บทความรับเชิญ

Jedi's Light Sabre: Site Specific Photocleavage of Proteins with Light

Apinya Buranaprapuk[1] and Challa V. Kumar[2,3*]

ABSTRACT

The ability of photoreagent to target and photocleave the protein backbone at a single site has been a considerable challenge for over decades. The structure of the photoreagent can be systematically modified to provide valuable information on the binding site recognition and photoreaction mechanisms. Many factors, for example, functional groups and overall charge present on the probe, conformations of the bound probe, protein size and amino acids present at the binding site, can affect the photocleavage reaction. A variety of spectroscopic methods (absorption, fluorescence and circular dichroism spectroscopies) were used to monitor the binding interaction. Peptide bond cleavage reactions were obtained from irradiation, at the probe absorption wavelength (~340 nm), of the probe/protein mixture, in the presence of an electron sink. The mechanism indicated the important role of aromatic cation radicals in the mechanistic path as obtained from laser flash photolysis studies. Computer docking studies can be used to provide strong support for the photocleavage and sequencing studies. The docked structures indicated the location of the probe in good proximity to the obtained cleavage site. In this short review, we focused mainly on few generic proteins, such as serum albumin, lysozyme and avidin. However, this strategy may be extended to other proteins with appropriate modifications, which could be useful in a rational design of novel photoreagents to target desired sites on proteins in future studies.

Keywords: Pyrene, Co(III)hexamine, serum albumin, binding constant, Scatchard

[1] Department of Chemistry, Srinakharinwirot University, Sukhumvit 23, Bangkok 10110, Thailand.

[2] Department of Chemistry, 55 N. Eagleville Road, U-3060, University of Connecticut, Storrs, Connecticut, USA. 06269-3060.

[3] Department of Molecular and Cellular Biology, 91 N. Eagleville Road, U-3125, University of Connecticut, Storrs, Connecticut, USA. 06269-3125.

*Corresponding Author: Challa Vijaya Kumar, challa.kumar@uconn.edu

Introduction

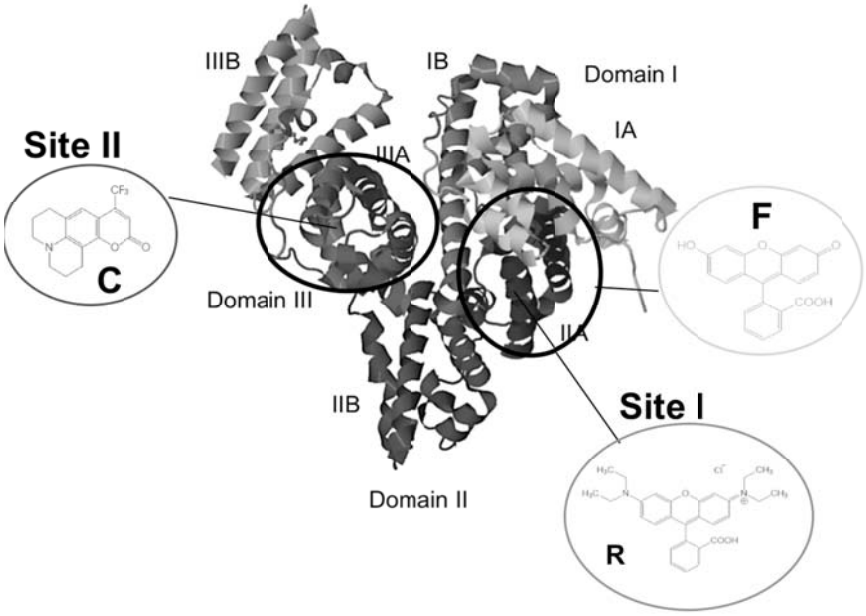
Proteins are made of a sequence of amino acids connected via peptide bonds and this sequence is important in defining the chemical and biochemical properties of proteins. Establishing this sequence of amino acids in a given protein is of fundamental importance in protein biochemistry. With the discovery of large numbers of new proteins on a daily basis makes this task even more challenging and mass spectrometry is a powerful that is being employed in the study of the structures of these proteins (proteomics) [1]. Despite the overwhelming power of mass spectrometry, sequencing a large protein is still a major challenge and conversion of such large proteins to smaller fragments that are more amenable for analysis by mass spectrometry is an important need. This task of converting large proteins into smaller fragments is routinely achieved using enzymes such as peptidases [2] or chemical reagents such as cyanogen bromide [3]. However, these approaches for making a set of peptides from a large protein have serious limitations [4]. To meet this challenge, photochemical reagents that cleave proteins upon activation with light were designed and tested in our group, nearly two decades ago and success with this approach has continued to nourish our work to the current day. The initial approach also included inorganic photoreagents which also demonstrated specificity or high selectivity with yields approaching 60% [5].

In this short review, a couple of examples are taken from our published work and the rational approach used as well as the results obtained have been analyzed from a different perspective. One interesting feature of these artificial photo proteases was that the daughter fragments were amenable to standard biochemical sequencing methods to identify the residues where the original protein has been cleaved [6]. Thus, mass spectral sequencing of the peptides can be connected readily to construct the original protein sequence. Thus, a photocleavage reagent needs to full fill a number of requirements for its utility for protein sequencing studies.

