Precise determination of heme binding affinity in proteins

Authors/Affiliations:

Galvin C.-H. Leung,[a] Simon S.-P. Fung,[a] Nicholas R. B. Dovey,[a] Emma L. Raven[b*] and Andrew J. Hudson[a*]

[a] *Department of Chemistry and the Leicester Institute of Structural & Chemical Biology, University of Leicester, Leicester, LE1 7RH, United Kingdom.*

[b] *School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom.*

Abstract

Accumulating evidence suggests a new role for cellular heme as a signalling molecule, in which interactions with target proteins are more transient than found with traditionally-defined hemoproteins. To study this role, a precise method is needed for determining the heme-binding affinity (or dissociation constant, K_d). Estimates of K_d are commonly made following a spectrophotometric titration of an *apo*-protein with hemin. An impediment to precise determination is, however, the challenge in discriminating between the Soret absorbance for the product (*holo*-protein) and that for the titrant (hemin). An altogether different approach has been used in this paper to separate contributions made by these components to absorbance values. The pure component spectra and concentration profiles are estimated by a multivariate curve-resolution (MCR) algorithm. This approach has significant advantages over existing methods. First, a more precise determination of K_d can be made as concentration profiles for all three components (*apo*-protein/*holo*-protein/hemin) are determined and can be simultaneously fitted to a theoretical-binding model. Second, an absorption spectrum for the *holo*-protein is calculated. This is a unique advantage of MCR and attractive for investigating proteins in which the nature of heme binding has not, hitherto, been characterised because the *holo*-protein spectrum provides information on the interaction.

Introduction

Heme has traditionally been viewed as a prosthetic group that binds tightly to several important proteins and confers upon them specific capabilities for either O_2-storage and transport (the globins), electron transfer (the cytochromes) or biological catalysis (peroxidases and dioxygenases). Such examples of heme-containing proteins are found in all living organisms and detailed structural and mechanistic information has been reported.[1-3] There are, however, reports emerging in the literature which suggest this traditional view of heme biology is too narrow and that heme has a much wider role in regulation of ion channels[4], transcription control of gene expression[5], circadian rhythms[6], sensing of small molecules[7] and signalling[8] in the cell. These new functions for heme are all transient in nature and, therefore a wider portfolio of proteins must be capable of binding heme. The heme concentrations within the cell must be dynamic, and will increase and decrease in selected organelles according to their needs in order to prevent the cytotoxic chemistry of free heme having deleterious effects on the cell. Chaperone proteins will be utilised for heme trafficking between sub-cellular compartments, where heme must be released or transferred to other receptors. The mode of interaction of heme with these proteins is likely to be altogether different to that characterised in globins, cytochromes and enzymes; for example, the binding of heme to any chaperone protein must have a relatively low affinity. Frontier topics in heme biology include the identification of chaperone and trafficking proteins; the characterisation of, hitherto unknown, mechanisms for heme-mediated processes in the cell;[9,10] and the quantification of bioavailable heme in cellular compartments.[11-13] The latter is relevant for investigating homeostasis, metabolism and toxicity of heme.

In order to characterise novel heme chaperones or other receptors in the cell, the binding affinity of heme must be determined precisely. Heme dissociation constants, K_d, given in the literature are typically obtained by titrating an *apo*-protein with heme while recording the UV-visible absorbance of the mixture in the region of the Soret bands (one band appears in the spectrum for the *holo*-protein and a second band for unbound heme). The binding curve is most commonly derived by plotting values of the difference absorbance at the peak of the Soret band for the *holo*-protein, and the K_d is obtained by fitting the plotted data to a theoretical binding model.[14 -22] An alternative approach for measuring the binding affinity can be used if the *apo*-protein is fused to green fluorescent protein (GFP). In this case, the emission band of GFP can be monitored to report on heme binding (this is because the emission from GFP is quenched by heme).[12] Hemin binding to a protein has also been monitored by surface plasmon resonance spectroscopy;[23] and, in (17), hemin binding to human serum album was monitored by fluorescence quenching of tyrosine residues using circular dichroism spectroscopy.

In this work, a straightforward approach has been used to analyse absorption spectra from a spectrophotometric titration. While the measured spectra were obtained from solutions containing varying amounts of an *apo*-protein, a *holo*-protein and unbound hemin, the data could be transformed

into individual spectra and concentration profiles for the pure components using a mathematical algorithm called multivariate-curve resolution (MCR) alternate-least squares analysis. The concentration profiles obtained in this way could be fitted to theoretical-binding curves; from which, the heme binding affinity in the protein could be determined. The precision of the MCR approach, compared with difference absorption spectroscopy, for determination of heme-binding affinities is discussed in this paper. Beyond determination of K_d, the MCR approach has added value for studying novel, or newly discovered, heme-binding proteins because the absorption profile for the pure *holo*-protein is calculated by the computational algorithm. This means that the nature of the heme-protein interaction can be deduced by comparing the peak position, and width, of the Soret band to other examples of well-characterised proteins.

Methods

Spectrophotometric titration of the apo-protein with hemin

Solutions of three different *apo*-proteins, horseradish peroxidase (HRP), ascorbate peroxidase (APX), and myoglobin (Mb), were titrated against a hemin solution. In each case, the volume of the protein stock solution (1 to 3 µM) was 850 µL and additions of the hemin solution (100 to 150 µM) were made in 1 µL increments. The preparation of hemin and protein solutions, and the procedure for the titration is described in the **Supporting Information**. In summary, the titrations were performed in a 1 cm path length quartz cuvette; with 1 mL of the buffer solution placed in a matched-reference cuvette. Following each addition of the titrant to the sample cuvette only, an absorption spectrum was recorded, in increments of 1 nm, from 200 to 700 nm. A separate measurement was made of the absorption spectrum for a 10 times dilution of the hemin titrant.

