Monitoring protein interactions and dynamics with solvatochromic fluorophores

Galen S. Loving¹, Matthieu Sainlos¹,² and Barbara Imperiali¹

¹Department of Chemistry and Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139-4307, USA. ²Centre National de la Recherche Scientifique, Université de Bordeaux, UMR 5091, Bordeaux, France.

Corresponding author: Imperiali, B. (imper@mit.edu)
Abstract

Solvatochromic fluorophores possess emission properties that are sensitive to the nature of the local microenvironment. These dyes have been exploited in applications ranging from the study of protein structural dynamics to the detection of protein-binding interactions. While the solvatochromic indole fluorophore of tryptophan has been utilized extensively for *in vitro* studies that advance our understanding of basic protein biochemistry, new extrinsic synthetic dyes with improved properties together with recent developments in site-selective methods to incorporate these chemical tools into proteins now open the way for studies in more complex systems. Herein we discuss recent technological advancements and their application in the design of powerful biosensors, which serve critical roles in modern cell biology and assay development.
Introduction

Fluorescence-based reporters are vital in the fields of molecular cell biology and protein biochemistry. The techniques that implement these reporters are based on fundamental photophysical principles including Förster resonance energy transfer (FRET) [1], fluorescence polarization (FP) [2], and fluorescence correlation spectroscopy (FCS) [3] as well as a set of fluorescent dyes exhibiting unique and varied photophysical properties [4]. Innovations in instrumentation have allowed fluorescence-based techniques to be applied in a wide array of formats ranging from high throughput screening assays [5] for drug discovery to the latest in state-of-the-art super-resolution microscopy [6] for in cellulo studies. Of particular interest are fluorescent dyes that possess emission properties that are responsive to physical changes in the local environment including pH, viscosity, biological analytes, and solvent polarity. By conjugating probes of this type to biologically relevant molecules, it is possible to obtain valuable information regarding the functions, activities and interactions of such species in the context of living systems with spatial and temporal resolution.

This review focuses on recent advances in the development, incorporation, and application of a specific class of environment-sensitive fluorophores that display solvatochromism (box 1). Solvatochromic fluorophores demonstrate emission properties (e.g. fluorescence lifetimes, emission wavelengths, and quantum yields) that are highly sensitive to the immediate environment. The dynamic behavior makes these species particularly well suited for investigating biomolecular interactions since it provides information on the state of a protein at the single amino acid level. For example, if a solvatochromic fluorophore is appended to the surface of a protein at a site that is involved in a transient interaction or that undergoes a
conformational change, then the probe will report binding or allosteric changes if these events are coupled to modifications in the local solvent sphere. Herein, particular attention is devoted to the incorporation of these tools into peptides and proteins for the development sensors of biomolecular dynamics. In surveying a selection of these applications, a brief analysis of the associated advantages and limitations of various probes is also presented.
Solvatochromic fluorophores commonly used in biological studies

Protein studies based on solvatochromism have evolved with the use of both intrinsic and extrinsic fluorescence species. For example, the fluorescence of the encoded amino acid tryptophan, has long been known to be environment dependent \cite{7, 8}, and has been widely used in folding and ligand-binding studies. However, the high-energy wavelengths required for indole excitation and the relative abundance of tryptophan in nature limits applications in complex systems, and this has prompted the design and application of extrinsic synthetic fluorophores with improved photophysical properties. These efforts have led to the emergence of a host of solvatochromic probes with diverse properties \cite{9}.

The structures and properties of some of the most common and promising solvatochromic fluorophores are summarized in Figure 1. Key physical parameters include extinction coefficients, excitation and emission wavelengths, quantum yields, size, hydrophobicity, and stability. Indeed, for specific applications, it is often challenging to identify a species that possesses all of the desired attributes. Fortunately, a number of solvatochromic fluorophores possessing overlapping properties have been developed, thereby offering researchers options in the selection of fluorophores for specific applications.

