Abstract Imaging of biochemical processes in living cells and organisms is essential for understanding how genes and gene products work together in space and time and in health and disease. Such imaging depends crucially on indicator molecules designed to maximize sensitivity and specificity. These molecules can be entirely synthetic, entirely genetically encoded macromolecules, or hybrid combinations, each approach having its own pros and cons. Recent examples from the author's laboratory include peptides whose uptake into cells is triggered by proteases typical of tumors, monomeric red fluorescent proteins and biarsenical–tetracysteine systems for determining the age and electron-microscopic location of proteins.

© 2004 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Molecular imaging; Fluorescence resonance energy transfer; Fluorescent protein; Polyarginine; Biarsenical; Tetracysteine

1. Introduction

The 20th century witnessed explosive progress in macromolecular biochemistry and genetics, which started with the rediscovery of Mendelian genetics and the recognition of biopolymers and culminated in the sequencing of complete genomes. However, genome sequences alone lack spatial and temporal information and are therefore as dynamic and informative as census lists or telephone directories. The challenge for the 21st century is to figure out how these casts of molecular characters work together to make living cells and organisms and how such understanding can be harnessed to improve health and wellbeing. I believe this quest will depend heavily on molecular imaging, which shows when and where genetically or biochemically defined molecules, signals or processes appear, interact and disappear, in time and space. Therefore, molecular imaging synergistically draws upon physics, chemistry, anatomy, physiology, biochemistry and genetics.

Our first significant contribution to molecular imaging was a series of organic chemical buffers and indicators (reviewed in [1,2]) for intracellular calcium (Ca\(^{2+}\)), which is a crucially important messenger inside cells. We also synthesized novel molecules to measure other intracellular signals such as sodium [3] and proton concentrations [4], gene expression [5] and membrane potential [6–8].

In the mid-1990s, we began to put most of our effort into developing genetically encoded macromolecular indicators, which are usually administered not as ready-made proteins but as genes, which tell the cell or organism to make the indicators to our specifications. The key building blocks for such genetically encoded indicators are mutants of the green fluorescent protein (GFP) [9] from the jellyfish *Aequorea victoria*. GFP was discovered as a protein by Shimomura [10], the gene was cloned by Prasher [11], and Chalfie’s and Tsuji’s labs [12,13] first reported heterologous expression. GFP and its relatives have become tremendously useful in many areas of molecular and cell biology because they provided the first means by which a simple gene could give rise to bright visible fluorescence. Whenever scientists want to make a cellular protein fluorescent, their first thought nowadays is to fuse the gene for their favorite protein to the gene for a fluorescent protein (FP), then put this composite gene back into the cell or organism of interest. If all goes well, the hybrid generates a chimeric protein in which the host component goes about its normal business, while the attached FP fluoresces and reports the presence and position of the pair [9] (see Fig. 1).

Dr. Roger Heim in my lab started working in 1992 on the GFP gene provided by Prasher [11]. We very much wanted to create mutants with brighter fluorescence, because the original GFP was dim, fickle and spectrally impure. We also wanted different colors to enable fluorescence resonance energy transfer (FRET), which inherently requires a pair of colors. Heim developed the first mutants in which the spectrum of GFP was simplified and enhanced [14]; one of these mutations (S65T), right next to the chromophore, is at the heart of the optimized GFPs now routinely used around the world. Heim also created blue and cyan-emitting mutants, BFPs and cyan fluorescent protein (CFPs), respectively [15,16]. Later we helped Prof. Jim Remington’s group to solve the X-ray crystallographic structure of GFP, which immediately suggested a...
way to push the emission to somewhat longer wavelengths [17]. Andrew Cubitt made the appropriate mutants, whose yellowish appearance led to the name of yellow fluorescent protein (YFPs). Currently, the best pair for FRET consists of the cyan and the yellow mutants, CFP and YFP, respectively.

2. Recent progress (1999–2004) in understanding and improving fluorescent proteins

For most applications of YFPs, it is important to minimize their sensitivity to pH, [Cl\(^{-}\)] and photo bleaching. We found a mutant of YFP (Q69M) with greater resistance to these factors and solved the crystal structure, which showed that the Met filled a cavity in previous YFPs [18]. We destroyed the weak dimerization of CFP/YFP to probe loose association of membrane-anchored versions in or out of lipid rafts [19].

