 = 0$); $\partial h_1 = \partial h_2 = \partial h_3 = \partial h$
3 (6 parameters)	anisotropic intra-trimer cooperativity; (no inter-trimer cooperativity)	K_0, k_1, k_2; $J_1 = J_2 = J_3 = 1$	ΔH_0, ∂g_1, ∂g_2; $\partial h_1 = \partial h_2 = \partial h_3 = 0$
4 (8 parameters)	anisotropic inter-trimer cooperativity; (no intra-trimer cooperativity)	K_0, ($k_1 = k_2 = 1$); J_1, J_2, J_3	ΔH_0, ($\partial g_1 = \partial g_2 = 0$); ∂h_1, ∂h_2, ∂h_3
5 (8 parameters)	anisotropic intra-trimer cooperativity; isotropic inter-trimer cooperativity	K_0, k_1, k_2; $J_1 = J_2 = J_3 = J$	ΔH_0, ∂g_1, ∂g_2; $\partial h_1 = \partial h_2 = \partial h_3 = \partial h$
6 (12 parameters)	anisotropic intra-trimer cooperativity; anisotropic inter-trimer cooperativity	K_0, k_1, k_2; J_1, J_2, J_3	ΔH_0, ∂g_1, ∂g_2; ∂h_1, ∂h_2, ∂h_3
(b) Subset of Models Constructed by the Systematic Reduction of Model 5			
5a (7 parameters)	limited anisotropic intra-trimer cooperativity; isotropic inter-trimer cooperativity	K_0, ($k_1 = k_2 = k$); $J_1 = J_2 = J_3 = J$	ΔH_0, ∂g_1, ∂g_2; $\partial h_1 = \partial h_2 = \partial h_3 = \partial h$
5b (6 parameters)	anisotropic intra-trimer cooperativity in enthalpy space only; isotropic inter-trimer cooperativity	K_0, ($k_1 = k_2 = 1$); $J_1 = J_2 = J_3 = J$	ΔH_0, ∂g_1, ∂g_2; $\partial h_1 = \partial h_2 = \partial h_3 = \partial h$
5c (6 parameters)	limited anisotropic intra-trimer cooperativity in equilibrium and enthalpy space; isotropic inter-trimer cooperativity	K_0, ($k_1 = k_2 = k$); $J_1 = J_2 = J_3 = J$	ΔH_0, ($\partial g_1 = \partial g_2 = \partial g$); $\partial h_1 = \partial h_2 = \partial h_3 = \partial h$
5d (5 parameters)	limited anisotropic intra-trimer cooperativity in enthalpy space only; isotropic inter-trimer cooperativity	K_0, ($k_1 = k_2 = 1$); $J_1 = J_2 = J_3 = J$	ΔH_0, ($\partial g_1 = \partial g_2 = \partial g$); $\partial h_1 = \partial h_2 = \partial h_3 = \partial h$
5e (6 parameters)	anisotropic intra-trimer cooperativity in equilibrium space only; isotropic inter-trimer cooperativity	K_0, k_1, k_2; $J_1 = J_2 = J_3 = J$	ΔH_0, ($\partial g_1 = \partial g_2 = 0$); $\partial h_1 = \partial h_2 = \partial h_3 = \partial h$

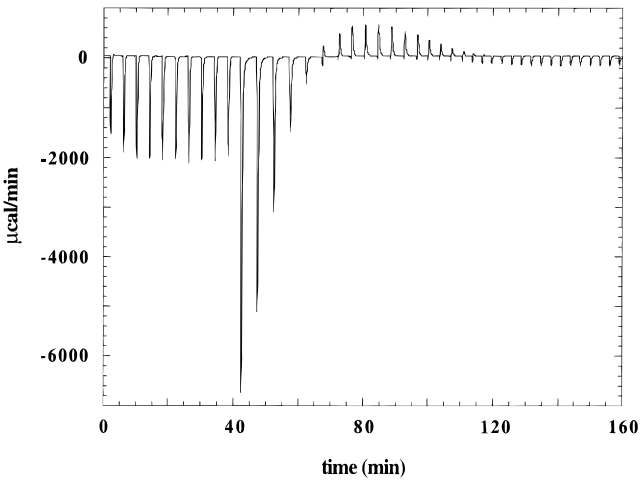

FIGURE 2: Isothermal titration calorimetry profile resulting from the titration of 200 mM phenol into 0.83 mM In_6Zn_2 at 25 °C. Solutions were 25 mM in Tris, pH = 8.0, and 50 mM in Cl^-. Variable injection volumes (1−15 μL) were used to increase the sampling density in the early part of the titration. Integration of the area under these peaks plotted versus total ligand concentration yields the derivative curves, an example of which is shown in Figure 4B.

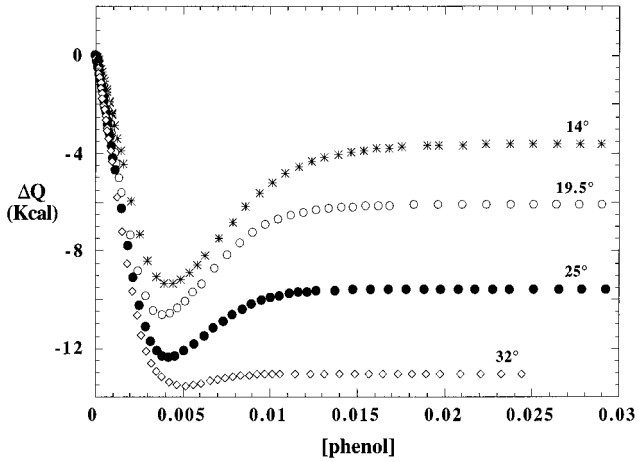

FIGURE 3: Dependence of temperature on the integral curves resulting from the titration of phenol into 0.83 mM In_6Zn_2. As the temperature increases, the second transition ($T_3R_3 \leftrightarrow R_6$) becomes more exothermic until it disappears completely at temperatures near 32 °C. Solutions were 25 mM in Tris at pH = 8.0 and 50 mM in Cl^-.

parametrized models of which the best minima contained parameters of unrealistic magnitude.