For example, the merocyanine dyes \cite{10} display exceptionally large extinction coefficients with long excitation wavelengths that are ideal for \textit{in cellulo} studies due to the minimization of damaging effects of UV light and the diminished potential for signal interference due to auto-fluorescence. However, these dyes are also large and exhibit rather subtle changes in fluorescence quantum yields and emission wavelengths in response to solvent polarity. In contrast, the PRODAN fluorophore \cite{9, 11}, is much smaller thereby reducing the likelihood that
the probe will negatively impact the native function or activity of the attached biomolecule and it exhibits emission wavelength shifts up to 100 nm. However, unlike the merocyanine derivatives, PRODAN is excited at wavelengths below 400 nm and has an appreciably smaller extinction coefficient.

The variation of the emission properties of solvatochromic fluorophores can be difficult to compare directly since most of the studies have been conducted under varying conditions and using different metrics (i.e. Lippert-Mataga plots [7] versus the $E_T(30)$ scale [12]). Typically, this is done by measuring the emission properties of the fluorophores under a range of solvents conditions. However, these studies are complicated by the fact that the solubility of many of these species is restricted to a narrow range of solvents. This limitation can be overcome by conjugating the fluorophore to another molecule to enhance solubility [13]; however the need still exists for standardized approaches for comparing the solvatochromic properties of all fluorophores. With that stated, there are a number of generalities that can be made regarding each species. For example, the dimethylaminophthalimide dyes exhibit extremely weak fluorescence in aqueous buffers [13-15] providing the advantage of low background signal until the occurrence of an event that perturbs the local environment. This creates the effect of “switch-like” fluorescence changes with the potential to produce 1000-fold increases in emission intensity [16]. The dapoxyl® dyes are noted for the ability to exhibit shifts in emission wavelength greater than 200 nm in response to changes in solvent polarity alone [17].

**Methods of incorporation into peptides and proteins**

The methods for introducing solvatochromic fluorophores into peptides and proteins are based on similar approaches for incorporating other unique functionalities such as reactive cross-
linking groups and affinity tags (e.g. biotin). However, in contrast to many other species, the insertion of solvatochromic fluorophores is topologically restricted to sites in the protein that preserve function and activity while permitting the dye to make necessary contacts that will result in measurable fluorescence changes. This consideration necessitates the use of methods that offer precise control over the dye placement. Insertion of solvatochromic fluorophores into peptides can generally be accomplished by the modification of standard solid phase synthesis approaches. For proteins, the most common methods can broadly be divided into three categories: 1) direct covalent modification; 2) incorporation of fluorescent amino acids \textit{via} semi-synthesis (expressed protein ligation); and 3) incorporation of fluorescent amino acids \textit{via} suppression of the amber (TAG) stop codon. Here, only a brief description of each approach is provided as these have been thoroughly reviewed elsewhere [9, 18, 19].

**Direct covalent modification of proteins.**

A large number of methods for site-selective chemical modification of proteins with extrinsic fluorophores have been developed [9, 18]. Cysteine and lysine reactive agents offer a convenient and direct method for labeling proteins and have been used extensively to conjugate a myriad of auxiliary groups into biomolecules. Most common among these are the thiol-selective electrophiles such as maleimides and $\alpha$-halocarbonyl compounds, along with the amine-selective acylating agents such as the $O$-succinimidyl esters (Figure 2a). In general, cysteine residues are ideal since this amino acid occurs relatively infrequently in proteins [19] and possesses excellent nucleophilic properties under most physiological conditions. Proteins with unique cysteines can be readily prepared using standard molecular biology techniques. In contrast, lysine residues are far more abundant [19] making selective site-specific labeling of recombinant proteins less practical.
Incorporation of unnatural amino acids via synthesis and semi-synthesis (expressed protein ligation)

Expressed protein ligation (EPL) [20, 21] is a powerful semisynthetic approach for incorporating non-native elements, such as solvatochromic fluorophores, site-specifically into proteins. The method involves expressing a truncated form of the target protein to which a synthetically prepared peptide is ligated at the N- or C-terminus. The synthetic peptide that is prepared constitutes the omitted portion of the native protein, but includes a fluorescent amino acid in place of another residue. The most commonly employed ligation method requires one of the two fragments to contain an N-terminal cysteine residue while the complementary fragment bears a C-terminal thioester. The N-terminal cysteine residue facilitates ligation through an initial, reversible transthioesterification step followed by an irreversible intramolecular S→N acyl transfer resulting in the formation of an amide bond (Figure 2b).