The cloning of the first red fluorescent protein, DsRed [20], revealed nothing about its post-translational biochemistry. We showed that DsRed is an obligate tetramer, which first becomes green-fluorescent before finally maturing slowly and incompletely to red [21]. Such tetramerization explained why most attempts to fuse DsRed to other proteins produce mistargeting or toxicity. Mass spectrometry told us the covalent structure of the chromophore, formed from the GFP chromophore by an unprecedented dehydrogenation of a C\(\alpha\)–N bond [22]. X-ray crystallographers [23,24] confirmed the structure by observing the subtle change of that C\(\alpha\) from tetrahedral to trigonal. The intimacy of the intersubunit contacts led to predictions that the tetramer would be impossible to break [25], but we succeeded [26]. The resulting monomeric red fluorescent protein, “mRFP1” has become very popular for many in vivo applications [27–30].

Although others have discovered coral fluorescent proteins of varied colors, all have been wild-type tetramers. Also improvements in mRFP1 brightness and photostability would be desirable. Therefore, Nathan Shaner has evolved mRFP1 into brighter monomers ranging in emission between 534 and 638 nm. For example, “mCherry” matures more quickly and completely than mRFP1, giving higher extinction coefficient and brightness, yet bleaches about 10-fold more slowly [77].

Lei Wang developed an alternative method for evolving novel protein properties directly in mammalian cells by harnessing somatic hypermutation, the process by which B lymphocytes constitutively mutate genes for immunoglobulin or other highly expressed proteins up to 10\(^6\) faster than normal [31]. mRFP1 was stably transfected into the B-cell line Ramos with expression controllable by doxycycline. Cells with progressively red-shifted fluorescence emissions were selected by fluorescence-activated cell sorting. After over 23 generations of selection, the emission peak gradually shifted to 648 nm, 36 nm longer than parental mRFP1 and beyond any found by semi-rational in vitro mutagenesis [78].

3. Fluorescent-protein-based indicators

**Redox.** James Remington’s lab invented redox indicators consisting of GFP mutants with cysteine pairs, whose reversible oxidation to a disulfide profoundly shifts the excitation spectrum to shorter wavelengths, enabling ratiometric measurement [32]. Colette Dooley [33] characterized and applied these indicators in live cells, targeted them to cytosol, plasma membrane and nucleus and increased their oxidation sensitivity by placing positive charges next to the cysteines to lower the p\(K_a\) of the latter. The redox indicators become substantially
oxidized in response to exogenous oxidants or oxidative bursts in HL60 cells. Despite considerable effort, we have not found responses to growth factor stimulation, suggesting that if reactive oxygen species (ROS) are as important in mitogenic signaling as widely believed, such ROS must be very transient or localized.

Ca2+. Geoff Baird showed [18,34] that GFPs could be circularly permuted and that conformationally sensitive receptor domains could be inserted within the GFP to generate indicators (e.g., ‘camgaroos’) with large-amplitude responses to binding of ligands such as Ca2+. Camgaroos have provided useful signals in *Drosophila* [35] and inspired higher-affinity Ca2+ indicators such as “G-CaMP” and “pericams” from other groups [36,37], which have become some of the most promising genetically encoded indicators of neuronal network activity [38,39].

Our FRET-based sensors of Ca2+, the “cameleons” [40–42], remain advantageous for their emission of ratiometric response and avoidance of the intrinsic pH-sensitivity of the camgaroo/G-CaMP/pericam family. Cameleons have revealed important biology in plants [43,44] and nematodes [45–47]. Transgenic mice expressing cameleons were the testbed for a new combinatorial chemistry strategy to disrupt PKA anchoring [54], largely mimics insulin’s preferential inhibition of β-adrenergic activation of PKA. We hypothesize that chronic insulin pretreatment disrupts preferential coupling of β-adrenergic receptors to localized cAMP and PKA, which would exemplify a novel mechanism for heterologous crosstalk between signal transduction pathways.