Examples of typical fits to integral curves and their resulting derivatives are shown for the titration of phenol into In_6Zn_2 (Figures 4, panels A and B, respectively). The fits obtained fall within the standard experimental error (largest for the first few injections) for nearly every point on the curve. Differences between experimental and calculated derivatives show noticeable errors in the first few points; however, similar residual profiles were observed between the experimentally-determined derivative and the experimental average profile associated with the same system.

A typical plot of the saturation constraint (γ; ratio of bound ligand to insulin monomer) versus ligand concentration is also shown for the phenol/In_6Zn_2 system (Figure 4C). The assumption that all (or nearly all) binding sites within the insulin hexamer are occupied at saturation was built into the fitting routine for all models.

Homotropic Intra-Trimer Cooperativity. The optimized parameters for Model 5b, which was used to fit the data for the various systems investigated, are listed in Table 4. The modeling results suggest that intra-trimer cooperativity is present in every system studied but to a limited degree, appearing only in the enthalpic coefficient space. No intra-trimer cooperativity in the equilibrium coefficient space ($k_1 = k_2 = 1$; Table 3b, Model 5b) was required to fit the data.

The intra-trimer cooperativity in the enthalpy coefficient space was observed to be complex and a function of the ligand and temperature. The incremental cooperative heat due to the presence of one intra-trimer ligand (∂g_1) is either endothermic or exothermic and is typically small in mag-

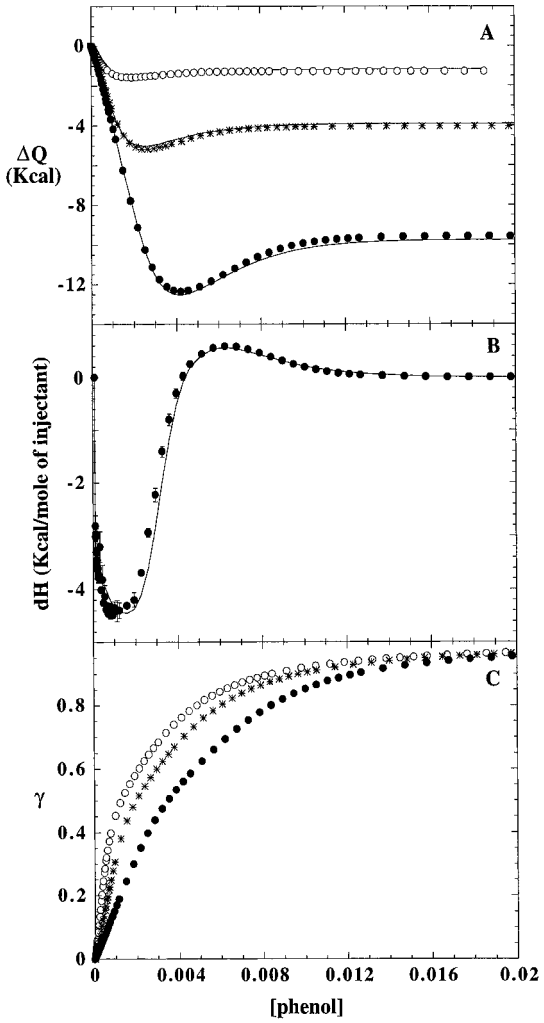

FIGURE 4: Representative example of the quality of fit obtained for the integrals (A) and derivatives (B) in the phenol/In_6Zn_2 system using Model 5b. Shown are the average of three experiments with the standard error. The solid line is the fit. For clarity, only one derivative is shown up to the point of saturation. The calculated ratio of bound ligand to insulin monomer (γ) versus phenol concentration is displayed in (C) for all three insulin concentrations studied. Inflections in γ represent the ligand concentration at which the $T_6 \leftrightarrow T_3R_3 \leftrightarrow R_6$ transitions occur. Solutions were 25 mM in Tris at pH = 8.0 and 50 mM in Cl^-. Initial hexamer concentrations were 0.83 (●), 0.33 (∗), and 0.10 mM (○).

nitude (<2 kcal/mol), with the exception of the resorcinol/In_6Zn_2 system at 14 °C, where it makes a relatively large contribution to the overall enthalpy (Table 4). The incremental cooperative heat due to the presence of two intratrimer ligands (∂g_2) was exothermic for all the systems studied and increases in contribution to the overall enthalpy as the temperature is increased (Table 4).

Homotropic Inter-Trimer Cooperativity. The modeling results (Table 4) also show that strong isotropic inter-trimer cooperativity in the equilibrium coefficient space exists in all the systems studied, *i.e.*, $J < 1$. The values obtained for the inter-trimer cooperativity parameter (J) were observed between 0.035 and 0.29, with all but two cases (phenol at 32 °C and *m*-cresol at 25 °C) being less than 0.1. As is shown in the mathematical mapping of the intrinsic and cooperative equilibrium parameters to the individual equilibrium constants (Table 2), values less than 1 for J indicate negative inter-trimer cooperativity. The temperature data for phenol and *m*-cresol binding suggest that the magnitude of this negative cooperative factor diminishes as temperature increases. The incremental cooperative heat due to the presence of an inter-trimer ligand(s) (∂h) was strongly endothermic in all cases.

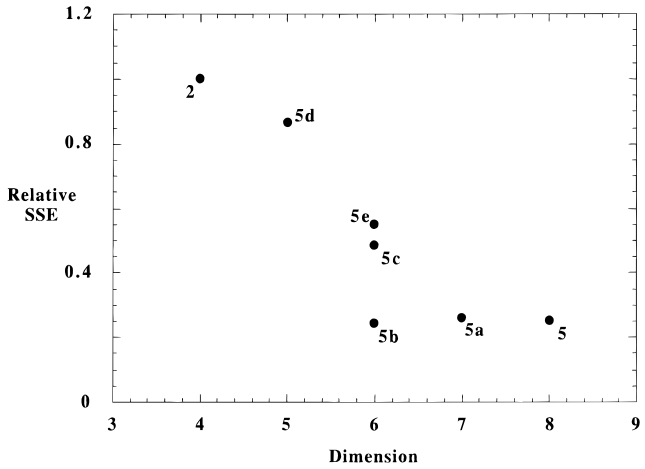

FIGURE 5: Relative sum of squared error between experimental average and calculated integrals and derivatives for specified models obtained from fitting ITC profiles generated by phenol binding to 2Zn−insulin at 14 °C. Data are not shown for all the lower dimensional models given in Table 3a because of significantly increased errors.