Insertion of solvatochromic fluorophores into peptides in order to yield the N-terminal cysteine-containing or C-terminal thioester peptides (as well as any peptide-based probe) is conducted either by the use of a pre-prepared fluorescent amino acid building block for SPPS or by on-resin derivatization of a peptide sequence containing orthogonally-protected amino acid with a reactive side chain. Many solvatochromic fluorophores have been integrated into unnatural amino acid building blocks for incorporation into both peptide and proteins through synthetic and semisynthetic approaches [13-15, 22-26]. The complementary expressed portion of the protein is obtained either by using a protease to reveal an N-terminal cysteine or using a “defective intein”-based approach to produce the C-terminal thioester [21, 27]. While EPL requires more synthetic manipulations and some specialized expertise, the technique offers the advantage of yielding material of greater homogeneity than that typically achieved through direct
chemical modification. In general, applications of EPL are confined to modification of amino acids within approximately 40 residues of the N- or C-termini of proteins due to the greater technical challenges involved in generating long peptide sequences or the development of three-segment ligation strategies [28].

**Incorporation of unnatural amino acids via suppression of the amber stop codon.**

Methods to expand the genetic code to include unnatural amino acids have been the source of intense study by several research groups [19]. One of the more successful approaches utilizes the amber stop codon (TAG) to encode a new amino acid [29]. Early embodiments of the method involved the semisynthesis of an artificial amino acyl tRNA (AA-tRNA) molecule designed to recognize the amber codon through base pairing of the anticodon loop. Although different methods have been applied to prepare the charged suppressor AA-tRNA\textsubscript{CUA} molecule, one of the more straightforward approaches utilizes semisynthetic approach wherein a dinucleotide (pdCpA) charged with the desired unnatural amino acid is enzymatically ligated (T4 ligase) to a tRNA\textsubscript{CUA} that is transcribed from a DNA template such that it lacks the two 3’-nucleotides. Once obtained, the artificial misacylated tRNA is introduced into a cell extract derived from *E. coli* (or in other cases from yeast or rabbit reticulocytes lysates) that is rich in the molecular machinery necessary for protein synthesis. The gene for the protein of interest, which has been mutated to incorporate the amber codon site-specifically, is then translated with the unnatural amino acid integrated into the protein at the desired position (Figure 2c). Drawbacks of the *in vitro* translation approach include the difficulty of synthesizing the aminoacylated dinucleotide (AA-pdCpA) and the overall protein yields. Furthermore, suppression efficiency can be low (20-30%) [19] and can vary widely depending on the nature of the amino acid, the gene to be
expressed, the site within the gene that the amber codon is located, and the protein expression system being used.

Currently, a great deal of attention is being focused on expanding the scope of this approach by evolving novel AA-tRNA synthetase/suppressor tRNA pairs in *E. coli* [30], yeast [31] and mammalian cells [32] that may recognize unnatural amino acids. This *in vivo* approach shows greater promise in yielding practical quantities of protein and has been utilized for the incorporation of a wide range of tyrosine derivatives. However, the unnatural amino acid of interest must first meet several criteria. It must be passively or actively transported into the host cell and lack any toxic activity. Furthermore, it should be orthogonal to the native AA-tRNA synthetases of the host such that it is not recognized as a substrate and used to misacylate one of the endogenous tRNAs. Despite these challenges, there are significant opportunities for future development of this methodology, which would provide exceptional control over the placement of solvatochromic amino acids into native proteins.

### Applications

#### Protein folding

Accurate folding is essential to protein function and valuable information regarding folding processes can be obtained by monitoring changes in the fluorescence properties of solvatochromic fluorophores incorporated into the primary sequence of a protein of interest [7]. In this case, the unfolded state would result in maximal exposure of the probe to the polar solvent environment while intermediate states or the final tertiary structure will result in increased fluorescence emission by lowering the solvent accessibility and the local polarity (Figure 3a).
**Structural information**

The sensitive emission properties of solvatochromic dyes can be used to obtain localized structural information on proteins by reporting the nature of the environment immediately surrounding the incorporation site (Figure 3b). This approach is complementary structural methods such as crystallography and NMR since it can yield detailed information regarding highly localized structural dynamics under physiologically relevant conditions or in environments that are beyond the scope of X-ray and NMR-based methods (e.g. large protein complexes, extremely dilute conditions, and with elements that exhibit significant dynamics). The method can provide information on hydration [33, 34], degree of solvent exposure or electrostatic environment of specific residues [24], and insight into interactions with membrane lipid bilayers [35]. An illustration of the potential of the approach is provided by the fluorescence mapping of a transmembrane complex involved in protein translocation across the mitochondrial inner membrane [36]. Analysis of the fluorescent properties of various NBD-labeled mutants of a specific transmembrane segment of the complex resulted in high-resolution information concerning the protein–conducting channel.