Deconvolution of UV-visible absorption spectra into pure components representing apo-protein, holo-protein and unbound hemin

A multivariate-curve resolution (MCR) algorithm[24,25] was applied to the sequence of UV-visible spectra recorded in a titration experiment. The purpose of the algorithm was to deconvolve the data set into (a) individual spectral profiles for the *apo*-protein, the *holo*-protein and unbound hemin, and (b) weighted concentrations for the three components in each of the raw experimental spectra. The weighted concentrations were used to illustrate the titration curve for addition of hemin to the *apo*-protein. An outline of the algorithm is given below and a full description in (25).

A matrix, **D** ($m \times n$), was constructed from the m absorption values in each of the n experimental spectra. **D** can be equated to the sum of the concentration profiles for the underlying components, c_k (n), multiplied by their pure spectral profiles, s_k (m):

$$\boldsymbol{D} = \boldsymbol{CS}^{\mathrm{T}} = \sum_{k=1}^{3} \boldsymbol{c}_k \boldsymbol{s}_k^{\mathrm{T}} \qquad (1)$$

The underlying components (k) are the *apo*-protein, the *holo*-protein and unbound hemin. An evolving factor-analysis routine provided the initial estimation of the pure spectral profiles, s_k, and an iterative least-squares method was used to optimise the fit to the experimental spectra, $\boldsymbol{D}$, by minimising the error matrix, $\boldsymbol{E}$:

$$E = D - \sum_{k=1}^{3} c_{k,trial} s_{k,trial}^{\mathrm{T}} \qquad (2)$$

Non-negative least-squares constraints were applied to both the spectral and concentration profiles. Tauler and co-workers developed the scripts and graphical-user interface using MATLAB software (MathWorks, Inc.);[26,27] the software release, managed by Tauler et al., can be downloaded from (28).

The analysis was performed on experimental UV/visible absorption spectra between 250 and 700 nm (i.e. $m = 451$); which excluded the strong bands observed below 250 nm. In addition to the experimental spectra recorded during the titration, a further spectrum for the diluted hemin stock solution was added in the final row of the matrix, **D**. Using these data, the spectral profiles of the pure components, calculated by MCR, were consistent with those expected for an *apo*-protein (peak at 280 nm), a *holo*-protein (peaks at 280 and, typically, ~400 nm) and unbound hemin (peak at 385 nm). Although the concentrations were calculated in arbitrary units by the multivariate algorithm, their profiles were also consistent with those expected from a titration experiment. It was found that, in some cases, the calculated concentration profile for the *apo*-protein gave a significantly poorer fit to a theoretical binding model compared with both the *holo*-protein and unbound hemin. Estimation of the *apo*-protein concentration is challenging because this component shares its principal absorption feature (at 280 nm) with the *holo*-protein, and this region of the absorption spectrum lies on the tail of the strong UV absorption bands for the aromatic rings in heme. As a consequence, the most reliable approach for determining binding affinities involved fitting a theoretical binding model to concentration profiles for the *holo*-protein and unbound hemin only.

Fitting of the concentration profiles to theoretical binding curves

The multivariate analysis of the experimental titration data yielded concentration profiles in arbitrary units for the *holo*-protein, c_{holo}, and unbound hemin, c_{heme}. A non-linear regression analysis was subsequently applied, whereby the concentration profiles were fitted to theoretical curves for heme-binding to an *apo*-protein. The dissociation constant, K_{d}, is given for each incremental addition of hemin in the titration ($i = 1, \ldots, n$) by:

$$K_{\mathrm{d}} = \frac{[apo]_i \times [\mathrm{heme}]_i}{[holo]_i} \qquad (3)$$

where $[apo]_i$, $[holo]_i$ and $[\mathrm{heme}]_i$ are the molar concentrations (in M) of the *apo*-protein, the *holo*-protein and unbound hemin. During the titration of hemin with the *apo*-protein, the total amount of protein (*apo* + *holo*) is fixed, and the total molar concentration of protein will be equal to the initial concentration for the *apo*-protein, p_0, adjusted for the degree of dilution by the titrant:

$$[apo]_i + [holo]_i = p_0 \times \frac{V_0}{V_0 + V_i} \qquad (4)$$

where V_i is the cumulative volume of the hemin solution added, and V_0 is the initial volume of the *apo*-protein solution. The total concentration of *holo*-protein and unbound hemin is given by:

$$[holo]_i + [\text{heme}]_i = h_0 \times \frac{V_i}{V_0 + V_i} \qquad (5)$$

where h_0 is the concentration of the standard hemin solution. Explicit equations for the molar concentrations of the *holo*-protein, $[holo]_i$, and unbound hemin, $[\text{heme}]_i$, can be derived from Eq. 3, 4 and 5; these are given in the **Supporting Information**.

The non-linear regression analysis employed a simplex algorithm[29] in order to find the best fit between the output from the multivariate analysis, c_{holo} and c_{heme}, and the two theoretical profiles for the molar concentrations. This was done by optimising the variables p_0 and K_d, along with two further variables representing scaling factors, α and β, for converting the arbitrary concentration values for the *holo*-protein and unbound hemin, respectively, into molar concentration values. The values of h_0, V_0 and V_i (for $i = 1, ..., n$) were known *a priori*. The optimised parameters were determined by iterative minimisation of the standard error of the fit (see **Supporting Information**). The analysis was performed using custom software written in Fortran 90; the software is available from the authors on request.[30]

Estimating the precision of K_d determined by fitting a binding model to titration curves

The standard error was calculated for the best fit of a theoretical binding model to titration curves, obtained by the MCR algorithm, for the *holo*-protein and unbound hemin. The standard error was subsequently examined in a series of optimisations, in which only the protein concentration (p_0) and the scaling factors (α, β) were treated as parameters. In this case, the K_d was fixed in the optimisations at a series of values displaced from the best fit to the titration curves. By plotting the standard error versus the value of K_d that was used in the optimisation, the precision for determination of K_d by the MCR approach could be determined. This was done by identifying the range of values of K_d across which the change in the standard error of the fit was negligible.