Intrinsic and Overall Thermodynamic Binding Parameters. The modeling data suggest that the rank order of intrinsic binding affinity, in the absence of cooperativity, of these ligands for insulin hexamer is resorcinol (8780 L/mol) > phenol (5040 L/mol) > *m*-cresol (3370 L/mol) (Table 4). The intrinsic heat of binding was exothermic for all the systems studied and ranged from −1.2 to −3.8 kcal/mol, with resorcinol being the most exothermic for the ligands studied at 14 °C (Table 4).

The overall thermodynamic parameters (ΔH_{net}, ΔG_{net}, and ΔS_{net}) for the various systems studied are shown in Table 4. The free energy of binding (ΔG_{net}) establishes the same order of ligand binding affinity stated above and also shows an increase in magnitude with temperature. The enthalpy (ΔH_{net}) is highly exothermic for resorcinol, decreases significantly for phenol, and becomes endothermic for *m*-cresol binding at 14 °C. As would be expected, increasing temperature increases the exothermic nature of ΔH_{net}. The entropy (ΔS_{net}) established a decreasing trend with temperature. The overall ligand binding process ($T_6 \leftrightarrow R_6$) is an enthalpy/entropy-driven process, *i.e.*, negative ΔH_{net}, positive ΔS_{net}, for each of the systems studied, with the exception of *m*-cresol binding at 14 °C, which is entropically driven, *i.e.*, positive ΔH_{net}, positive ΔS_{net} (Gilli et al., 1994).

Temperature Effects. Isothermal titration calorimetry profiles have been obtained at 14, 19.5, 25, and 32 °C for phenol; 14, 25 (data not shown), and 40 °C (data not shown) for resorcinol;[3] and 14 and 25 °C for *m*-cresol binding to In_6Zn_2. The binding profiles at a single insulin concentration as a function of temperature are shown for the phenol/In_6Zn_2

[3] The data generated by resorcinol binding to 2Zn−insulin at 25 °C was not fit with our models due to a thermally silent second transition. The apparent loss of the second transition is due to a balancing of endothermic and exothermic contributions in the binding of preservative to the second trimer; however, spectroscopic studies (data not shown) indicate that the second transition still occurs. Consequently, the calorimetric data for resorcinol binding to 2Zn−insulin at 25 °C could

Table 4: Equilibrium and Enthalpic Parameters That Reproduce the Derivatives and Integrals Resulting from the Titration of Phenol, *m*-Cresol, and Resorcinol into In_6Zn_2 at the Given Temperatures Using Model 5b

	resorcinol (14 °C)	*m*-cresol (14 °C)	*m*-cresol (25 °C)	phenol (14 °C)	phenol (19.5 °C)	phenol (25 °C)	phenol (32 °C)
K_0 (L/mol)	8780 (±520)[a]	3370 (±340)	2230 (±160)	5040(±420)	4360 (±380)	3420(±190)	2040 (±370)
k_1	1.0	1.0	1.0	1.0	1.0	1.0	1.0
k_2	1.0	1.0	1.0	1.0	1.0	1.0	1.0
J	0.07 (±0.00)	0.08 (±0.01)	0.15 (±0.01)	0.04 (±0.00)	0.07 (±0.01)	0.09 (±0.01)	0.30 (±0.05)
ΔH_0 (kcal/mol)	−2.60 (±0.00)	−1.20 (±0.16)	−2.90 (±0.34)	−1.21 (±0.03)	−3.05 (±0.43)	−3.58 (±0.34)	−3.79 (±0.38)
∂g_1 (kcal/mol)	−6.76 (±0.10)	0.71 (±0.78)	1.20 (±2.21)	−1.84 (±0.28)	1.60 (±1.81)	0.83 (±1.68)	−2.83 (±0.53)
∂g_2 (kcal/mol)	−0.28 (±0.15)	−6.04 (±0.97)	−7.16 (±0.33)	−4.54 (±0.08)	−5.91 (±0.59)	−5.62 (±0.41)	−6.45 (±1.48)
∂h (kcal/mol)	4.90 (±0.06)	6.67 (±0.55)	8.01 (±0.31)	5.66 (±0.18)	7.03 (±0.09)	7.28 (±0.08)	9.94 (±1.19)
ΔH_{net} (kcal/mol)	−14.8	2.14	−5.29	−3.03	−5.83	−9.22	−11.5
ΔG_{net} (kcal/mol)	−26.6	−23.4	−24.1	−23.4	−24.5	−24.7	−25.5
ΔS_{net} (cal/(mol·deg))	41.0	89.1	63.0	71.1	63.8	51.8	46.0
K_{net}	1.76×10^{20}	6.45×10^{17}	4.63×10^{17}	6.45×10^{17}	1.98×10^{18}	1.27×10^{18}	1.83×10^{18}

[a] Results of the sensitivity analysis (see Materials and Methods) are given in parentheses for each adjustable parameter. These values represent the variability allowed in the parameter before the calculated fit falls outside the standard experimental error for the curves.

system in Figure 3. In all cases, an increase in temperature decreases the relative magnitude of the second endothermic phase. Upon further increase in temperature, the second transition becomes discernible as an exothermic contribution to the profile. This was verified for only resorcinol, because the transition temperature is considerably lower than the phenol system in which the transition temperature exceeded 40 °C (data not shown).

The temperature-dependent data for the phenol/In_6Zn_2 system are well fit by Model 5b, and the results are shown in Table 4. The intrinsic equilibrium constant (K_0) decreases with increasing temperature; however, the inter-trimer cooperativity in the equilibrium coefficient space (J) increases with temperature, but remains as a negative cooperativity factor. The intrinsic heat of binding of phenol, in the absence of cooperativity, becomes more exothermic with increasing temperature. An enthalpy/entropy compensation plot (not shown) indicates that the binding is a spontaneous enthalpy/entropy-driven process. The binding is more entropically-driven at low temperature and enthalpically-driven at high temperature. A plot of overall enthalpy (ΔH_{net}) as a function of temperature is linear and yields a heat capacity change (ΔC_p) of −460 cal/(mol·deg).

pH and Buffer Effects. The effect of pH and Tris buffer on the titration profiles is minimal. For the phenol/In_6Zn_2 system, the integrated ITC profiles compared between pH = 8.0 and 25 mM Tris and pH = 7.5 with a solution containing no buffer are the same within experimental error (data not shown). Also, the heat generated by the ionization of Tris buffer is negligible (nanocalories) because the change in pH associated with these reactions is approximately 0.1 unit when carried out in the absence of buffer.