In addition to protein sequencing studies, the above photoreagents could also be useful for identifying biological interfaces. For example, when a protein binds to a solid surface some of its residues are buried at the protein-solid interface and may not be accessible for interactions with the reagent. Thus, the residues at this interface can be readily identified when the photocleavage patterns of the protein in solution are compared with the photocleavage patterns when the protein is bound to the solid. Identifying such interfaces is a major unmet challenge in biology. This understanding could be important to manipulate and control biological processes in the condensed phases, at room temperature, as in cell communication, cell adhesion, cell motion, gene expression/regulation, and other key biological phenomenon [7].

Protein examined in these studies

There are a number of proteins that could be chosen for these studies, but one of the very first proteins chosen for these studies was bovine serum albumin (BSA), among others. BSA (Scheme 1) is a 65 kDa, major serum protein, and it transports nutrients and most drugs entering the circulation [8]. Thus, it is logical to study it for interactions with organic molecules. It has a 3-dimensional heart shaped structure which is made up of three domains and each domain harbors multiple, major binding sites. These binding interactions are supported by an army of hydrophobic, hydrogen bonding and electrostatic interactions [9].

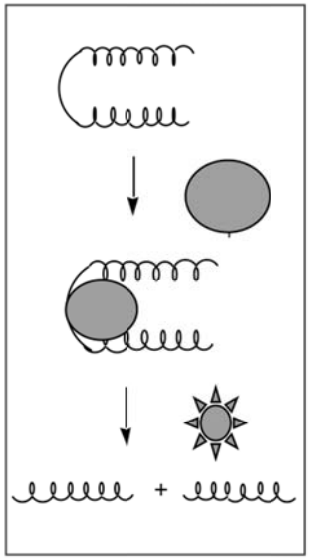


Scheme 1 Three-dimensional structure of bovine serum albumin with specific binding domains and subdomains as established in the literature [10]. The binding sites of Coumarin 450c, Fluorescein and Rhodamine on this protein are shown.

Besides being an inexpensive protein to play with in a chemistry lab, its well-known structure [10] is useful for modeling and testing the binding of ligands [11]. Thus, designing small organic molecules that could bind to BSA is perhaps much more facile than with other proteins. For these various reasons, BSA has been chosen as one of the first proteins to be examined in our laboratories for the photochemical studies described here.

The Approach

One approach to address the above important problem is to design molecular reagents that are capable of light absorption and generate reactive intermediates that can induce the peptide bond cleavage. If the reactive intermediates are generated in the solution, they would react indiscriminately at multiple sites on the protein and result in multiple fragmentations at the backbone. On the other hand, if the reagent binds to the protein at one or few sites and then generates the reactive intermediates then, the peptide cleavage could occur at fewer sites, under a controlled manner (Scheme 2).

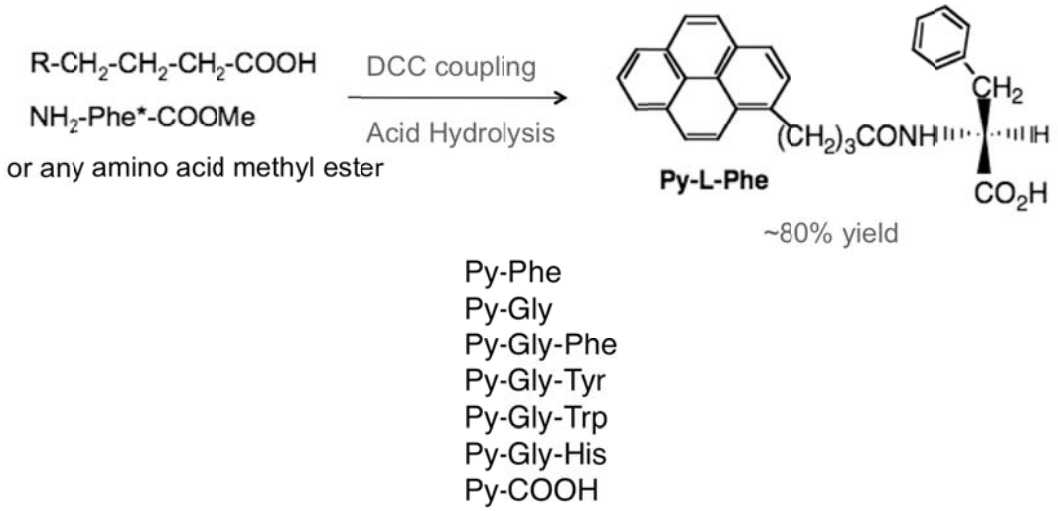


Scheme 2 Photoreagent (blue) binds to a protein at a particular site and then and cleaves the backbone upon irradiation (blue star). The cleavage reaction is more likely to occur at the reagent binding site rather than at a distant site, thereby controlling the selectivity of the photocleavage reaction.