Calculation of difference absorbance at the Soret peak and fitting to a theoretical binding model

A direct comparison was made between the MCR approach for analysis of absorption spectra in a heme titration, and the alternative approach of experimental measurement of difference absorbance.

From each of the absorption spectra measured in the titration experiment, the corresponding difference absorption spectrum was calculated by subtracting the absorption spectrum measured for hemin (scaled according to the cumulative volume of titrant added). By generating the difference absorption spectra from the original measured data, an unbiased comparison between the methods of multivariate-curve resolution and difference absorption could be made. The difference absorbance at the wavelength corresponding to maximum extinction for the *holo*-protein was fitted to a theoretical, single-site, binding model. In this case, the parameters used in the fitting algorithm were the dissociation constant, K_d, and a scaling factor to convert difference absorption into concentration values for the *holo*-protein; in this case, the protein concentration was fixed. The number of degrees of freedom in the non-linear fitting of a concentration profile from difference absorbance (i.e. for the *holo*-protein only) is considerably less than in the fitting of the two concentration profiles from MCR (i.e. for the *holo*-protein and unbound hemin). This was the reason why the protein concentration could not be included as an extra parameter in the fitting algorithm.

The precision for determining K_d by difference absorption was also estimated. This was done by examining how the standard error of the fit changes in a series of optimisations in which only the scaling factor was treated as a variable. Fixed values of K_d displaced from the best fit to the titration curve were used in these calculations, and the precision for determination of K_d was estimated, as before, by identifying the range of values of K_d across which the change in the standard error of the fit was negligible.

Results

Multivariate-curve resolution approach for determining the heme-binding affinity of a protein

The UV-visible spectra recorded in a titration of *apo*-HRP with a standard solution of hemin are shown in **Figure 1 (A)**. Multivariate-curve resolution (MCR) was subsequently used to deconvolve the sequence of spectra into the underlying spectra, and concentration profiles, for the *apo*-protein, the *holo*-protein and unbound hemin; these are shown in **1 (B)** and **(C)**. A fitting of three components captured at least 99.9% of the variance in the experimental data. The residual obtained by subtracting the weighted sum of the pure spectral profiles (calculated by MCR) from the raw experimental spectrum recorded after addition of 16 μL of titrant is shown in **1 (D)** (where the concentrations of *apo* and *holo*-protein are approximately equal). This particular example is representative for the twenty UV/visible spectra shown in **1 (A)**. The signal remaining in **1 (D)** appears to be stochastic, without any significant trace of absorption features for the Soret bands of unbound hemin and the *holo*-protein, and hence supports the suitability of the 3-component model used by the multivariate algorithm.

The pure spectral profiles, shown in **Figure 1 (B)**, for the components are characteristic of the *apo*-protein, the *holo*-protein and unbound hemin, respectively. The peak position of the Soret band, calculated by the MCR algorithm, for *holo*-HRP is 405 nm (**1 (B)**). This value is consistent with the measured value of 403 nm reported in (10). Although the concentration profiles, shown in **1 (C)**, for the three components are calculated by the multivariate algorithm in arbitrary units, they have the characteristic shape of titration curves for the *apo*-protein, the *holo*-protein and unbound hemin. Therefore, these data were fitted to a theoretical, single-site, binding model. For reasons outlined in the **Methods**, the component corresponding to the *apo*-protein was not included in the fitting procedure. The analytical fits to the concentration profiles assigned to the *holo*-protein and unbound hemin are shown in **1 (E)**. Also included in the figure is the theoretical profile for the concentration of the *apo*-protein. The calculation optimised 4 parameters in the single-site binding model: the protein concentration (p_0), K_d, and scaling factors (α and β) for conversion of the concentrations obtained by MCR into quantities with appropriate units of μM. The ability to determine the initial protein concentration from the analytical fit of the binding model is advantageous because the active form can be discriminated from the inactive form in the original protein sample (i.e. partially degraded proteins in which every molecule is no longer able to bind heme). The best fit to the concentration profiles was obtained for a K_d of 0.96 μM and a protein concentration of 2.512 μM. The standard error for the best fit to the data was 3.02×10^{-2} μM.

The precision of the MCR algorithm for determination of K_d was estimated by examining how the standard error (s) changed in a series of optimisations, in which only the protein concentration (p_0)

and the scaling factors (α, β) were treated as variables. Fixed values of K_d displaced from the optimised value of 0.96 μM were used in these calculations. The standard error, s, obtained for the best fit to the binding model using values of K_d between 0.75 and 1.2 μM is shown in **Figure 1 (F)**. The standard error has a sharp minimum indicating that K_d can be precisely determined by using the MCR algorithm with subsequent fitting of the titration curves to a theoretical binding model. The minimum in the distribution of s is indistinguishable between 0.91 and 0.98 μM; and this range represents the precision for determination of K_d for *apo*-HRP.

In this example, it is likely that the precision for fitting theoretical binding curves to concentration profiles obtained by MCR analysis is not the most significant origin of uncertainty in the value of K_d. The uncertainty in reported values of K_d will also reflect the accuracy of the concentration of the hemin titrant. Any error in the concentration of the standard hemin solution will lead to a proportional error in the estimated value of K_d.

Comparison of multivariate-curve resolution to difference absorption for determining the heme binding affinity of a protein

A comparison can be made between the fitting of a theoretical binding model to the concentration profiles obtained by multivariate-curve resolution and difference absorption. In earlier work, a K_d of 0.24 μM was determined for *apo*-HRP by difference absorption.[10] Difference absorption spectra for the titration of *apo*-HRP with hemin are shown in **Figure 2 (A)**. The difference absorbance, ΔA_{405}, at 405 nm reports on the concentration of *holo*-HRP during the titration. In **2 (B)**, values of ΔA_{405} have been plotted against the cumulative volume of the titrant added. The parameters used to fit the profile for ΔA_{405} to a theoretical binding model were the dissociation constant, K_d, and a scaling factor for the linear transformation of ΔA_{405} into concentration values for the *holo*-protein. The protein concentration was set to 2.512 μM. The optimised value of K_d was 0.21 μM and the value of the standard error (s) for the fit was 4.01×10^{-2} μM. The theoretical curve, and the scaled experimental data, is shown in **2 (C)**. The value of K_d obtained by difference absorption has the same order of magnitude as the value obtained by multivariate curve resolution (i.e. 0.96 μM).