**Sensors for small molecules**

Sensors for small molecules or ions can be developed by incorporating environment-sensitive fluorophores into proteins or protein domains that naturally bind target analytes such as bacterial periplasmic binding proteins [37]. In this application, strategic placement of the chromophore into the macromolecule affords fluorescence changes that are directly coupled either with a displacement of the fluorophore by the analyte or to conformational changes that occur in response to analyte binding (Figure 3c). This approach has been utilized to develop sensors for various classes of analytes including carbohydrates (glucose [38-41] and maltose [42, 43]), ions
(nickel [44], zinc [45, 46], sulfate [47], and inorganic phosphate [48]), amino acids [49], signaling small molecules (autoinducer-2 involved in quorum sensing [50] or acyl-CoA [51]), steroids [52] and peptide-oligonucleotides [53].

Reporting conformational states

As a common regulatory mechanism, many proteins—in particular enzymes and ion channels—alternate between different stable conformations, which constitute different functional states. These modulations are typically the consequence of various signals including interactions with binding partners (e.g. protein, small molecule ligand or ion) or via post-translational modification. Solvatochromic fluorophores can be utilized to report these dynamic changes in protein structure providing direct information on protein activity (Figure 3d). Activation of proteins has thus been monitored for various calcium-binding proteins [54, 55]. Conversely, monitoring inactivated state of enzymes by this approach can be of interest for inhibitor screening. For example, acrylodan modification of a cysteine in a critical regulatory loop region of the cSrc kinase enabled development of a direct binding assay for identifying small molecule inhibitors that specifically stabilize the inactive conformation of the kinase [56].

An elegant application of this general approach has been pioneered by Isacoff with the development of voltage/patch-clamp fluorometry for the study of voltage- or ligand-gated ion channels [57-59]. By using site-directed labeling with an environment-sensitive fluorophore, subtle conformational changes in a specific region of the channels can be monitored using fluorescence in conjunction with the modulation of the currents resulting from the control of the gating state either with a voltage/patch clamp or application of a ligand. The high resolution and fast-time scale information (in comparison to the channel motions) provided by these unique...
tools, has provided new insight into the structural rearrangements that underlie channel activity [60, 61].

**Direct reporting of post-translational modifications**

Post-translational modifications such as phosphorylation can be directly reported using a proximal solvatochromic fluorophore, which reports on local changes in polarity. This approach has been utilized to monitor the activity of various kinases (myosin kinase [62] and PKC [63]) by appending an environment-sensitive fluorophore to the N-terminus of the Ser/Thr kinase substrate sequence. A complementary approach has also been reported for reporting phosphatase activity [64]. Similarly, assays for protein prenyltransferase activity have been developed by using peptides incorporating the environment-sensitive dansyl group into a cysteine-containing substrate sequence [65].

**Reporting binding interactions**

Specific protein-protein interactions can be reported by monitoring solvatochromic fluorophores positioned within or near an interaction interface, which can significantly alter the environment by the formation of new contacts between residues (*e.g.* via hydrogen bonding and/or displacement of solvent molecules) (Figure 3e,f). This approach has been widely exploited by labeling fragments or minimal protein sequences derived from one binding partner while preserving the specificity determinants required to establish the interaction (Figure 3e). While this constitutes a convenient way of designing viable probes before attempting to implement the methodology with full-length proteins, the resulting minimalized tools can also provide valuable information such as affinity constants or sensors for a particular state of one of the binding partners. Furthermore, since peptides lack the ability to form complex structural elements, the fluorophores generally remain largely exposed to the solvent environment until the
probe interacts with a partner biomolecule, thus minimizing background fluorescent signal, which might occurs when the dye is appended on globular proteins. Probes that reports binding to common protein-interaction domains have been developed for SH2 [15] and PDZ domains [66] by incorporating the solvatochromic dyes within peptides sequences derived from cognate ligands while maintaining their native specificity across different domain family members. Using the same general principle, phosphorylation of specific sequences can be monitored indirectly via the resulting interaction with a cognate phosphopeptide binding domain as illustrated with 14-3-3 [67] and SH2 domains [68]. Similarly, protein activation can be observed when the active state induces binding to other proteins. Hence by labeling a recognition domain that only binds the GTP-bound, activated form of Cdc42, the Hahn laboratory developed a sensor for monitoring the activation of the endogenous GTPase [69].