Both the above approaches have their advantages and disadvantages reactive intermediates generated in the solution may never reach the protein during their lifetime. On the other hand, radical intermediates generated on the protein surface may also induce unintended crosslinking of the peptide backbone or form undesirable protein aggregates. These issues can be minimized if the photoreagent binds to a particular site in the interior of the protein where its reactivity is controlled to produce only a limited number of products.

With this simple notion, we have synthesized a few different photochemical probes based on the pyrenyl chromophore. The pyrene ring system has strong absorption bands in the near UV region of the electromagnetic spectrum and hence, the photoprobe derived from this chromophore is less likely to be sensitive to the ambient light of the laboratory. This avoids the use of elaborate methods to work in a dark room which is not so convenient. At the same time, when this chromophore is excited around 340 nm, there is very little chance that the aromatic residues of the protein would interfere. That is, any photoreaction observed will not be due to light absorption by the protein but likely to be due to light absorption by the photoreagent containing the pyrenyl chromophore. Light absorption by the aromatic residues of the protein are generally below 300 nm and thus, these would not interfere with the above experiment.

Few of the photochemical probes synthesized in our laboratories and found to cleave the peptide backbones of at least two different proteins are listed in Scheme 3. These are amino acid derivatives of the above pyrene chromophore, which are prepared by the standard solid state synthesis methods that are used to make peptides. For example, the amino terminus of the any amino acid methyl ester can be coupled with dicyclohexyl carbodiimide (DCC) to the COOH group of the pyrene-butyric acid to give the corresponding adduct and the methyl ester is readily hydrolyzed in dilute acid to provide the final product (Scheme 3). Either of the optical isomers may be chosen for the amino acid esters, thereby providing significant control over the chirality of the photoprobe.



Scheme 3 Synthesis of pyrenyl peptides derived from amino acid methyl esters and pyrene butyric acid by standard methods of peptide synthesis. A few of the photoprobes synthesized by this approach are listed below.

The purpose of using the amino acids in the probe design is three-fold: (1) adopt a well known synthetic strategy to prepare a number of derivatives of the photoprobe so that a structure-activity correlations can be sought; (2) the amino acid residue can provide a chiral center and thereby providing opportunities to examine the chiral discrimination provided by the protein environment for improved control over the selectivity for binding and subsequent photocleavage, and (3) the amino acids of the probe might interact favorably with the amino acids of the protein backbone and this interaction might promote the binding affinities of the probes. The data, obtained thus far, appear to support these expectations. Thus, a number of different amino acids were coupled with the pyrenyl chromophore and the corresponding photoprobes are shown in Scheme 3.

Results

Absorption titrations

Prior to conducting the photocleavage reaction, one needs to make sure that the photoprobe has certain affinity for the target protein and also evaluate the number of binding sites that the probe might reside at. Very weak binding affinities would imply that the probe would mostly reside in the solvent and any reactive intermediates generated by light activation would then shower the protein with these. Such a situation might result in protein crosslinking rather than desired cleavage of the peptide backbone of the target protein. On the other hand, strong to moderate binding of the photoprobe with the target protein would ensure that the reagent is more likely to be localized on the protein and hence, any derived chemistry would be directed at particular sites on the protein. This latter result could produce specific cleavage patterns near the probe binding site and hence, this pattern could be controlled by adjusting the probe structure in a rational manner to target different sites on the protein.

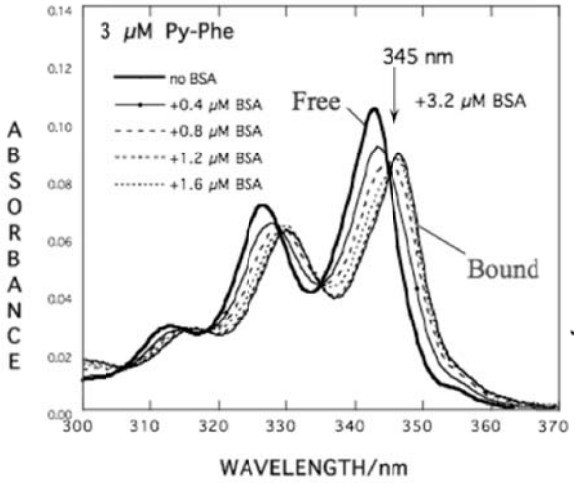
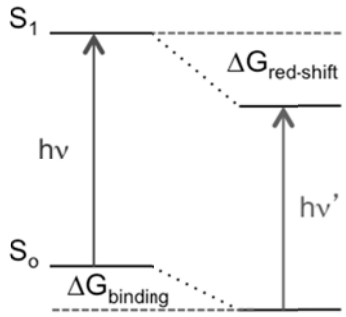


Figure 1 Titration of a concentrated solution of bovine serum albumin (BSA) into a solution of the photoprobe (Py-L-Phe), while monitoring the absorption spectrum of each mixture. The spectra are corrected for small changes in the volumes and overlaid to illustrate small but significant changes.