The precision for determining K_d by difference absorption was estimated by examining how the standard error (s) changed in a series of optimisations, in which only the scaling factor was treated as a variable. Fixed values of K_d displaced from the best fit value were used in the calculations. The value of s obtained for the best fit to the binding model using values of K_d between 0.005 and 5.0 μM is shown in **Figure 2 (D)** (solid line). The standard error (s) has a broad distribution in which the minimum cannot be clearly distinguished. This suggests that K_d cannot be precisely determined by fitting a binding curve to difference absorption values. In the initial optimisation, the calculation

converged to a value of K_d of 0.21 μM which is located on an inflection point on the curve for the standard error shown in **2 (D)**. For comparison, the corresponding distribution for the standard error obtained by fitting the concentration profiles from MCR is reproduced, as a dotted line, in **2 (D)**. The sharp minimum in this distribution is in stark contrast to the results obtained from difference absorption and illustrates the greater precision that can be obtained for determination of K_d by analysing the concentration profiles produced by MCR.

An advantage of MCR-ALS is that the concentration profile for both unbound hemin (c_{hemin}) and the *holo*-protein (c_{holo}) can be determined from the experimental absorption spectra (see **Figure 1(E)**). The theoretical, single-site, binding model can then be fitted to both of these sets of data points. In contrast, the concentration of the *holo*-protein alone is estimated from the difference absorbance profile, ΔA_{405}, providing fewer data points for fitting to the binding model. The concentration of unbound heme can only be inferred from the difference absorbance profile, ΔA_{405}, after fitting to a theoretical, single-site, binding model (**2 (C)**). As a result, the precision of fitting is much higher for the MCR-ALS method compared to the difference absorption method. The standard error of the fit, s, is indistinguishable between the K_d of 0.91 and 0.98 μM for MCR-ALS; the range of K_d values for difference absorption in which s is indistinguishable is much wider, namely 0.005 to 5.0 μM. For the example illustrated in **Figure 2**, the best fit K_d for *apo*-HRP has been underestimated by difference absorption spectroscopy (*i.e.* a best fit value of 0.21 mM was obtained by difference absorption and a best fit value of 0.96 mM was obtained by MCR-ALS). An underestimate can be a consequence of the difference absorption profile, ΔA_{405}, showing a degree of curvature that is not reflective of the actual concentration profile for the *holo*-protein. Due to a greater degree of curvature in the difference absorption profile, the fitting to a single-site binding model will scale-up the *holo*-protein concentration and scale down the heme concentration.

Extension of the multivariate approach for determining the heme binding affinity of other proteins

Following evaluation of the multivariate procedure for analysis of titration data for HRP, the approach was repeated for other heme-binding proteins. The results obtained for ascorbate peroxidase (APX) and myoglobin (Mb; sperm whale) are shown in **Figure 3**. The latter was titrated against hemin and reduced heme under anaerobic conditions. The optimised fit of a single-site binding model to the concentration profiles obtained by MCR analysis of the UV-visible absorption spectra gave values for K_d of 0.19 μM (*apo*-APX + hemin), 0.60 μM (*apo*-Mb + hemin) and 3.1 fM (*apo*-Mb + reduced hemin). The K_d value for binding of ferric hemin to *apo*-Mb is higher than typical values reported in the literature (for example[31]). Culbertson and Olson have discussed a model that accounts for this discrepancy,[31] whereby Mb exists in equilibrium between three states *native*, *intermediate* and

unfolded (*i.e.*, most structured to least structured) with different values of K_d. A low K_d will be measured in studies of heme dissociation from the *native* state of *holo*-Mb. In contrast, the experimental results reported in **Figure 3 (C)** involved the addition of hemin to *apo*-HRP, and the K_d measured is likely to represent the process of heme binding to *intermediate* or *unfolded* states, which has been reported to have a much higher K_d than the *native* state.[31] The peak of the Soret band, calculated by the MCR algorithm, was found to be 405 nm for *holo*-APX, 408 nm for *holo*-Mb (ferric) and 432 nm for *holo*-Mb (ferrous). Full calculated spectra are provided in the **Supporting Information**. The peak position of the Soret band for the *holo* proteins compare favourably with experimental data previously reported: 406 nm for *holo*-APX[10]; 409 nm for *holo*-Mb (ferric)[32]; 434 nm for *holo*-Mb (ferrous)[32].

Discussion and conclusion

Spectrophotometric titration of the *apo*-protein with hemin is a universal approach for estimating the binding affinity of heme to a protein. In order to make a precise determination of the binding affinity though, the Soret absorbance for the *holo*-protein must be discriminated from that for unbound hemin. This is challenging due to the small red-shift in the peak positon of the band (typically, 20 nm) following heme binding to the protein, and the significant overlap between the edges of individual bands. Difference absorption is the method that has been most widely used to address the problem. This is where a solution containing a mixture of the analyte and titrant is referenced to a blank solution into which an equivalent addition of the titrant is made (alternatively, the protein can be titrated into both a standard sample of hemin and reference[17]). This method adjusts the measured absorbance value for the peak of the Soret band of the *holo*-protein; however, whilst the intention is to compensate for any contribution to the absorbance from unbound hemin, an over correction will actually be made. This is because the amount of unbound hemin in the sample cuvette will be less than that present in the reference cuvette due to the fraction that binds to the protein. As a consequence of this, the binding curve that is derived by plotting values of the difference absorbance (at the peak of the Soret band for the *holo*-protein) against the amount of titrant added will not be accurate; and, consequently, the K_d obtained by fitting the plotted data to the theoretical binding curve for the *holo*-protein will be subject to a significant uncertainly. The magnitude of the error in the *holo*-protein concentration, estimated by difference absorption, will accumulate during the titration experiment and, hence, will be largest after the addition of an equivalent amount of hemin to the protein solution. This is likely to account for why many examples of titration data reported in the literature appear to be truncated, where the titration curve does not extend entirely into the asymptotic region corresponding to complete conversion of protein into the *holo* form (see [10,15-17,21,30]). This will also have a significant impact on how precise experimental data can be fitted to a theoretical binding curve.