DISCUSSION

It has been proposed that phenol-like molecules could be utilized to stabilize the insulin hexamer, consequently increasing the shelf-life stability of insulin preparations (Derewenda et al., 1989) and possibly extending the time action of the insulin formulation (Smith & Ciszak, 1994). Therefore, it is essential to quantitate the intrinsic binding constant of these phenolic ligands to assess the inherent hexamer stabilization effect that different ligands possess. This study provides a methodology for analytically determining a quantitative solution to the complex binding patterns of phenolic ligands to insulin hexamers. Previous microcalorimetric studies were performed to investigate phenolic ligand binding to 2Zn−insulin hexamers (McGraw & Lindenbaum, 1990); however, the concentrations of ligand used in these studies began at, or near, saturating levels, and therefore the data only reflect binding to the second trimer. Consequently, comparisons between results are limited.

As a reminder, the information available from ITC is limited to the resultant heat produced by an injection. In this study, these heats are the sum of several processes (ligand binding, conformational change from extended to α-helical, and the release of water molecules) that cannot be individually isolated. If the binding of a ligand produces exothermic and endothermic heats of equal magnitude, a monophasic binding profile may result and valuable information may be lost. Consequently, care must be exercised in performing and interpreting not only the data, but the parameters calculated by the model as well. Thus, the experimental conditions (especially temperature) as well as a sensitivity analysis of the minima must be considered in the fitting routine.

Modeling. The modeling results from this study demonstrate that the free energy of preservative binding to In_6Zn_2 is controlled by negative, isotropic inter-trimer cooperativity and that intra-trimer cooperativity is less important, but does contribute to the binding (Model 5b). These quantitative results support the qualitative statements in previous studies that positive intra-trimer cooperative binding of phenolic ligands is not significant in the In_6Zn_2 system (Krüger et al., 1990) and that strong, negative inter-trimer cooperativity modulates the binding (Krüger et al., 1990; Roy et al., 1989).

It is notable that, in order to fit much of our data, the dimensionality of Model 5 could be reduced in the equilibrium coefficient space, but not in the enthalpy coefficient space. This is not surprising since the enthalpy coefficient space is directly measured; however, the equilibrium coefficient space is influenced by both enthalpic and entropic contributions ($\Delta G = \Delta H - T\Delta S$). Consequently, entropy/enthalpy compensation can occur that can affect the parametrization of the equilibrium coefficient space.

be fit with Model 2 which is an underestimate of the actual cooperativity that exists in the system. Data generated at 14 °C and preliminary data at 40 °C (data not shown) for resorcinol have a detectable second transition that is most adequately fit with Model 5b.

As previously stated, the heats generated in these reactions are attributable to several factors including conformational changes, ligand binding, and hydration perturbation. Under certain experimental conditions with specific ligands, *e.g.*, resorcinol binding to In_6Zn_2 at ~25 °C or phenol binding to In_6Zn_2 at ~40 °C, the second binding phase appears to be absent, suggesting that negative cooperativity no longer exists. For both of these systems at 14 °C, the profiles have the usual exothermic/endothermic phases. However, increasing the temperature reduces the magnitude of the endothermic/second phase until a transition temperature is reached, after which the second transition reappears as an exothermic profile. Spectroscopic evidence (not shown), CD data (Jacoby et al., 1993), and these temperature-dependent studies indicate that the second transition is thermally silent at the transition temperature; *i.e.*, exothermic and endothermic processes that compose the heat signal are of equal magnitude. Thus, these calorimetric constraints need to be considered in the analysis of the binding isotherms.

Sensitivity analysis was performed on the optimal minima to estimate the local curvature (or constraint) around each of the fitting parameters. For each system, the variability allowed in the inter-trimer cooperativity factor (J) was quite small (Table 4). This small variability indicates that the model is very sensitive to small inflections and changes in slope occurring in the derivative and integral profiles and that the relative magnitude of the second transition does not indicate the degree of inter-trimer cooperativity present in the system.

From the sensitivity analysis (Table 4), the least constrained minima were observed with *m*-cresol (14 °C) and phenol (32 °C) data sets. The increased variability observed in the parameters of these two data sets most likely arises from larger experimental errors. The larger experimental error in the ITC profiles generated by *m*-cresol binding at 14 °C can be, in part, attributed to the low solubility of *m*-cresol, which made it difficult to saturate the insulin hexamer at the highest insulin concentration.

The identification of poorly constrained parameters with the sensitivity analysis also suggests that a possible reduction in the parametric space of the model can be achieved. However, with the exception of resorcinol binding at 14 °C and phenol binding at 32 °C, it was found that any reduction from Model 5b significantly increased the sum of squares error in the fit. With resorcinol at 14 °C and phenol at 32 °C, the quality in the fit to the data with Model 5d was nearly as good as with Model 5b. However, this lower dimensional model was eliminated based on its inability to satisfy the ΔH_{net} constraint in the phenol system at 32 °C, and on a small but noticeable deviation in the quality of the calculated derivative from the experimental profile in the region of the second transition for the resorcinol system at 14 °C (data not shown).

A lack of constraint in the parameter ∂g_1 was observed in all cases. However, using the hierarchical model protocol to reduce the parametrization and assessing the fits based on the sum of squares error criteria as described above, it appears that the extra degree of freedom offered by ∂g_1 is necessary to obtain the best possible fit to the data. The increased variability in ∂g_1 may be due to the reduced magnitude of the parameter.