Thus, we have examined the binding of Py-L-Phe shown in Figure 1 in absorption titrations with bovine serum albumin while recording the absorption spectra, after equilibration. The titration of a solution of BSA into a solution of Py-L-Phe (3 μM) for example, indicated significant changes in the absorption spectrum of the photoprobe in the 300-370 nm region. Care is exercised in keeping the volume of the mixture the same or corrections are made to the observed spectra so that they could be compared readily.

A number of changes are immediately clear in the spectral changes accompanying the binding of Py-L-Phe with BSA. As the titration progressed, the peak positions of the absorption spectra are gradually red-shifted, and as a result the spectra intersected at 345 nm. This intersection point is important because the absorption of the probe at this wavelength is invariant with respect to the protein concentration. This intersection is called the isosbestic point or the point of constant absorption. There few other such isosbestic points shown in Figure 1 and they are significant in terms of what is happening in the solution as the titration is progressing.

The presence of isosbestic points implies that there is equilibrium between two distinct chemical species and as the titration progresses, one species is being converted into the other. Our interpretation of this situation is that the free probe is being converted to the bound probe, as one adds more and more protein to the solution, thus, shifting the chemical equilibrium toward the bound state. For example, if there is a third species in significant concentrations which absorbs differently from either of these species then the isosbestic point will be defocused or diffuse or shifts as the titration continues. Thus, the spectral data strongly suggest that the free probe is gradually being converted to the bound form and that there is one major species in the solution.



Scheme 4 Changes in the energies of the electronic states of Py-L-Phe as it binds to bovine serum albumin. The binding free energy may be computed from the red-shift and energies of the absorption peak positions of the free ($h\nu$) and bound ($h\nu'$) probes, as shown.

Any precipitation of the probe-protein complex would have resulted in upward-moving baseline which was not observed here, and the peak absorbances would have decreased significantly. In addition, close examination of the titration mixture did not reveal any obvious precipitation. Thus, one could conclude that the free probe is being converted smoothly to the bound probe and that the bound probe absorbs at longer wavelengths than the free probe. This red-shift in the absorption is significant because it implies a decrease in the energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) when the probe nestles into the protein matrix (Scheme 4).

The energy level diagram depicted in Scheme 4 provides an interesting approach to measure the binding free energy and it is simply the difference in the quantities depicted on the right vs the left halves of the scheme. For example, the sum of the binding free energy ($\Delta G_{binding}$) and the energy corresponding to the absorption peak position of the free probe ($h\nu$) should be equal to the energy corresponding to the red-shift ($\Delta G_{red\text{-}shift}$) plus the energy of the red-shifted absorption transition ($h\nu'$). However, this analysis assumes that the same excited electronic state is populated in the free and bound forms and that the shift is solely due to the binding of the probe to the protein.

The decrease in the free energy of the ground state (HOMO) is due to binding of the probe to the protein where the protein matrix replaces the solvent around the chromophore, while the decrease in the free energy of the excited state (LUMO) is also due to the solubilization of the excited state by the protein matrix. Note that the stabilization of the HOMO is less than the stabilization of the LUMO and hence $h\nu > h\nu'$, the concomitant red shift in the absorption spectrum much greater than that of the ground state. Thus, greater stabilization of the excited state by the protein when compared to that of the ground state which resulted in the red-shift of the absorption spectrum implies that the protein interior is more conducive to the excited state than to the ground state. There would have been a blue shift, if the opposite has been true.

Thus, one should not conclude that there is no binding if there is no shift in the absorption spectrum when a probe is titrated into the protein solution. That could be true but if both HOMO and LUMO are stabilized or destabilized to the same extent, then there would be no change in the energy gap between the HOMO and LUMO. Hence, there will be no shift in the absorption peak position. This scenario can happen and so, the absorption titration data are to be interpreted more carefully. Nevertheless, one important and unexpected dividend from the above treatments is the possibility for the direct assessment of the binding free energy of the probe with the protein, from simply analyzing the above absorption data and the peak positions.

The peak absorption wavelengths are to be first converted into the corresponding frequencies (energy units) and the red shift is to be obtained. Then, one can estimate the binding free energy as given by equation 1, deduced from the energy equality argument explained above.

$$\Delta G_{binding} = \Delta G_{red\text{-}shift} + h(\nu' - \nu) \text{..... equation 1}$$

The important assumption made in the above equation is that the electronic transition is between the same states in the free probe and the bound probe. If this is not true, then appropriate corrections need to be made. Secondly, one can use equation 1 for all the vibrational bands observed in the absorption spectrum and hence, multiple measurements of binding free energy is possible, from a single set of absorption titrations. Thus, a more accurate, average value for the binding free energy can be obtained without a calorimeter. In addition, we need not assume that there is an equilibrium between the bound and free probes, only that we need to ensure complete binding occurs so that there is no free probe in the final solution of the titration. By adding excess protein, one can often capture all the free probe and convert to the bound probe so that its absorption spectrum could be obtained more reliably.