The alternative approach to difference absorption is to monitor the absolute absorption value at 405 nm during the course of the titration (i.e. against a reference of the buffer solution only). In this case, the red edge of the band due to unbound hemin will make a direct contribution to the measured absorption value. The concentration of unbound hemin in solution will be high in a titration experiment involving an *apo*-protein with a low binding affinity, or in any titration experiment near/beyond the equivalence point. Under these conditions, the concentration of *holo*-protein will be significantly overestimated, which will impact on the quality of fit between the theoretical binding curve and the experimental data. This approach can be extremely accurate, however, if the extinction coefficients at 385 nm and 405 nm are known *a priori* for both the holo-protein and unbound hemin; then, the contribution made by the tail of the absorption band, at 405 nm, for unbound hemin can be subtracted.[32] In addition to determination of K_d, the titration data obtained by difference absorption

spectroscopy[19,20,33] or by directly monitoring the absorption peak[34 -39] have been used to deduce stoichiometry of protein-heme complexes. As explained above, both of these methods for analysing the absorption bands in a titration experiment are not ideal.

In this paper, the dissociation constant, K_d, for the prosthetic group in hemoproteins has been estimated by a spectrophotometric titration of the *apo*-form of the protein with hemin, and subsequent analysis of the raw experimental data by a multivariate-curve resolution method. Using this approach, the Soret absorbance for the *holo*-protein can be discriminated from that for unbound hemin. This is something that can only be achieved approximately by difference absorption spectroscopy. The multivariate method generated accurate titration plots for a number of different proteins, horseradish peroxidase, ascorbate peroxidase and myoglobin, with both ferric and ferrous heme. The fitting of the titration data to a theoretical-binding curve provided values for K_d. The precision for fitting theoretical curves to MCR concentration profiles *vs.* difference absorbance profiles was tested by examining how the standard error changes in a series of optimisations when the K_d for the theoretical-binding curve is adjusted around the best-fit value. While the standard error exhibits a sharp minimum for the best-fit parameters to MCR concentration profiles, the standard error exhibits an extremely broad minimum in the fitting to difference absorbance profiles. Thus, the method of multivariate-curve resolution is recommended for more precise determination of the binding affinity of heme to proteins. It is strongly recommended under circumstances in which the active protein concentration in unknown. The determination of K_d by difference absorption is particularly imprecise if the protein concentration is used as a parameter in fitting the titration curves to a theoretical binding model. The calculation of the absorption spectrum for the pure *holo*-protein is a further advantage of the MCR approach. As new heme chaperone and heme regulatory proteins are being discovered, this could enable the nature of the heme binding to be deduced, from the peak position and width of the Soret band, for these novel, or newly discovered, proteins.