Despite the advantages of solvatochromic fluorophores over alternative more pertubating approaches, examples of the use of these tools in the context of monitoring binding of full-length proteins are still limited. This is essentially because, in addition to the generation of efficient probes, the ultimate applications are technically challenging as they involve studying these interactions in native or biologically relevant conditions. Examples of successful approaches include systems where one of the binding partners is a peptide such as for class II MHC complexes [16], opioid receptors [70] or the OppA protein of *L. lactis* [71]; for polymerizing proteins such as actin [72]; and for the development of antibody based-biosensors [73].
Design considerations when incorporating solvatochromic fluorophores into proteins

For certain applications, the precise placement of a solvatochromic fluorophore within the structure of a protein is restricted to a particular region of interest. Examples include investigations of protein-folding dynamics, where detailed information regarding the orientation of protein domains is required, as well as structural studies to map the location of solvent-exposed residues. Alternatively, the optimal placement of the solvatochromic fluorophore in a binding partner may be readily identified in that a well-defined hydrophobic interaction is already known. An excellent example of this is in the development of a series of fluorogenic probes to detect peptide loading onto class II MHC proteins involved in the activation of the adaptive immune response (Figure 4a) [16]. Crystallographic studies of the class II MHC protein, HLA-DR1, in a complex with a short peptide fragment (HA) derived from influenza revealed a pronounced hydrophobic pocket (P1) that could bind large aliphatic or aromatic side chains such as tyrosine. This position in the HA peptide chain was replaced with the fluorescent amino acids 4-\textit{N,N}-dimethylaminophthalimidoalanine (4-DAPA) [14] and 6-\textit{N,N}-dimethylamino-2,3-naphthalimidoalanine (6-DMNA) [15]. Upon binding to HLA-DR1, the peptide probes exhibited $10^2$ to $10^3$-fold increases in fluorescence intensity without perturbing the affinity or biological activity of the native complex.

Alternatively, the choice of dye placement must be determined empirically and if possible guided by structural information regarding the interaction sites of interest. However, in most cases, determining the optimal site of fluorophore placement is challenging even when high-quality structural data is available. When considering a binding event, the pockets and clefts that form the interaction interface can differ by size, shape and charge thereby accommodating some
fluorophores while limiting others. This has lead many researchers to apply screening approaches to simultaneously optimize both placement and fluorophore-type. This is particularly valuable with peptide-based ligands since the highly modular nature of solid-phase peptide synthesis, enables the rapid preparation of peptide libraries, which allows critical variables such as position, linker length and nature of the chromophore to be optimized. This approach has proven highly successful in the development of fluorogenic peptide-based sensors designed to report specific PDZ domains (Figure 4b) [66].

In applications that involve more complex interactions such as monitoring structural changes within a large protein however, the use of peptide-based probes may be inadequate. In these cases, screens can be performed by using site-directed mutagenesis to introduce uniquely-reactive cysteine residues, which can be labeled with thiol-reactive dyes. In these experiments, the goal is to identify residues that may be mutated without significantly disrupting the native structure or function of the protein. Furthermore, it is important that the appended fluorophore is positioned appropriately such that it is capable of producing a measurable signal without negatively interfering with the interaction of interest. Recently, a screening approach that produced an effective fluorescent sensor for the detection of a signaling molecule used in bacterial quorum sensing has been reported [50]. A variety of cysteine mutants were prepared from two protein receptors known to bind different forms of the bacterial autoinducer II (AI-2). The receptors, LuxP and LsrB, exhibit structural similarity and consist of two domains linked together by a hinge region with the ligand-binding site located at the interface. The cysteine residues introduced at the periphery of this site were labeled with an assortment of solvatochromic fluorophores including derivatives of dansyl, prodan, NBD, dapoxyl, and
PyMPO. One construct, a LuxP mutant (T137C) labeled with a dapoxyl derivative, was identified to possess the desired properties.