Fluorescence titrations

In addition to the absorption studies, one could also examine the fluorescence spectra of the pyrenyl probes to assess their binding affinities. For example, gradual titration of a concentrated solution of the protein into the probe solution resulted in a decrease in the intensities of the fluorescence. Note that the samples were excited at the isosbestic point, equal absorbance, so that the amount of light absorbed by the sample remains independent of the protein concentration. Under these conditions, one obtained the following fluorescence spectra for Py-L-Phe.

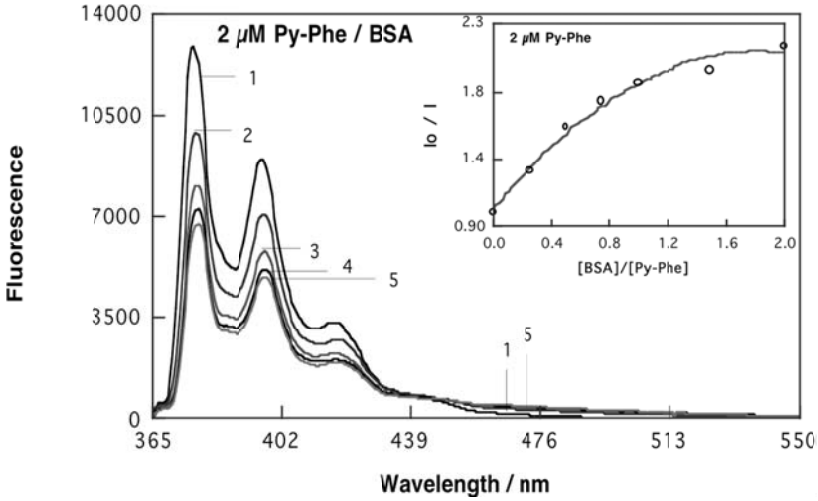
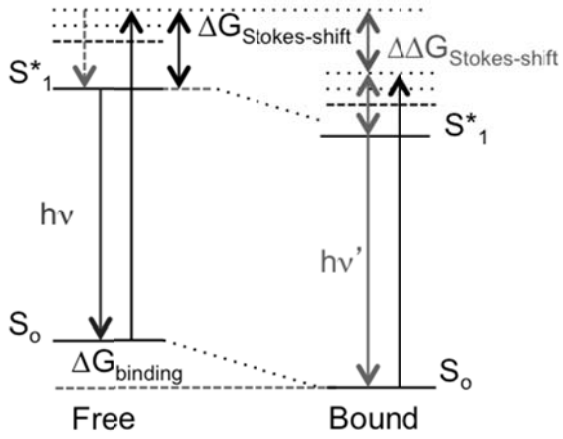


Figure 2 The fluorescence titrations of Py-L-Phe (2 μM) with bovine serum albumin (0-5 μM, in equal increments) while exciting at the isosbestic point shown in Figure 1.

The fluorescence data shown in Figure 2 have two separate aspects. One is that as the protein is titrated into the probe solution, there has been a gradual decrease in the intensities of the bands but there have been no detectable shifts in the peak positions. The other aspect is that a new band appeared above 470 nm, which is broad and weak. This latter band is common for pyrenyl chromophores where an excited state complex with a ground state of the probe or with another molecule is responsible for this new emission band. But this aspect is not discussed any further but we will get back to the major spectral changes noted above. That is, the emission peaks remained the same or that the associated energies of the emission appears to be nearly the same. This observation immediately implies that the energy gap between the singlet excited state that is radiant and that of the ground state are exactly the same for both the free probe and the bound probe. This appears to be in direct contradiction to the discussion under the absorption titrations, presented above. A closer look at the energy levels of the states responsible for the fluorescence makes the situation clearer (Scheme 5).



Scheme 5 Comparison of the energy levels of the initial and final states of Py-L-Phe emission when the probe is free and bound to bovine serum albumin. The energy gap between the absorption maximum and the emission maximum is the Stoke's shift.

The emission data imply a similar situation where the equation 1 can be recast in terms of the emission energies, instead of the absorption energies. However, one needs to include the energies corresponding to the Stoke's shift of the emission spectra with respect to the absorption spectra, in the case of the free probe as well as the bound probe. Thus, we have equation 2 where $\Delta G_{binding} + h\nu + \Delta G_{Stoke's\ shift} = \Delta G_{red\text{-}shift} + \Delta G'_{Stoke's\ shift} + h\nu'$. Where, $\Delta G'_{Stoke's\ shift}$ is the shift of the emission spectrum of the bound species from its absorption spectrum, and $\Delta G_{red\text{-}shift}$ is the red shift observed in the corresponding absorption spectra discussed above (Scheme 5). Thus, the red-shift in the absorption spectrum and the Stoke's shifts of the emission bands of the free and bound probes are directly related to the binding free energy.

$$\Delta G_{binding} = \Delta G_{red\text{-}shift} + \Delta G'_{Stoke's\ shift} - \Delta G_{Stoke's\ shift} + h(\nu' - \nu) \text{ ..... equation 2}$$

An interesting scenario is depicted from the data in Figure 2, where the emission peak positions did not shift upon titrating the probe with the proteins or $h\nu = h\nu'$. Then, from equation 2 we get the binding energy as: $\Delta G_{red\text{-}shift} + \Delta G'_{Stoke's\ shift} - \Delta G_{Stoke's\ shift}$. Thus, $\Delta G_{binding} = \Delta G_{red\text{-}shift} + \Delta\Delta G_{Stoke's\ shift}$. Thus, the binding free energy is the sum of the red shift and the differences in the Stoke's shifts, an elegant scenario. Thus, two alternative ways of estimating the binding free energies are obtained, independent of any calorimetric determinations.