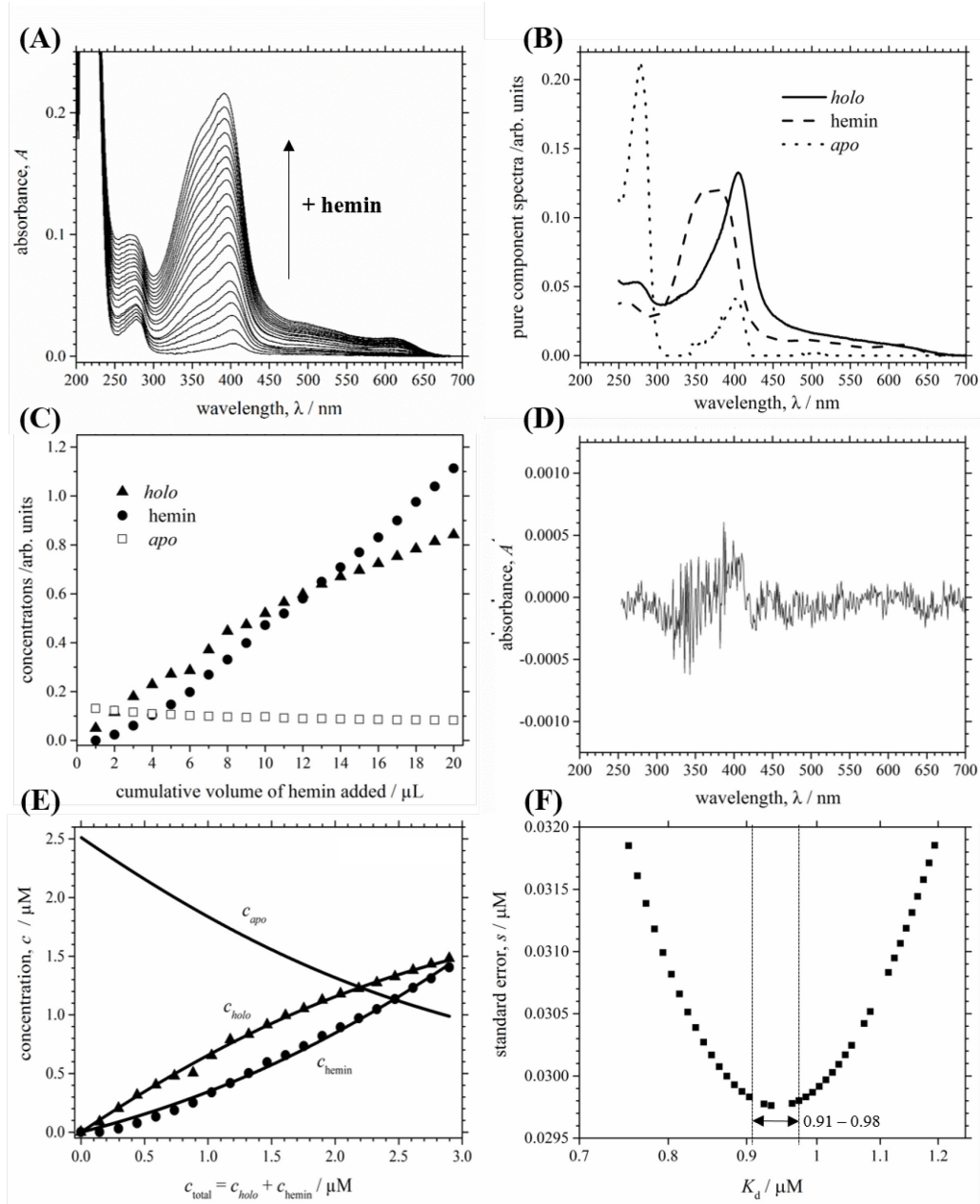
Figure 1: (A) Data from a spectrophotometric titration of *apo*-HRP with hemin. **(B)** and **(C)** The pure spectral profiles and concentration profiles for three components obtained by multivariate-curve resolution. **(D)** The residual obtained by subtracting the weighted sum of the pure spectral profiles (calculated by MCR-ALS) from the raw experimental spectrum recorded after addition of 10 μL titrant. **(E)** Optimised fit of a single-site binding model for HRP to the concentration profiles obtained by multivariate analysis of the UV-visible absorption spectra. The fitting was optimised to *holo*-HRP and unbound hemin concentrations only (the theoretical profile for *apo*-HRP is included for reference); the standard error of the optimised fit to both concentration profiles was 3.02×10^{-2} μM. The optimised value for K_d was 0.96 μM. **(F)** The standard error obtained for the optimised fit to the same concentration profiles, in which K_d is fixed at a series of different values in the calculation. The global minimum can be clearly distinguished in the range 0.91 to 0.98 μM.
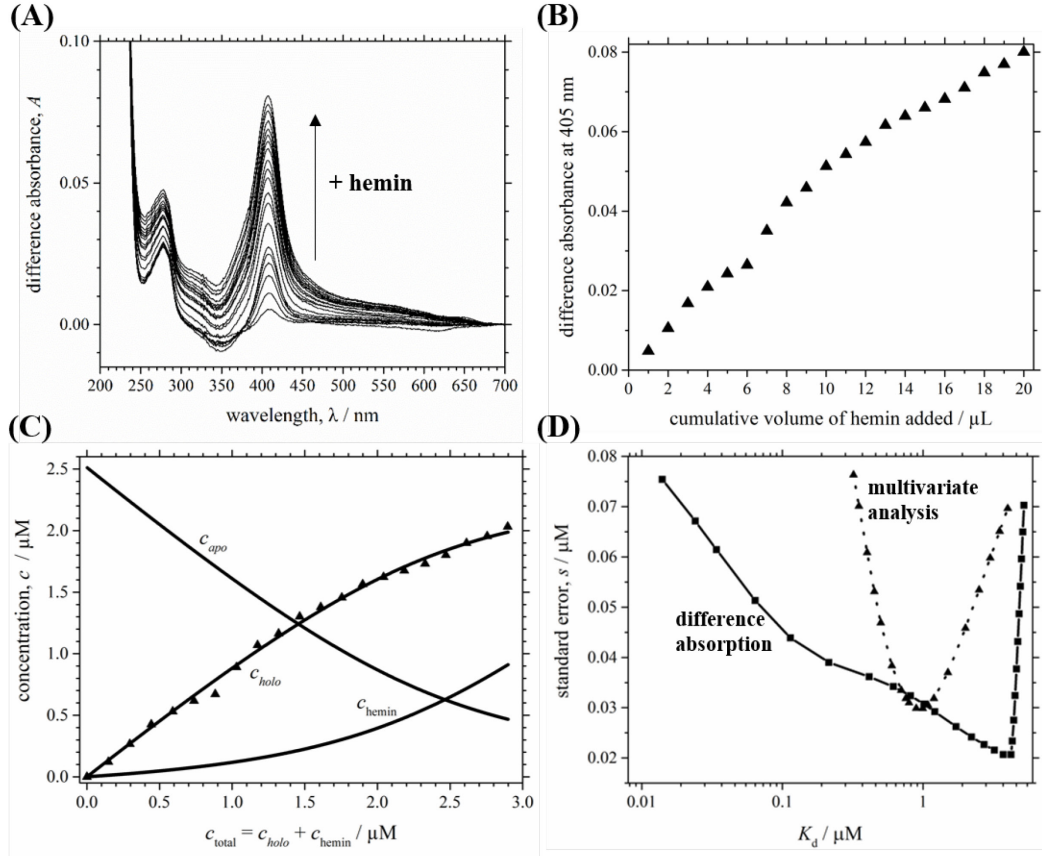
Figure 2: (A) UV-visible difference absorption spectra generated from the same raw data shown in Figure 1 (A.) for a titration of *apo*-HRP with hemin. **(B)** The profile for the difference absorbance at 405 nm, which reports on the increase in concentration of *holo*-HRP during the titration. **(C)** Optimised fit of a single-site binding model for HRP to the difference absorbance at 405 nm. The difference absorbance reports on the concentration of *holo*-HRP during the titration (the theoretical profile for *apo*-HRP and unbound hemin are included for reference); the standard error of the optimised fit was 4.01×10^{-2} µM. The optimised value for K_d was 0.21 µM. **(D)** The standard error obtained for the optimised fit to the difference absorbance at 405 nm (ΔA_{405}; solid line), in which K_d is fixed at a series of different values in the calculation. The global minimum is considerably broader than that obtained from the same analysis performed on concentration profiles calculated by multivariate-curve resolution (shown as a dotted line for reference).
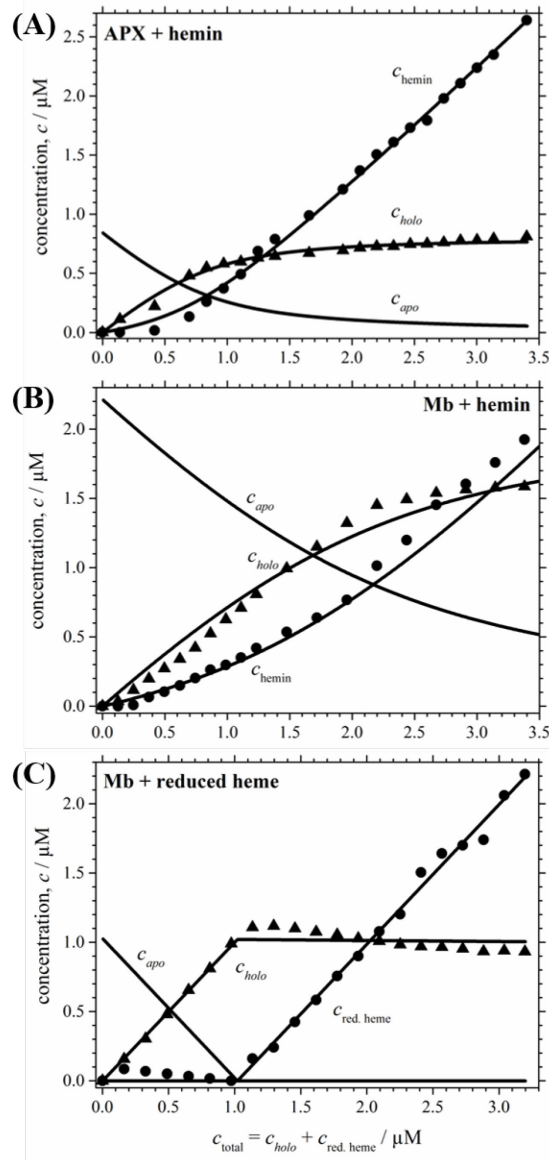
Figure 5: Optimised fit of a single-site binding model for **(A)** APX and **(B)** & **(C)** Mb to the concentration profiles obtained by multivariate analysis of the UV-visible absorption spectra. The fitting was optimised to concentrations for the *holo*-protein and either unbound hemin, in **(A)** & **(B)**, or reduced hemin, in **(C)**. The theoretical profile for the *apo*-protein is included for reference. The optimised value for K_d were 0.19 µM (APX + hemin), 0.60 µM (Mb + hemin) and 3.1 fM (Mb + reduced hemin).