As with peptide probes, the linker length of a thiol-reactive dye also represents an essential variable to be explored when optimizing the fluorescent response of a biosensor. A systematic study conducted using the calcium binding protein calmodulin revealed that linker-type can exert a dramatic influence on both the measured fluorescence change and the degree of background emission generated by the construct (Figure 4c) [74].

**Perspectives**

Since the early protein studies that utilized the advantages of solvatochromism by relying primarily on the intrinsic fluorescence of tryptophan [7], significant progress have been achieved in this field as a result of the development of new extrinsic synthetic fluorophores with improved fluorescence properties together with major advances in the methods for site-selectively incorporating the dyes into proteins. Recent applications have shown that unique information can be obtained on the dynamics of proteins involving conformational or activity changes, which makes these environment-sensitive tools particularly useful with respect to common dyes or genetically encoded fluorescent proteins. In particular, the small size of the chromophores minimizes the perturbation induced by their incorporation, which allows studies to be conducted on quasi-native proteins. Furthermore, unlike FRET-based [1] or fluorescent protein complementation approaches [75], only a single chromophore is required, therefore providing access to systems involving multiple binding partners and to monitoring of endogenous proteins. Finally, the high sensitivity to changes in the local environment provides information at the amino acid level. However, the exploitation of solvatochromic fluorophores to study
biomolecular systems is still limited by the challenges encountered when trying to incorporate them into complex systems. In this context, with new extrinsic small chromophores that can display up to a 1000-fold increase in fluorescence [16] and methods that illustrate the viability of these tools in live cells [69], it can be anticipated that integration of the recent advances and improvement in the methods to incorporate fluorophores into proteins (as well as to deliver the labeled protein into cells) will foster studies that rely on these potent tools in cellular environment.
Acknowledgments

We apologize to the colleagues whose work we were unable to cite owing to space constraints. Research in Imperiali group is supported by the Cell Migration Consortium (GM64346) and the NSF (CHE-0414243), The support of Biotechnology Training Program (T32-GM08334) to G.S.L, and the Marie Curie postdoctoral Fellowship program to M.S. (PICK-CPP) are also gratefully acknowledged.

Additional material (to be moved after references)

Box 1. Solvatochromism

The effect of solvent polarity on the emission properties of a solvatochromic fluorophore are generalized in the Jablonski diagram, which depicts the energies of the different electronic states of the system. Beginning in the bottom left corner of the diagram, the fluorophore (large oval, grey) resides in the electronic ground state ($S^0$) surrounded by a sphere of solvent molecules (small ovals, white). The electric dipole moment of the fluorophore is indicated by the black arrow with the “+” sign at positive terminus of the dipole. Upon absorbing a photon of the appropriate energy ($h\nu\lambda$), the system is rapidly promoted to an excited singlet state ($S^1$). During this event, the system adopts a new electronic configuration with a dipole moment that differs significantly from that of the ground state. In this case, the dipole has increased in magnitude. This process of electronic excitation occurs on a time scale that is much faster than that of the motions of atomic nuclei (Frank-Condon principle [76]). In the picosecond timeframe, the molecules of the solvent sphere reorient dipoles to accommodate the now larger dipole of the
fluorophore resulting in a more highly ordered arrangement (upper left corner). This step, termed solvent relaxation, ultimately lowers the energy of the excited singlet state while simultaneously destabilizing the ground state thereby narrowing the energetic gap between the two states. When the system finally returns to the ground state through a fluorescence event, the emitted photon is of a much longer wavelength (i.e. lower energy, $h\nu_F$) than that which was originally absorbed during excitation. The degree of solvent relaxation increases with increasing solvent polarity.