Saturation Constraint, γ. As Figure 4C illustrates, γ approaches unity at the end of the titration (as it is constrained to within a tolerance of 5% of unity at saturation), and the inflections in the curves indicate the ligand concentration at which the $T_6 \leftrightarrow T_3R_3 \leftrightarrow R_6$ transitions occur. The steep, initial slope indicates that a greater percentage of the total ligand present in solution is bound to the hexamer at this stage of the titration. A representative example of the percent ligand bound as a function of total ligand concentration is shown in the inset of Figure 6 for phenol binding at 25 °C. As much as 80% of the total ligand present in solution is bound to the hexamer at the early stages of the titration. This clearly shows the inaccuracy of the assumption of total and free ligand concentrations being equal in previous modeling procedures.

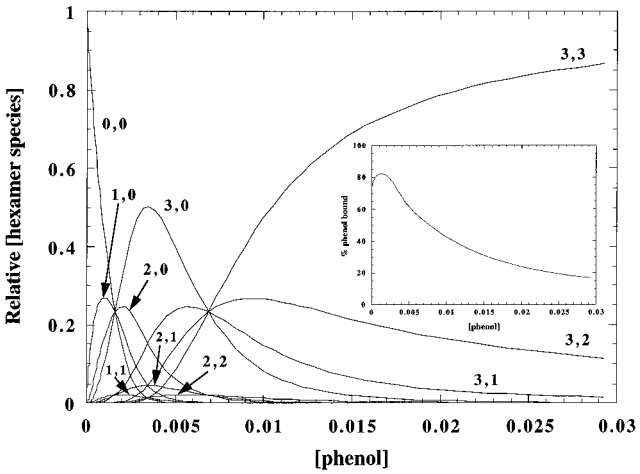

FIGURE 6: Concentrations of insulin hexamer species as a function of total phenol present in solution. The sum of all 10 species is 1.0 at any given phenol concentration. The inset shows the percent phenol bound to the 2Zn−insulin hexamer as a function of total phenol concentration. These curves were calculated using the equilibrium constants obtained from modeling the data generated by the binding of phenol to the 2Zn−insulin hexamer at 25 °C (5.0 mM insulin, 50 mM Cl^-, 25 mM Tris, pH = 8.0).

Intrinsic Binding Constant. In the In_6Zn_2 system at 25 °C, the intrinsic binding constant for phenol ($K_0 = 3420$) is comparable to, but slightly greater than, that of *m*-cresol ($K_0 = 2230$) (Table 4). This is consistent with results reported by several other workers who have used spectroscopic methods to measure the binding of phenolic ligands (Choi et al., 1993; Jacoby et al., 1993). The modeling of CD spectroscopic binding curves by Jacoby and co-workers produced a similar trend for the intrinsic equilibrium constant for phenol and *m*-cresol (Jacoby et al., 1993). The apparent order of ligand ability to bind to the insulin hexamer is established as resorcinol ≫ phenol ≥ *m*-cresol. This order has also been observed by several other groups using various spectroscopic techniques (Choi et al., 1993; Dodson et al., 1993; Jacoby et al., 1993) and supports the results of this quantitative treatment.

The tighter binding of resorcinol is due to the participation of the second hydroxyl group in the formation of additional hydrogen bonds to the hexamer. X-ray crystallographic data reveal that resorcinol binds to the insulin hexamer in a similar fashion as phenol, with hydrogen bonds to the ^{A6}Cys and ^{A11}Cys; however, a water molecule acts as a hydrogen-bonding bridge between the second hydroxyl group on resorcinol and the carbonyl group of the ^{A11}Cys of the same insulin monomer (G. D. Smith, personal communication). In addition, a somewhat weaker hydrogen bond between the second hydroxyl group and the imidazole nitrogen of a ^{B5}His

on an adjacent monomer has also been identified. This increased affinity (relative to phenol and *m*-cresol) of a ligand to bind to the insulin hexamer through additional hydrogen bonds is consistent with data from other bifunctional ligands, including 3-aminophenol (unpublished results) and 7-hydroxyindole (Dodson et al., 1993).

The weaker K_0 for *m*-cresol binding to the insulin hexamer may be due to steric hindrance by the *m*-methyl group. This steric hindrance could affect orientation, alter the hydrogen bonding pattern or length, and/or cause less favorable contacts relative to phenol. These factors would account for the reduced heat released and lower equilibrium constant (intrinsic and overall) for *m*-cresol.

Increasing the temperature results in an expected decrease in K_0, because the intrinsic heat of binding (ΔH_0) is exothermic. The exothermic intrinsic heat of binding (Table 4), observed under all conditions with all ligand/hexamer systems, is believed to be due to the formation of hydrogen bonds and van der Waals contacts between the ligand and the insulin hexamer, as well as the extension of the B-chain α-helix from B9 through B19 to B1 through B19 upon phenolic ligand binding (Derewenda et al., 1989; Scholtz et al., 1991). The greater value for resorcinol binding at 14 °C relative to phenol and *m*-cresol is due, in part, to the formation of additional hydrogen bonds. The increasing exothermic response (ΔH_0), with increasing temperature, is consistent with the perturbation of the hydration layer of the insulin hexamer, one of numerous sources for heat loss/generation. Crystallographic evidence from T_3R_3 structures suggests that the amount of ordered water in the vicinity of the hydrophobic ligand binding site is greater in the T-state than in the R-state (G. D. Smith, personal communication). Therefore, low temperatures should order the hydration shell around hydrophobic groups, which will lead to an increased endothermic response upon ligand binding, thereby reducing the exothermic nature of ΔH_0. This is what is observed and reported in Table 4.

Homotropic Intra-Trimer Cooperativity. Zinc insulin hexamers display only a limited degree of intra-trimer cooperativity in binding phenolic ligands (Table 4). This may, in part, be due to the ligand field stabilization energy for zinc complexes. With zinc, there is no observed difference in the LFSE between octahedral and tetrahedral coordination in aquo complexes (Huheey, 1983). The lack of an initial energy barrier with respect to octahedral → tetrahedral conversion makes it difficult to observe intra-trimer cooperativity in the equilibrium coefficient space with the In_6Zn_2 systems, and consequently Model 5b is sufficient to fit the data reported for these 2Zn−insulin and preservative systems.