[1] J.H. Dawson, Probing structure-function relations in heme-containing oxygenases and peroxidases, Science 240 (1988) 433-439.

[2] S.G. Sligar, Glimpsing the Critical Intermediate in Cytochrome P450 Oxidations, Science 330 (2010) 924-925.

[3] T.L. Poulos, Heme Enzyme Structure and Function, Chem. Rev. 114 (2014) 3919-3962.

[4] E.L. Raven, M.J Burton, S.M. Kapetanaki, T. Chernova, A.G. Jamieson, P. Dorlet, J. Santolini, P.C. Moody, J.S. Mitcheson, N.W. Davies, R. Schmid, E.L. Raven, N.M. Storey, A heme-binding domain controls regulation of ATP-dependent potassium channels, Proc. Natl. Acad. Sci. USA 113 (2016) 3785-3790.

[5] S. Hira, T. Tomita, T. Matsui, K. Igarashi, K., M. Ikeda-Saito, Bach1, a heme-dependent transcription factor, reveals presence of multiple heme binding sites with distinct coordination structure, IUBMB Life 59 (2007) 542-551.

[6] G.S. Lukat-Rodgers, C. Correia, M.V. Botuyan, G. Mer, K.R. Rodgers, Heme-based sensing by the mammalian circadian protein CLOCK, Inorg. Chem. 49 (2010) 6349-6365.

[7] T. Shimizu, D. Huang, F. Yan, M. Stranava, M. Bartosova, V. Fojtíková, M. Martínková, Gaseous O_2, NO, and CO in signal transduction: structure and function relationships of heme-based gas sensors and heme-redox sensors, Chem. Rev. 115 (2015) 6491-6533.

[8] A. Verma, D.J. Hirsch, C.E. Glatt, G.V. Ronnett, S.H. Snyder, Carbon monoxide: a putative neural messenger, Science 259 (1993), 381-384.

[9] A.R. Reddi, I. Hamza, Heme Mobilization in Animals: A Metallolipid's Journey, Acc. Chem. Res. 49 (2016) 1104-1110.

[10] X. Yuan, N. Rietzschel, H. Kwon, A.B. Walter Nuno, D.A. Hanna, J.D. Phillips, E.L. Raven, A.R. Reddi, I. Hamza, Regulation of intracellular heme trafficking revealed by subcellular reporters, Proc. Natl. Acad. Sci. USA 113 (2016) E5144-5152.

[11] Y. Song, M. Yang, S.V. Wegner, J. Zhao, R. Zhu, Y. Wu, C. He, P.R. Chen, A Genetically Encoded Fret Sensor for Intracellular Heme, ACS Chem. Biol. 10 (2015) 1610-1615.

[12] D.A. Hanna, R.M. Harvey, O. Martinez-Guzman, X. Yuan, B. Chandrasekharan, G. Raju, F.W. Outten, I. Hamza, A.R. Reddi, Heme dynamics and trafficking factors revealed by genetically encoded fluorescent heme sensors, Proc. Natl. Acad. Sci. USA 113 (2016) 7539-7544.

[13] D.A. Hanna, O. Martinez-Guzman, A.R. Reddi, Heme Gazing: Illuminating Eukaryotic Heme Trafficking, Dynamics, and Signaling with Fluorescent Heme Sensors, Biochem. 56 (2017) 1815-1823.

[14] I. Rublevskaya, M.D. Maine, Interaction of Fe-protoporphyrin IX and heme analogues with purified recombinant heme oxygenase-2, the constitutive isozyme of the brain and testes, J. Biol. Chem. 269 (1994) 26390-26395.