In some instances, the fluorophore will return spontaneously to the ground electronic state through a thermal (non-radiative) decay process ($k_{nr}$) that competes with fluorescence. A common feature among many solvatochromic fluorophores is the tendency to exhibit a marked increase in non-radiative decay as the energy gap separating the $S^0$ and $S^1$ states is reduced. This effect is particularly apparent in polar protic solvents such as water resulting in a decrease of the fluorescence quantum yield. The mechanisms for such processes are varied and can include a range of events such as internal charge transfer (ICT), tautomerization, isomerization, and intersystem crossing (ISC) to an excited triplet state ($T^1$). When non-radiative decay competes strongly with fluorescence in polar solvents, the fluorophore can exhibit sensitive “switch-like” emission properties upon perturbations to the ordering of the solvent sphere.
Figure 1. Established solvatochromic fluorophores used in peptide and protein studies.

Figure 2. Methods for site-selective and site-specific incorporation of solvatochromic fluorophores into protein. (a) Direct labeling of a solvent exposed cysteine residue using a thiol-reactive agent [9, 18]. (b) Incorporation of an unnatural amino acid possessing a solvatochromic fluorophore as the side chain group via suppression of the amber stop codon. A tRNA molecule designed to recognize and read-through the amber codon is charged with the unnatural amino acid [19, 29]. (c) Expressed protein ligation involves a semi-synthetic approach typically requiring either the N- or C-terminal end of the protein to be prepared by solid phase peptide synthesis (SPPS). The chromophore can be inserted either as an amino acid building block (e.g. Fmoc-protected amino acid) during SPPS or after by labeling a side-chain residue with an appropriate electrophilic derivative of the chromophore (represented as a red circle). The peptides are then ligated to the portion of the protein construct that was expressed from a recombinant gene product [20, 77, 78].

Figure 3. Applications of solvatochromism for biomolecular protein studies. (a) Folding studies. (b) Mapping of a protein local environment and/or solvatation. (c) Sensors for small molecule analytes. The sensors are generally designed by exploiting protein domains that intrinsically bind a select analyte resulting in fluorescence change due to a conformational change in the protein or the displacement of the fluorophore (illustrated here with the D-glucose/D-galactose-binding protein from *E. Coli*, 2FW0 for the apo state and 2FVY for the glucose-bound state). (d) Reporting protein structural changes. In response to a signaling event, the protein of interest will undergo a conformational change that corresponds to a different functional state, which may be monitored by an appropriately positioned solvatochromic fluorophore (illustrated with calmodulin, 1DMO apo state and 1UP5 calcium-bound state). (e)
Fragment- or peptide-based probes for the monitoring of protein interactions. A minimal fragment of one of the binding partners can be labeled with a solvatochromic fluorophore to report interactions. In the case of transient interactions, the signaling event that will induce the interaction can affect either one of the partners thus providing information on function or activity. This approach is illustrated here with the Crk SH2 domain (1JU5) that binds to phosphopeptides sequences and that can be used to report either the activity of a kinase or the phosphorylated state of a substrate. (f) Reporting protein-protein interactions. A similar approach to panel (e) can be applied to full-length proteins instead of fragments. In this case advantage can be taken of the larger interaction interface between two proteins compared to peptide-based probes that do not adopt secondary and tertiary structures.

**Figure 4.** Design considerations. (a) Replacement of a conserved hydrophobic/aromatic residue by a solvatochromic fluorophore. Alignment of crystal structures of the (4-DAPA)-HA (2IPK, stick representation of the peptide in light blue with 4-DAPA highlighted in orange) and HA (1JWU, stick representation in teal) peptides bound to HLA-DR1 protein (class II MHC protein, surface representation), adapted from [16]. The replacement of the conserved ligand aromatic residue that occupies the P1 pocket of the HLA-DR protein by the 4-DAPA and 6-DMNA amino acids yielded highly efficient fluorogenic probes (over 1000-fold increase in fluorescence upon binding) without significantly affecting the specificity or affinity compared to the native interaction [16]. The crystal structures illustrate the ability of the small size 4-DAPA solvatochromic amino acid to replace the tyrosine of the native ligand. The fluorogenic probes have enabled the monitoring of *in vivo* regulation of cell-surface peptide-binding activity of class II MHC proteins in primary dendritic cells. (b) Screening for optimal positioning of the environment-sensitive fluorophore. Top: crystal structure of the third PDZ domain of PSD-95
with a modeled bound decapptide ligand derived from the PDZ domain-binding motif of Stargazin in stick representation (NTANRRTPV, adapted from 1TP3). Critical residues for PDZ domain-mediated interactions (at position 0 and -2) are represented in orange with red numbering. The 4-DMAP fluorophore was inserted systematically in each non-critical position with a diaminobutyric acid linker (*: except for -3 and -1, where a diaminopropionic acid linker was used instead). The respective fluorescence increases observed upon binding to the cognate PDZ domain are presented in the bar graph. The screening approach yielded a probe with a ~90-fold fluorescence increase by insertion of the fluorophore at position -5 [66]. (c) Screening for optimal linker length between the protein backbone and the environment-sensitive fluorophore 4-DMN. Cysteine labeling agent analogues derived from 4-DMN were incorporated into monocysteine mutants of calmodulin (illustrated with S38C and E11C mutants in the calcium-bound state, 1UP5) and compared at each position for the effect of the linker length on the ability of the solvatochromic fluorophore to report changes in its local protein environment upon binding of calcium [74].