As previously mentioned, anisotropic intra-trimer cooperativity in the enthalpy coefficient space was essential for modeling the binding isotherms under all conditions investigated. Analysis of the results indicates this enthalpic intra-trimer cooperativity is complicated and dependent on the ligand and temperature. The incremental contribution to the enthalpy by a single ligand binding within a trimer is not obvious and requires further investigation. However, different ligands influence the enthalpic contributions by altering the hydration state, the hydrogen bonding pattern to the ligand, and the coordination geometry to varying degrees, as demonstrated by resorcinol, phenol, and *m*-cresol binding at 14 °C (Table 4). The effects of the *m*-hydroxyl group of resorcinol on the incremental contribution to the enthalpy by a single ligand binding within a trimer is highlighted by the increased magnitude of ∂g_1 relative to phenol and *m*-cresol ligand systems (Table 4). The incremental cooperative heat contributed by the presence of two ligands was exothermic in every case. These data suggest that the water associated with the third ligand site on a trimer is sufficiently perturbed by the binding of the two previous ligands that the endothermic/entropic gains of water release are not achieved by the binding of the third ligand.

The increased variability in ∂g_1 makes it difficult to evaluate the temperature effects on intra-trimer cooperativity with a single phenolic ligand bound. However, the exothermic cooperative factor ∂g_2, the cooperative correction factor in the presence of the phenolic ligands, was essentially constant at temperatures greater than or equal to 19.5 °C. The combination of higher temperatures and degree of saturation within the trimer clearly disrupt the hydration layer at the third site, reducing the endothermic contributions and increasing the exothermic contributions to binding.

Homotropic Inter-Trimer Cooperativity and the Binding Pathway. It has been observed that negative inter-trimer cooperativity is present in the binding of phenolic ligands to zinc insulin hexamers (Krüger et al., 1990; Roy et al., 1989). This suggests that ligand binding occurs through the sequential conversion of trimers. Although our models have no predisposition for a specific binding pathway, the resulting inter-trimer parameter (J) suggests this trimer-wise pathway is preferred (Table 4). This is illustrated in Figure 6 for phenol binding at 25 °C, where it is clear species (1,1), (2,1), and (2,2) are never present in significant amounts. The population of species (3,2) is approximately 12% of the relative hexameric concentration, which corresponds to approximately 2% of the sites left unligated.[4]

The isotropic nature of the inter-trimer cooperativity is a reasonable expectation considering that intermediate ligated species such as the (1,1), (2,1), and (2,2) hexamers exist at relatively low concentrations compared to the other species and therefore make considerably smaller contributions to the ITC profile (Figures 1 and 6). The isotropic nature of J also suggests that negative cooperativity is a function of the trimer and not of the individual insulin monomers.

The overall equilibrium constant, K_{net}, for phenol remains constant at temperatures greater than or equal to 19.5 °C (Table 4) despite the increasing exothermic response in ΔH_{net}. The increasing trend in J with temperature compensates for the changes observed in K_0, yielding a constant K_{net}. These results suggest that negative inter-trimer cooperativity is

[4] Figure 6 also gives the appearance that the system is not yet saturated, which contrasts the ITC profiles. This is an artifact of the fit. The γ constraint is not rigorously forced to unity at saturation, but is allowed a variance of 5% (*i.e.*, γ can range between 0.95 and 1.0 at saturation) because of the obvious experimental errors in measuring both insulin and ligand concentrations. Typical limiting values of γ fall between 0.97 and 1.0 (Figure 4C). Because γ is a function of the free ligand concentration, it directly influences the calculation of the curves in Figure 6. However, because less than 3% of the calculated sites are unsaturated, the tolerance in γ will only affect the populations of (3,1) and (3,2) species, significantly. Because of this range allowed in γ, only populations in excess of 3% of total sites at saturation can be considered significant.

much less of a factor at higher temperatures.[5] A more dynamic hexamer may be better able to accommodate the six carboxylate groups of the ^{B13}Glu residues that are in close proximity in a central cavity of the R-state insulin hexamer, which has been identified, in part, to be responsible for the negative inter-trimer cooperativity observed in insulin (Bloom et al., 1995; Jacoby et al., 1993; Wollmer et al., 1989).

The incremental cooperative heat due to inter-trimer bound ligands was unfavorably endothermic in all cases. These results are also consistent with the hypothesis that charge repulsion created by the carboxylate groups of the ^{B13}Glu residues is partly responsible for the observed negative inter-trimer cooperativity.

Overall Thermodynamics. Thermodynamic values calculated from the equilibrium and enthalpic parameters show that ligand binding to the insulin hexamer is favorable and spontaneous (Table 4). At 25 °C, phenol and *m*-cresol binding to In_6Zn_2 are enthalpy/entropy-driven processes (Gilli et al., 1994). The large positive changes in entropy are consistent with water release upon the binding of phenolic ligands to the hexamer. This is further substantiated by the overall entropy change (ΔS_{net}) observed as a function of temperature, which shows increased entropic contributions to binding at lower temperature (Table 4) that have been attributed to an increased ordering of water around hydrophobic groups in the protein (Ross & Subramanian, 1981). At higher temperatures, the binding process becomes more enthalpically-driven due to the loss of this ordered water around the hydrophobic residues in the binding site reducing the associated endothermic/entropic contributions upon binding.

The temperature dependence of phenol binding to the zinc insulin hexamer indicates a strong dependence for ΔH_{net} (Table 4). The heat capacity (ΔC_p) in this limited range is −460 cal/(mol·deg). This large negative value is also consistent with the "hydrophobic effect" or the disruption of ordered water from the protein surface (Ross & Subramanian, 1981; Sturtevant, 1977).

In summary, using a hierarchical series of models in fitting ITC data, we can make direct quantitative comparisons between structurally similar yet chemically different systems. The broad dynamics of this family of models are demonstrated in their ability to extract subtle energetic behavior from the data, such as cooperativity. Finally, it should also be relatively simple to use the core of the model in fitting other complex binding systems by adjusting eqs 1 and 2 to correct for differences in equilibria, *i.e.*, number of binding ligands, structural representation, etc.

ACKNOWLEDGMENT

The authors wish to thank Mark L. Brader and Maryanne Wagner for critical review of the manuscript.