Binding Isotherms

In addition to the above spectral analysis, the absorption and emission titration data are generally analyzed using Scatchard [12] model to extract binding constants and the number of binding sites on the protein. This model assumes that the binding is discrete or that binding to one site does not influence binding to another site on the same protein molecule. This is called the non-cooperative binding. Secondly, binding is assumed to be reversible and the data are obtained under equilibrium conditions so that the binding constant is calculated based on the chemical equilibrium established during the binding event. Each protein molecule is assumed to have a finite number of non-interacting binding sites for a given ligand. The number of binding sites are thus, expected to be integral numbers and when $n > 1$, the binding sites do not overlap with each other. With these assumptions, the Scatchard equation is written as in equation 3.

$$r/C_{free} = K_b\ (n\text{-}r) \text{ where } r = C_{bound}/[\text{Protein}] \text{ ..... equation 3}$$

There are a few terms in equation that need to be identified and they are, the number of binding sites (n, unitless) per protein molecule that are identical and non-overlapping, the binding constant (K_b, units of M^{-1}) and the binding density (r, dimension less) which is defined as the ratio of total ligand bound (molar) per total protein (molar) available for binding.

One last thing to discuss is how to estimate the above quantities from the experimental data. For example, the concentration of the total protein is directly obtained from the titration data, after appropriate correction to the changes in the total volume, as the titration progressed. The other is the concentration of the ligand bound at each concentration of the ligand added to the titration. Usually this is obtained from the extinction coefficients of the bound and free ligands at the wavelength of analysis and the total absorbance of the sample at that wavelength.

The total absorbance measured at any wavelength is the sum of the absorbances of the free and bound ligands at that wavelength. These absorbances are, in turn, equal to the concentration of the corresponding probe (free or bound) multiplied by their molar extinction coefficients. One also has the relation that the total ligand concentration is equal to the sum of the concentrations of the free and bound ligands, at each protein concentration. Using these equations, one transforms the above Scatchard equation in terms of absorbances or directly calculates the concentrations of the free and bound ligands at each of the protein concentrations. The extinction coefficient of the bound ligand at the wavelength of analysis is obtained by adding excess protein to the ligand solution such that the ligand is completely consumed by the protein. Or this is also obtained by plotting the observed absorbance of the sample at the wavelength of analysis vs inverse of the protein concentration. By extrapolation of this plot to infinite concentration of the probe, one obtains the Y-axis intercept which gives the absorbance of the fully bound probe and from this, one obtains its extinction coefficient. Either way, one obtains a set of r and ligand concentrations and obtains the Scatchard plot shown in Figure 3.

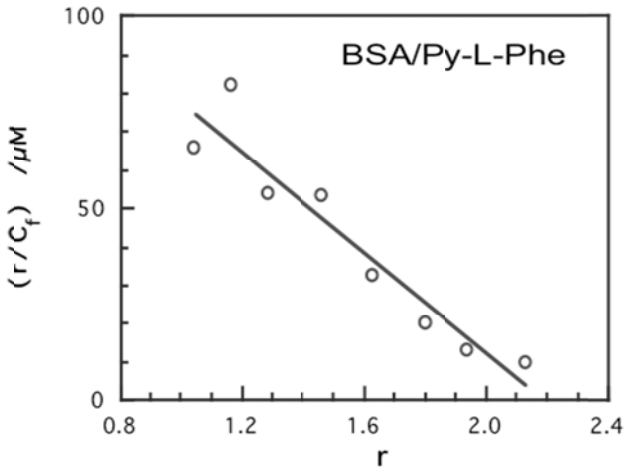


Figure 3 The Scatchard plot for the titration shown in Figure 1. The binding constant estimated from these data was $6.5x10^7$ M^{-1}.

Linear fits to the binding isotherms by the Scatchard equation provided the corresponding binding constants and the number of binding sites. These are collected in Table 1 for comparisons.