**Glutamate.** Optical indicators of glutamate would be useful to image glutamate without concerns of up/downregulation or desensitization of endogenous receptors and to map the spatial spread of glutamate. Andrew Hires has created glutamate-sensitive fluorescent reporters (GluSnFRs) by fusing CFPS, GtII (a periplasmic glutamate-binding protein from *E. coli*) and YFP. Seminal modifications vary the *Kd* for glutamate from 0.15 to 700 μM. When GluSnFR is expressed on cultured hippocampal neurons but not localized at synapses, he and Yongheng Zhu in Charles Stevens’ lab can image electrically evoked [Glu] transients, which vary with extracellular Ca2+ and Mg2+ as expected for synaptic release and are amplified 3-fold by a glutamate reuptake inhibitor. Averaging across the dendritic surface shows responses to single action potentials, peaking within 50 ms and returning to baseline within 200 ms. Averaging across multiple stimuli shows some dendritic hotspots that respond most strongly to the first action potential of a short spike train, while other hotspots prefer the second, consistent with heterogeneity of paired-pulse facilitation vs. depression across synaptic sites.

4. Biarsenical–tetracysteine labeling system

Although FPAs have enormously advanced molecular and cell biology, they do have important limitations. They are at least 220 amino acids in size, so that they are often bigger than the host protein to which they are fused. Their only basic function is fluorescence, whereas artificial small molecule probes encompass a much wider range of useful properties. Ideally, one would like to combine the genetic targetability of fluorescent proteins with the versatility and small size of organically synthesized dyes. A solution to this challenge is to tag the protein of interest with a very small peptide domain, whose essential core is just two cysteines, two spacer amino acids

activation of Src in endothelial cells touching adherent beads pulled by laser tweezers.

Jin Zhang’s [51] original PKA reporter AKAR1 had given mostly irreversible responses in cells. Replacing the 14-3-3 domain, which binds phosphoserine very tightly, with an FHA1 domain, which binds phosphothreonine less tightly, generated a kinase activity reporter, 2nd version (AKAR2), which is quite reversible. Our foremost application of AKAR2 is to explore insulin antagonism of PKA in adipocytes, which has potential significance for obesity. Jerrold Olefsky’s lab, our collaborators, had shown [53] that chronic insulin exposure enhances cAMP generation by β2-adrenergic stimulation but had no effect on cAMP elicited by forskolin, a direct agonist for adenylate cyclase. We now find that chronic insulin pretreatment of 3T3-L1 adipocytes delays PKA activation by β-adrenergic stimulation despite the enhanced cAMP production. The divergence between total cAMP vs. PKA activity is seen only with β-adrenergic stimulation, not with other ways of raising cAMP, e.g., forskolin or photolysis of caged cAMP. Endogenous CREB phosphorylation agrees qualitatively with the AKAR2 responses. Ht31, a dominant-negative peptide known to disrupt PKA anchoring [54], largely mimics insulin’s preferential inhibition of β-adrenergic activation of PKA. We hypothesize that chronic insulin pretreatment disrupts preferential coupling of β-adrenergic receptors to localized cAMP and PKA, which would exemplify a novel mechanism for heterologous crosstalk between signal transduction pathways.

**Kinase activities.** We created indicators for kinases such as epidermal growth factor receptor, Src, Abl [50], protein kinase A (PKA) [51] and protein kinase C [52]. Generically, these indicators fuse CFP, a phosphoaminoacid-binding domain, a substrate peptide and YFP. Peter Wang, a postdoc bridging Shu Chien’s and my lab, has improved the specificity of the Src reporter by replacing its original substrate peptide by one derived from p130cas. He is seeing interesting localized
specific proteins [61–64]. This motif can be labeled in living intact cells with membrane-permeant dyes that carry two arsenic atoms along one edge. Each pair of cysteines binds to one of the arsenics, probably causing the peptide to bend into a hairpin conformation. We have synthesized a variety of bisarсенical dyes that can fit into the same peptide domain, but the most useful are a green-emitting fluorescein-based dye called FLAsH (for fluorescein-based arsenical hairpin binder) and a red-emitting resorufin-based analog called ReAsH (for resorufin-based arsenical hairpin binder). FLAsH and ReAsH are both able to cross mammalian cell membranes. They are delivered pre-bound to an antidote, 1,2-ethanedithiol, that minimizes toxicity by protecting endogenous thiols from the arsencis. The antidote also keeps the dyes relatively non-fluorescent until they find their target tetracysteine domains. A single tetracysteine site can bind either FLAsH or ReAsH with picomolar affinity, but not both at the same time. Once bound, the dyes remain stably attached for days under normal conditions.