[15] A. Wilks, M.P. Schmitt, Expression and characterization of a heme oxygenase (HmuO) from Corynebacterium diphtherae, J. Biol. Chem. 273 (1998) 837-841.

[16] T. Duncan, Y. Osawa, R.K. Kutty, G. Kutty, B. Wiggert, Heme-binding by *Drosophila* retinoid- and fatty acid-binding glycoprotein (RFABG), a member of the proapolipophorin gene family, J. Lipid Res. 40 (1999) 1222-1228.

[17] R. Kannan, D. Sahal, V.S. Chauhan, Heme-Artemisinin Adducts Are Crucial Mediators of the Ability of Artemisinin to Inhibit Heme Polymerization, Chem. & Biol. 9 (2002) 321-332.

[18] T. Muramoto, N. Tsurui, M.J. Terry, A. Yokota, T. Kohchi, Expression and Biochemical Properties of a Ferredoxin-Dependent Heme Oxygenase Required for Phytochrome Chromophore Synthesis, Plant Physiol. 130 (2002) 1958-1966.

[19] X. Zhang, M. Sato, M. Sasahara, C.T. Migita, T. Yoshida, Unique features of recombinant heme oxygenase of Drosophila melanogaster compared with those of other heme oxygenases studied, Eur. J. Biochem. 271 (2004) 1713-1724.

[20] K. Wakasugi, Human Tryptophanyl-tRNA Synthetase Binds with Heme To Enhance Its Aminoacylation Activity, Biochem. 46 (2007) 11291-11298.

[21] N. Gupta, S.W. Ragsdale, Thiol-disulfide Redox Dependence of Heme Binding and Heme Ligand Switching in Nuclear Hormone Receptor Rev-erbβ, J. Biol. Chem. 286 (2011) 4392-4403.

[22] I.C. Soto, F. Fontanesi, R.S. Myers, P. Hamel, A. Barrientos, A Heme-Sensing Mechanism in the Translational Regulation of Mitochondrial Cytochrome *c* Oxidase Biogenesis, Cell Metabolism 16 (2012) 801-813.

[23] E. Karnaukhova, S. Rutardottir, M. Rajabi, L.W. Rosenlöf, A.I. Alayash, B. Åkerström, Characterization of heme binding to recombinant α1-microglobulin, Front Physiol. 5 (2014) 465.

[24] A. de Juan, R. Tauler, Multivariate Curve Resolution (MCR) from 2000: Progress in Concepts and Applications, Critical Reviews in Analytical Chemistry 36 (2006) 163-176.

[25] A. de Juan, J. Jaumot, R. Tauler. Multivariate Curve Resolution (MCR). Solving the mixture analysis problem, Anal. Methods 6 (2014) 4964-4976.

[26] J. Jaumota, R. Gargalloa, A. de Juan, R. Tauler, A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB, Chemometrics and Intelligent Laboratory Systems 76 (2005) 101-110.

[27] J. Jaumot, A. de Juan, R. Tauler, MCR-ALS GUI 2.0: New features and applications. Chemometrics and Intelligent Laboratory Systems 140 (2015) 1-12.

[28] R. Tauler, A. de Juan, J. Jaumot, Multivariate Curve Resolution Homepage, http://www.cid.csic.es/homes/rtaqam/tmp/WEB_MCR/welcome.htm.

[29] W.H. Press, W.T. Vetterling, S.A. Teukolsky, B.P. Flanner, Numerical recipes in Fortran 77, Cambridge University Press, 2nd Edition (1992).

[30] A.J. Hudson, andrew.hudson@leicester.ac.uk.

[31] D.S. Culbertson, J.S. Olson, Role of Heme in the Unfolding and Assembly of Myoglobin, Biochem. 49 (2010) 6052-6063.

[32] B.A. Springer, S.G. Sligar, High-level expression of sperm whale myoglobin in Escherichia coli, Proc. Nati. Acad. Sci. USA 84 (1987) 8961-8965.

[33] P. Kumari, D. Sahal, V.S. Chauhan, Dendrimeric Template of *Plasmodium falciparum* Histidine Rich Protein II Repeat Motifs Bearing Asp→Asn Mutation Exhibits Heme Binding and β-Hematin Formation, PLoS ONE 9 (2014) e112087.

[34] C.J. Reedy, M.L. Kennedy, B.R. Gibney, Thermodynamic characterization of ferric and ferrous haem binding to a designed four-a-helix protein, Chem. Comm. 2003, 570–571.

[35] J. Zhuang, A.R. Reddi, Z. Wang, B. Khodaverdian, E.L. Hegg, B.R. Gibney, Evaluating the Roles of the Heme *a* Side Chains in Cytochrome *c* Oxidase Using Designed Heme Proteins, Biochem. 45 (2006) 12530-12538.

[36] A.M. Thompson, A.R. Reddi, X. Shi, R.A. Goldbeck, P. Moënne-Loccoz, B.R. Gibney, T.R. Holman, Determination of the Heme Affinity for Yeast Dap1p, and its Importance in Cellular Function, Biochem. 46 (2007) 14629–14637.

[37] A.R. Reddi, C.J. Reedy, S. Mui, B.R. Gibney, Thermodynamic Investigation into the Mechanisms of Proton-Coupled Electron Transfer Events in Heme Protein Maquettes, Biochem. 46 (2007) 291-305.

[38] W.J. Huber III, W.L. Backes, Quantitation of Heme Oxygenase-1: Heme Titration Increases Yield of Purified Protein, Anal. Biochem. 373 (2008) 167-169.

[39] Y. Guo, G. Guo, X. Mao, W. Zhang, J. Xiao, W. Tong, T. Liu, B. Xiao, X. Liu, Y.Feng, Q. Zou, Functional identification of HugZ, a heme oxygenase from *Helicobacter pylori*, BMC Microbiology 8 (2008) 226.