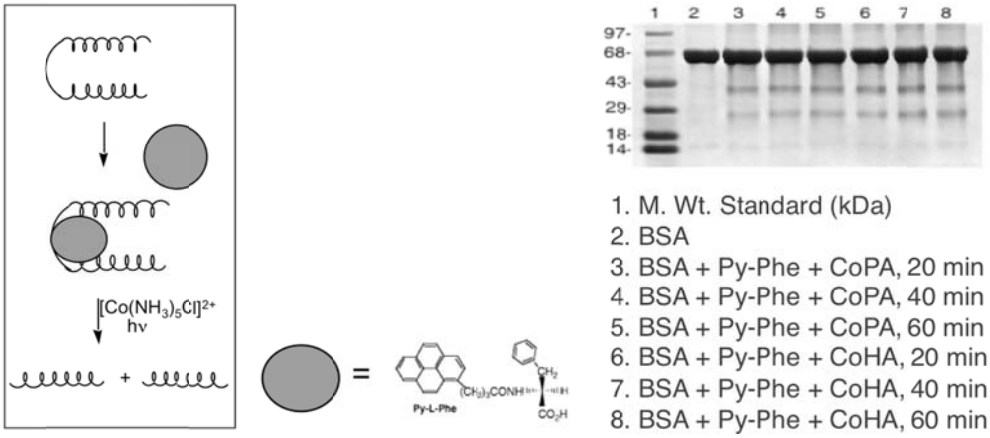
Table 1 Binding constants (K_b) for some pyrenyl probes with BSA.

Py-L-Phe	**Py-D-Phe**	**Py-Gly-Phe**	**Py-Gly-Gly-Phe**	**Py-Phe-Gly-Gly**
$6.5x10^7$ M^{-1}	$5.3x10^5$ M^{-1}	$1.5x10^6$ M^{-1}	$1.5x10^6$ M^{-1}	$1.5x10^7$ M^{-1}
Py-Gly-Phe	**Py-Gly-Tyr**	**Py-Gly-Trp**	**Py-Gly-His**	**PMA-L-Phe**
$1.5x10^6$ M^{-1}	$4.3x10^6$ M^{-1}	$6.7x10^5$ M^{-1}	$3.6x10^5$ M^{-1}	$3.58x10^5$ M^{-1}

The data in Table 1 led to some interesting comparisons: binding was non-cooperative in all cases examined and comparisons of the K_b values provided insight into the contributions of specific groups on the ligand toward the binding free energy. For example, Py-L-Phe binds with an affinity that is two orders of magnitude greater than that of Py-D-Phe, this clearly points that the chiral center of phenyl alanine moiety is significant. The tail wags the dog is an appropriate analogy here. Such a minute difference in the orientation of the chiral group depletes the affinity to a significant extent and this is the first man-made molecule to show 100-fold chiral discrimination. This was a world record when first published but there could be other examples by now [5]. The optical isomers used in Table 1 are the L-isomers, unless stated otherwise. The binding affinities are also affected by the linker length, charge or the recognition element. The binding is also decreased when a larger hydrophilic element is introduced into the side chain (Trp, Tyr, His) instead of a hydrophobic residue such as Phe [13].

Protein Photocleavage Studies

One other detail that required to be fixed is to trick the excited state of the pyrenyl chromophore to undergo photochemical reactions to generate reactive intermediates. This was readily accomplished by redox-quenching of the excited state with another reagent that preferred to stay in the solvent and had no interest in binding to the protein. That is, the excited state of our photoreagent is quenched by an electron deficient metal complex that robs an electron from the photoreagent and generates the corresponding pyrenyl cation radical. The cation radicals are generally highly reactive and abstract hydrogen atoms from their surroundings or the solvent. The radical intermediates, thus generated on the protein, are expected to induce further reactions and eventually led to the cleavage of the peptide backbone.



Scheme 6 The photoproducts arising from the irradiation of bovine serum albumin with Py-L-Phe in the presence of $[Co(NH_3)_5Cl]^{2+}$ (or CoPA) (lanes 3-6) or in the presence of $[Co(NH_3)_6]^{3+}$ (or CoHA) (lanes 6-8), as the quenchers.

An example of the above strategy is illustrated in SDS-PAGE of BSA (right figure, Scheme 6), where irradiation of a solution (at 345 nm, 20-60 min) of bovine serum albumin (BSA, 15 μM) with our photoreagent (Py-Phe, 15 μM) in the presence of electron hungry Co(III)polyammine complexes indicates the formation of only two product bands with lower molecular weight than the parent protein. The observed two photocleaved fragments suggests a single cut in the peptide backbone, and the molecular weights of the two product bands sums up to that of BSA.

Conclusion and Discussion

A number of small organic molecules have been designed, synthesized and tested for their binding to proteins which indicated several examples where the photoreaction can be successfully induced. The binding affinities, however, depended on probe structure in a complex manner, and could not be rationalized by any simple binding model. While their affinities differed, they all responded substantially on binding to the protein, in terms of their photophysical properties. Protein binding resulted in substantial changes in the absorption and fluorescence spectra and the data have been quantified using Scatchard equation. In all the cases examined here, except for Py-Phe/BSA, all systems indicated a single binding site with moderate to high affinities. Binding to the protein also buried the chromophore in the protein matrix and its access to the solvent was successfully probed in fluorescence quenching studies with CoHA. Furthermore, the degree of protection offered by the protein against quenching with CoHA correlated inversely with the yield of the photocleavage reactions. Thus, some access to the bound chromophore was essential to initiate the photoreaction.