Pulse-chase labeling and singlet-oxygen-(1O2)-mediated electron microscopic photoconversion of tetracysteine-tagged connexin-43 revealed the life cycle of this major constituent of gap junctions [58]. Analogous pulse-chase labeling of neuronal AMPA receptors (GluR1 and GluR2) gave many insights into their regulation and trafficking, including direct evidence that they can be synthesized de novo in microsurgically isolated dendrites [59]. Protein folding and aggregation have been monitored in real time within fairly intact bacteria [60]. Intense illumination of FLAsH or ReAsH can inactivate the tetracysteine-tagged protein, providing a genetically targeted strategy for sudden (and potentially localized) functional knockout of neuronal biomolecules.

Improved tetracysteine sequences. Brent Martin has completed a FACS-based genetic screen to identify the optimal X1–6 surrounding the Pro-Gly tetracysteine motif, i.e., X1X2X3CCPGCCX4X5X6 fused to the N-terminus of GFP. A mammalian cell-based library of >10^7 clones was screened based on ReAsH binding and fluorescence in the face of increasing levels of dithiol antidote. After 16 rounds of selection by FACS, several superior sequences emerged, including our present favorite, FLNCCPGCCMEP. The new motif expresses more efficiently, exhibits ~3 orders of magnitude increase in affinity as judged by resistance to dithiols and demonstrates a ReAsH quantum efficiency of 0.47, about 2-fold higher than earlier peptides. The new tetracysteine peptide displays >25-fold improvement in contrast over previously published peptides in living cells, enhancing the contrast to levels approaching that of GFP. These improvements have been transferable to either terminus of GFP as well as to direct fusions to β-actin and α-tubulin. The new tetracysteine motif significantly reduces background and enhances sensitivity, which should enable imaging or photoactivation of a much broader spectrum of cellular proteins.

New applications. Oded Tour and Stephen Adams have bound a bisarsenical, low Ca^{2+}-affinity derivative of Calcium Green to tetracysteine-tagged voltage-gated Ca^{2+} channels, as an alternative to cameleon fusions, to sample the nanometer-sized local [Ca^{2+}] domains at the channel mouths. Total internal reflection microscopy excludes background signals including channels that are not yet inserted into the plasma membrane. Preliminary indications are that the local Ca^{2+} concentrations are not as high as widely supposed from mod-

eling studies and that channels clustered in patchy hot spots rather than randomly distributed.

5. Aptamer labeling

Jeremy Babendure discovered that two series of RNA aptamers (38–54 nucleotides in length) developed for other purposes [65, 66] could each bind a different series of poorly fluorescent dyes with 0.12–86 µM K_d's while enhancing the fluorescence by >2000 and ~100-fold, respectively [67]. Such aptamers would constitute compact fusion tags for intracellular RNAs except that cells unfortunately contain other constituents that enhance the fluorescence of these triphenylmethane dyes. Stephen Adams made related dyes that show extremely little fluorescence in cells, but those dyes would require new aptamers.

6. Synaptic plasticity

Varda Lev-Ram discovered a new form of long-term potentiation (LTP) at the parallel fiber (PF) to Purkinje cell (PC) synapse, elicited by 1 Hz PF stimulation without a rise in PC [Ca^{2+}]_cyto. Unlike the previously known cAMP-mediated presynaptic form of LTP [68], the new LTP is postsynaptic, nitric oxide-mediated and a mutually reciprocal resetting mechanism for classic coincidence-dependent long-term depression [69, 70].

7. Tumor imaging

For clinical applications, one would prefer not to have to introduce genes or be limited to optical detection. Arginine-rich sequences are known [71–76] to mediate uptake of a wide variety of contrast agents into cells and tissues in vivo. Tao Jiang found that such uptake can be prevented by appending...
References