References

32. Liu, W., et al. (2007) Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat. Methods* 4, 239-244


Figures for the manuscript entitled:

Monitoring protein interactions and dynamics with solvatochromic fluorophores

Galen S. Loving, Matthieu Sainlos and Barbara Imperiali

Box 1.
Figure 1. Established solvatochromic fluorophores used in peptide and protein studies

<table>
<thead>
<tr>
<th>Fluorophore family</th>
<th>$\varepsilon$ (cm$^{-1}$ M$^{-1}$)</th>
<th>$\lambda_{max}$ (nm)</th>
<th>MW range (g mol$^{-1}$)</th>
<th>Charge (+/-)</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRODAN</td>
<td>$2.0 \times 10^4$</td>
<td>390</td>
<td>200</td>
<td>neutral</td>
<td>[9]</td>
</tr>
<tr>
<td>NBD</td>
<td>$2.5 \times 10^4$</td>
<td>480</td>
<td>180</td>
<td>neutral</td>
<td>[9]</td>
</tr>
<tr>
<td>Coumarin derivatives</td>
<td>$2.4 \times 10^4$ – $5.0 \times 10^4$</td>
<td>376 – 419</td>
<td>160 – 215</td>
<td>neutral</td>
<td>[9]</td>
</tr>
<tr>
<td>Dimethylaminophthalimides</td>
<td>$6.5 \times 10^3$ – $8.8 \times 10^3$</td>
<td>378 – 408</td>
<td>190 – 240</td>
<td>neutral</td>
<td>[13,14,15]</td>
</tr>
<tr>
<td>Merocyanines</td>
<td>$1.1 \times 10^3$ – $2.2 \times 10^3$</td>
<td>570 – 605</td>
<td>440 – 500</td>
<td>+ / – neutral</td>
<td>[10]</td>
</tr>
<tr>
<td>Dapoxyl* derivatives</td>
<td>$2.7 \times 10^4$</td>
<td>366 – 370</td>
<td>280 – 330</td>
<td>neutral</td>
<td>[17]</td>
</tr>
<tr>
<td>PyMPO</td>
<td>$2.3 \times 10^4$</td>
<td>412</td>
<td>250</td>
<td>+</td>
<td>[9]</td>
</tr>
</tbody>
</table>
Figure 2. Methods of incorporation

a) Direct conjugation to cysteine mutant

b) Translational incorporation via nonnatural amino acid

c) Semi-synthetic incorporation (N-terminus)

Semi-synthetic incorporation (C-terminus)
Figure 3. Applications of solvatochromic fluorophores for protein studies

a) unfolded ↔ folded
b) solvent exposed ↔ buried

c) open/unbound → closed/bound
d) inactive conformation ↔ active conformation

e) phosphorylated ↔ non-phosphorylated
f) binding event ↔ binding event
Figure 4. Design considerations

a) Design considerations

b) Design considerations

c) Design considerations

Fluorescence fold increase

Position of the solvatochromic fluorophore

Linker and fluorophore:

Linker length

E11C

S38C

4-DMAP

P-2 P-1 P2 P3 P4 P6 P8 tyrosine

4-DMAP

3 Å 6 Å (1) 6 Å (2) 9 Å

3 Å 6 Å (1) 6 Å (2) 9 Å