The newly created N-terminus of the photoproduct was amenable to conventional sequencing studies, which is an important result with ramifications about the application of these photoreagents in biochemical studies. However, the newly created C-terminus was not sequencible but these constrained helped us in narrowing down the possible mechanistic schemes that could be devised. Another significant observation was that many of the photoreagents shared the same cleavage site on BSA and lysozyme, with some exceptions. This could be because they all had the large hydrophobic pyrenyl group but the side chain played an important role. Longer chains allowed the probe to access residues that are farther and new cleavage sites were noted. Along the same lines, the attachment of biotin in the side chain allowed Py-Biotin to be directed to the biotin binding site and this is likely due to the high affinity of biotin to its binding site on avidin.

Computer docking studies provided strong support for the photocleavage and sequenching studies. The docked structures indicated good proximity of the photoreagent to the observed cleavage site and the experimental data have validated the success of the computer modeling. In future studies, computer modeling may be used in an iterative manner, to target specific sites on a given protein target. We are not there yet, but with increased computational power and availability of reliable and sophisticated algorithms should make this a viable strategy to photocleave proteins at desired sites.

In summary, the photoreagent design should consider several factors, such as potential binding site on the target protein, desired structural features of the photoreagent to bind at the binding site and the excited state properties. The excited state need to live long enough to induce the desired reaction at the protein binding site but it also should have the necessary chemical reactivity. CoHA played an important role in the examples presented here, where it generated a long lived radical intermediate from the pyrenyl chromophore, which successfully induced the photocleavage of the protein backbone.

Acknowledgments

We thank the National Science Foundation for the financial support of the work (EAGER DMR-1441879) and the University of Connecticut OVPR for the Research Excellence Award.

References

1. Li, H., Nguyen, H. H., Loo, R. R. O., Campusano, I. D. G., and Loo, J. A. 2014. An Integrated Native Mass Spectrometry and Top-Down Proteomics Method that Connects Sequence to Structure and Function of Macromolecular Complexes. *Nature Chemistry*. 10: 139-148.
2. Rawlings, N. D., and Barrett, A. J. 2005. Peptidases, Families and Clans. Wiley Online Library. Available from URL: https://doi.org/10.1002/047001153X.g306216. 15 April 2005.
3. Crimmins, D. L., Mische, S. M., and Denslow, N. D. 2001. Chemical Cleavage of Proteins in Solution. Wiley Online Library. Available from URL: https://doi.org/10.1002/0471140864.ps1104s19. 15 May 2001.
4. Kumar, C. V., and Buranaprapuk, A. 1997. Site-Specific Photocleavage of Proteins. *Angewandte Chemie-International Edition in English.* 36: 2085-2087.
5. Buranaprapuk, A., Malaikaew, Y., Svasti, J., and Kumar, C. V. 2008. Chiral Protein Scissors Activated by Light: Recognition and Protein Photocleavage by a New Pyrenyl Probe. *Journal of Physical Chemistry B.* 112: 9258-9265.
6. Kumar, C. V., Buranaprapuk, A., Opiteck, G. J., Moyer, M. B., Jockusch, S., and Turro, N. J. 1998. Photochemical Protease: Site-Specific Photocleavage of Hen Egg Lysozyme and Bovine Serum Albumin. *Proceedings of the National Academy of Sciences of the United States of America.* 95: 10361-10366.
7. Kumar, C. V., and Buranaprapuk, A. 1999. Tuning the Selectivity of Protein Photocleavage: Spectroscopic and Photochemical Studies. *Journal of the American Chemical Society.* 121: 4262-4270.
8. Quinlan, G. J., Martin, G. S., and Evans, T. W. 2005. Albumin; Biochemical Properties and Therapeutic Potential. *Hepatology.* 41: 1211-1219.

9. Jisha, V. S., Arun, K. T., Hariharan, M., and Ramaiah, D. 2010. Site Selective Interactions: Squaraine Dye-Serum Albumin Complexes with Enhanced Fluorescence and Triplet Yields. *Journal of Physical Chemistry.* 114: 5912-5919.
10. Majorek, K. A., Porebski P. J., Dayal, A., Zimmerman, M. D., Jablonska, K., Stewart, A. J., Chruszcz, M., and Minor, W. Structure and Immunological Characterization of Bovine, Horse, and Rabbit Serum Albumins. *Molecular Immunology.* 52: 174-182.
11. Patra, S., Samthosh, K., Pabbathi, A., Samanta, A. 2012. Diffusion of Organic Dyes in Bovine Serum Albumin Solution by Fluorescence Correlation Spectroscopy. *Royal Society of Chemistry Advances.* 2: 6079-6086.
12. Scatchard, G. 1949. The Attractions of Proteins for Small Molecules and Ions. *Annals of the New York Academy of Sciences.* 51: 660-672.
13. Buranaprapuk, A., Kumar, C. V., Jockusch, S., and Turro, N. J. 2000. Photochemical Protein Scissors: Role of Aromatic Residues on the Binding Affinity and Photocleavage Efficiency of Pyrenyl Peptides, *Tetrahedron.* 56: 7019-7025.