[5] The temperature dependence of the second transition could underestimate the point of saturation and affect the accurate determination of the various binding parameters. However, it should be noted that the intrinsic binding parameters, K_0 and ΔH_0, are primarily determined from the first few points in the integral profile, meaning that these two parameters are minimally influenced by the terminal values. Therefore, the two most significant fitting parameters, K_0 and ΔH_0 (Table 2), are minimally affected as the magnitude of the second transition decreases. Despite the diminished second transition at elevated temperatures, sufficient information is available in the binding profile to estimate these parameters in the temperature ranges studied. Additional work is planned to develop models that include the temperature dependence on the equilibrium constants.

APPENDIX A

Annihilation of Random Error along the Integral Path. We have examined the exclusive use of either integrals or derivatives in our fitting and have found that the minima obtained are not significantly different (average difference in parameters was ≤10%). It should be noted that the addition of the short curves (low preservative concentration range) to the data was critical for the convergence of the observed integral and derivative minima (especially, K_0 and ΔH_0). We have decided to present the minima obtained from fitting integrals exclusively due to their inherently decreasing error along the experimental curve and strong limiting behavior at saturation on the fit. The proof is quite simple.

It is well-known that for any two independent random variables x and y, the standard deviation can be expressed as follows (van der Varden, 1960):

$$\sigma_{x+y}{}^2 = \sigma_x{}^2 + \sigma_y{}^2 \tag{A1}$$

This formula can be easily extended to any number of random variables with the condition that any pair of variables are mutually independent. Also, for a given function $f(x)$, its integral over n points with an equal interval between points can be expressed as

$$\int_a^b f(x)\,\mathrm{d}x \approx \sum_{i=0}^{n-1} f(x_i)\,\Delta x_i \tag{A2}$$

where $f(x_i)$ corresponds to a value of the function, f, at any fixed point inside the interval Δx_i (*i.e.*, in the middle of Δx_i), and

$$\Delta x_i = \frac{b-a}{n}$$

where $b - a$ represents the interval of integration. The dispersion of eq A2 yields

$$D\left[\int_a^b f(x)\,\mathrm{d}x\right] = \frac{(b-a)^2}{n^2} \sum_{i=0}^{n-1} D[f(x_i)] \tag{A3}$$

where $D(x) = \sigma_x{}^2$ for any random variable x. With the assumption of *independent* and *equal* magnitude errors in every experimental point in the derivative curve, one obtains the following result:

$$\sigma_{tot} = \frac{L_{max}}{n^{1/2}}\,\sigma_0 \tag{A4}$$

where σ_{tot} is the final point-wise error at the end of the integrated curve, σ_0 is the error associated with each point in the derivative curve, and L_{max} is the final ligand concentration. Examination of eq A4 reveals that the point-wise error in the integrated profile will decrease linearly with the square root of the number of experimental points (*i.e.*, number of injections). The assumption of independent (random) errors in each point in the derivative profile is a reasonable expectation for a typical analytical measurement such as ours. The assumption that the errors are equal for each point in the derivative curve is certainly not true. However, due to increasing injection volumes as the titration progresses, the error in the heat measured per injection decreases cor-

respondingly. Therefore, σ_{tot} would only decrease further if the actual errors were used in the analysis. Due to the fact that experimental errors are largest in the critical first few injections for the derivative profile, and the above result, we will report values of the optimized parameters resulting from fitting integral data exclusively.

As an aside, it should be noted that the integral curve clearly has a strong correlation among the points toward the end of the curve; *i.e.*, the variations in eq A4 are not independent among the points in the integral curve (the transformation matrix is triangular with unit value nonzero entries). However, we do *not* use the assumption of independence of distribution in our analyses. The optimization algorithm attempts the best projection from the parameter space to minimize the residual function but does not invoke any assumed error distribution (use of the least-squares fitting procedure is to be regarded here as a projection operator). The sensitivity analysis is a separate procedure operationally constructed to test explicitly the point-wise variability of the fit by variation in each parameter.

APPENDIX B

Algorithm for Estimation of the Sensitivity to the Optimal Coordinates. Using the optimal minimum in $\bar{X}^* = (X^*(1), X^*(2), ..., X^*(K))$ K-dimensional space obtained by the application of the method described in the text, we generated N random points (using pseudo-random numbers with uniform distribution from the IMSL library for the Cray 2) with the following constraints:

$$\bar{\alpha}(i) < \bar{X}^*(X^*(i)) < \bar{\beta}(i), \qquad i = 1, ..., K \tag{B1}$$

$$0 < F(x) - F(\bar{X}^*) < \epsilon(\bar{X}^*) \tag{B2}$$

$$\text{abs}[\Delta H_{net}(\bar{X}^*) - \Delta H_{net}(x)] < \delta\, \Delta H_{net}(\bar{X}^*) \tag{B3}$$

where $\bar{\alpha}_i = \bar{X}^*(i)\,(1 - \Delta)$ and $\bar{\beta}_i = \bar{X}^*(i)\,(1 + \Delta)$. The set of inequalities B1 represent a "local box" around the equilibrium coefficients of the optimal point $\bar{X}^*$. The vectors, $\bar{\alpha}$ and $\bar{\beta}$, represent the lower and upper bounds of the box, respectively. The size of the sensitivity box is directly related to the variability in the equilibrium parameters obtained from the results of multiple fitting trials. The minima from multiple fitting trials were clustered using Statistical Discovery Software, JMP v3.1 (SAS, 1995), and the standard deviations around the clustered means computed to estimate the box size around the optimal fitting parameters. The clustered optimal minimum was generally composed of several very similar minima with equally good fits that resulted from different minimization starting points, weighting functions, and size of the area searched. In most cases, the size of the box was kept to within ±10% of the values for the optimal equilibrium parameters. If the percent standard deviation from the optimal point calculated for the equilibrium parameters obtained from the randomly selected points approached the boundaries of the box, then a larger box was selected. This procedure was repeated until convergence of the estimates for the variability, with respect to the parameters K and *J*, was observed within a given box size. Our intention is to estimate the sensitivity of each parameter in the vicinity of the optimal minimum, which has been previously obtained. The set of inequalities B2 establishes an upper limit (ceiling) for the objective function, $F(x)$, at any vector point x. The value of this boundary, $\epsilon(\bar{X}^*) = \xi\, F(\bar{X}^*)$, is the weighted average of the standard errors associated with each experimental point in a given integrated titration profile. The same weights used in the original fit are used to calculate $\epsilon(\bar{X}^*)$:

$$\epsilon(\bar{X}^*) = \alpha(1)\,\sigma(1) + ... + \alpha(j)\,\sigma(j), \quad j = 1, 2, 3 \tag{B4}$$

where $\alpha(j)$ are the statistical weights used on a set of curves at a particular insulin concentration (0.07, 0.26, and 0.67 for 0.6, 2.0, and 5.0 mM insulin solutions, respectively), and $\sigma(j)$ are the averages of the standard point-wise error over the set of experimental data for a particular curve. The inequality B2 reflects the central idea in our implementation of a theoretical sensitivity analysis. We are interested in estimating the possible variation in the coordinates around the optimal minimum; however, the resulting variability in the parameters is still constrained by requiring the resulting fits to not deviate outside of the standard error bar associated with the integrated experimental data. The inequality B3 constrains the randomly selected coordinates (generated within our "box") to have the proper limiting behavior at the end (tails) of the integrated curves. The addition of the ΔH_{net} constraint was necessary, as it was found that many of the randomly selected coordinates within our box did not fall within the proper limits of the overall enthalpy. Basically, all the theoretical curves of interest have to be nearly flat in their tails, and the parameter ΔH_{net} is a good determinant for this. We have chosen (or calculated) the following ranges for the parameters above:

$$0.1 \le \Delta \le 0.3$$

$$0.03 \le \xi \le 0.05$$

$$0.01 \le \delta \le 0.05$$

Restrictions B2 and B3 represent a filter of those coordinates chosen randomly within the box set up in B1. The sensitivity analysis is implemented on the filtered set of coordinates by the calculation of the standard deviation for each parameter (coordinate of the objective function). The greater the variability assigned to a particular parameter, then the optimum coordinate, $\bar{X}$, is less robust with respect to that parameter and may also indicate that the dimensionality of the model should be reduced.

It should be noted that by gradually increasing the size of the box (inequality B1), one can try to use this tool for obtaining an answer to the key question about the uniqueness of the global minima for the fitting problem. If there is no convergence in the estimates of the variability in the equilibrium parameters with increasing box size, then other minima may exist which cannot be eliminated from consideration. All attempts thus far have not been able to find any other minima using the optimal model that satisfies all of the constraints placed on the system.

In our opinion, this sensitivity analysis is a good complementary tool to the fitting procedure itself, and we use this analysis to evaluate the family of models with the intention to find the optimal model (see Materials and Methods) for the whole problem.

REFERENCES

Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N., & Vijayan, N. M. (1988) *Philos. Trans. R. Soc. London B 319*, 369−456.

Bloom, C. R., Choi, W. E., Brzovic, P. S., Ha, J. J., Huang, S. T., Kaarsholm, N. C., & Dunn, M. F. (1995) *J. Mol. Biol. 245*, 324−330.

Brader, M. L., & Dunn, M. F. (1991) *Trends Biochem. Sci. 16*, 341−345.

Brange, J., & Langkjær, L. (1992) *Acta Pharm. Nord. 4*, 149−158.

Choi, W. E., Brader, M. L., Aguilar, V., Kaarsholm, N. C., & Dunn, M. F. (1993) *Biochemistry 32*, 11638−11645.

Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., & Swenson, D. (1989) *Nature 338*, 594−596.

Dodson, E. J., Dodson, G. G., Hubbard, R. E., Moody, P. C. E., Turkenburg, J., Whittingham, J., Xiao, B., Brange, J., Kaarsholm, N., & Thogersen, H. (1993) *Philos. Trans. R. Soc. London A 345*, 153−164.

Frank, B. H., & Veros, A. J. (1968) *Biochem. Biophys. Res. Commun. 32*, 155−160.

Gilli, P., Ferretti, V., Gilli, G., & Borea, P. A. (1994) *J. Phys. Chem. 98*, 1515−1518.

Goldman, J., & Carpenter, F. H. (1974) *Biochemistry 13*, 4566−4574.

Huheey, J. E. (1983) in *Inorganic Chemistry* (Wasserman, M., Ed.) pp 359−463, Harper & Row, New York.

Jacoby, E., Krüger, P., Karatas, Y., & Wollmer, A. (1993) *Biol. Chem. Hoppe-Seyler 374*, 877−885.

Krüger, P., Gilge, G., Çabuk, Y., & Wollmer, A. (1990) *Biol. Chem. Hoppe-Seyler 371*, 669−673.

McGraw, S. E., & Lindenbaum, S. (1990) *Pharm. Res. 7*, 606−611.

MicroCal Inc., (1993) in *Origin for the Omega: Tutorial Guide, Version 2.9*, pp 65−69, MicroCal, Northampton, MA.

Pekar, A. H., & Frank, B. H. (1972) *Biochemistry 11*, 4013−4016.

Ross, P. D., & Subramanian, S. (1981) *Biochemistry 20*, 3096−3102.

Roy, M., Brader, M. L., Lee, R. W.-K., Kaarsholm, N. C., Hansen, J. F., & Dunn, M. F. (1989) *J. Biol. Chem. 264*, 19081−19085.

Scholtz, J. M., Marqusee, S., Baldwin, R. L., York, E. J., Stewart, J. M., Santoro, M., & Bolen, D. W. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2854−2858.

Schön, A., & Freire, E. (1989) *Biochemistry 28*, 5019−5024.

Smith, G. D., & Ciszak, E. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 8851−8855.

Sturtevant, J. M. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2236−2240.

van der Varden, B. L. (1960) in *Mathematical Statistics*, p 23, Inostrannaia Literatura, Moscow.

Varshavsky, A. D., Beals, J. M., Birnbaum, D. T., Dodd, S. W., & Saxberg, B. E. H. (1996) *Math. Modeling Sci. Comput. 6*, *in press*.

Wollmer, A., Rannefeld, B., Johansen, B. R., Hejnaes, K. R., Baldschmidt, P., & Hansen, F. B. (1987) *Biol. Chem. Hoppe-Seyler 368*, 903−911.

Wollmer, A., Rannefeld, B., Stahl, J., & Melberg, S. G. (1989) *Biol. Chem. Hoppe-Seyler 370*, 1045−1053.

BI9600557