Thank you very much for this honor that you have given me and for putting on such a nice symposium. It covered so many topics, ranging from organic synthesis, including tert-butyllithium mentioned in the talk of Professor Urano, all the way up to clinical medicine, with a very lively discussion at both ends and everything in between. So it is splendid to see such interdisciplinary efforts and the enthusiastic involvement of young people, not just old folks like me. Of course the young people represent the future of our subject. The students in the audience are the real stem cells of biology who will keep our subject alive. Professors like me with grey hair are pretty much set in our ways, or terminally differentiated. There are only two possible long-term fates for terminally differentiated cells — necrosis or apoptosis — take your pick. So the students are the ones who will carry on the tradition.

I have chosen three diverse topics to cover, again knowing that different people in the audience will have different interests. Some of these were hinted at yesterday, but I will try to go into a little more detail today, though I apologize for the repetition. One topic is how we are now progressing with red fluorescent proteins, the work primarily of Nathan Shaner and Lei Wang. I will also update you on our hybrid system based on arsenic in which we can combine molecular biological targeting via tetracysteines with organic chemical versatility. The final area is our effort to transition into in vivo imaging of tumors via their local proteolytic activities. We hope soon to be lessening our dependence on optics and to begin working with magnetic resonance imaging and nuclear medicine, so that this approach may eventually have some clinical application.

As I introduced yesterday, genetically targeted labels have the advantages that they are made by the organism itself under the control of the DNA that we put in. Therefore these reporters can be continuously made for the life of the organism, transmitted to subsequent generations, and very precisely targeted to particular cells or subcellular locations using cell specific promoters or targeting sequences. Genetic encoding also means we can exploit a wide variety of natural proteins as building blocks for sensors. Any protein that undergoes a controllable conformational change or protein-protein association or dissociation gives us a chance to generate an indicator out of that system. Thus genetically encoded indicators have been demonstrated for many important signals, including membrane potential, intracellular calcium, cyclic nucleotides, and redox status at least at the level of thiols and disulfides, although maybe not quite the same electron transport chain as we just heard about from Dr. Fujibayashi. Other indicators monitor exocytosis and endocytosis, neurotransmitters, and I will briefly mention protein conformational changes. I will not have time to talk about reporters for kinase and phosphatase activities, though they are a very interesting subject on their own. I will briefly mention imaging of protein turnover and trafficking, and lastly protease activation, though the latter is with small molecular probes.

The red fluorescent protein DsRed from corals, first cloned by Matz et al., has attracted much attention as a label providing contrast with jellyfish green fluorescent protein (GFP) and as a route to longer wavelengths that could penetrate tissues further. The problem with DsRed however, is that it was made by the coral for its own purposes, which are not the same as our purposes. We assume the coral just wanted to color its cells red, though even now we do not really know exactly why. This is a
favorite topic of argument whenever people who work on fluorescent proteins get together. We try to speculate why the corals and the jellyfish have these proteins, but we cannot “read their minds”. We are only grateful that they evolved such proteins. The corals, however, probably want to maintain very high photostability under tropical sunshine, which may be the reason why all the coral proteins in their native form have turned out to be tight tetramers. Such tetramerization is bad for us as biological users, because if we fuse such a protein to our protein of interest, the latter is dragged into the tetrameric complex, which can be very messy and ungainly. Thus if our protein participates in its own protein-protein interactions, there is the potential for a massive cross-linking. An example is the gap junctional protein connexin-43, which has to form a hexamer in one cell to combine with a hexamer from another cell to make the functional gap junction. If one fuses connexin 43 to green fluorescent protein (GFP), the jellyfish protein which only weakly dimerizes, one can get functional gap junctions, which correctly traffic to the border between two cells. By contrast, if we make the same fusion to the original coral protein DsRed, the fusion just precipitates around the Golgi apparatus as a mess of insoluble aggregates, never reaching the plasma membrane or forming gap junctions.\(^2\) When crystal structures of DsRed were solved,\(^3,4\) they revealed so much buried surface area in the tetramer interfaces that disruption would be very difficult. Indeed, attempts to make mutations at the interface destroyed the fluorescence.\(^5\) Other problems of DsRed are that it is slow to mature and always retains a major fraction of the green component, which is annoying when we want to do two color staining. If our red fluorescent protein (RFP) always retains some green, then we cannot independently detect a low amount of GFP.

These problems were solved stepwise. The first improvement came from the laboratory of Ben Glick,\(^6\) who evolved a mutant that matured more quickly but still incompletely, and was still tetrameric. Rob Campbell in my lab then devised a form of directed evolution where he introduced charged residues to disrupt one interface at a time, then randomly generated additional mutations to restore at least a small amount of fluorescence, which is easy to screen for visually. He then made new sets of mutations iteratively, about eleven rounds of evolution in as many months, eventually arriving at 33 mutations or 15% of the protein (Fig. 1). The resulting monomeric red fluorescent protein\(^2\) is the most popular fluorescent protein my lab has ever made; every day I get another few requests by email. Indeed the main thing my secretary does these days is to fulfill those requests.

Now, we and many other people still want more colors. There are two obvious choices. One is the very fundamental approach, as Dr. Miyawaki described earlier in this symposium, to go somewhere like Okinawa with lots of corals, and clone proteins with new colors. This
is the fundamental way to get true diversity, because the corals have had several hundred million years to figure out lots of variations. Unfortunately, they are almost always tetramers, leaving a lot of hard work to make them individually into monomers. Therefore we decided to try the alternative strategy, to see if we could change the color of our existing monomer. Such wavelength shifting proved surprisingly easy and produced a rainbow of yellow to deep red colors (Fig. 2), which we named according to the fruits with matching colors. There is no one best protein amongst these, just like there is no one best fruit in the grocery store that is so good that you do not have to buy any other fruits. Depending on your tastes and requirements, different proteins have different pluses and minuses. For example, “mOrange” has a particularly high quantum yield as a monomer, which makes it a good fluorescence resonance energy transfer (FRET) acceptor. However, it is markedly quenched by pH’s lower than 6.7. This acid sensitivity is undesirable for using mOrange as a simple marker in the presence of pH variations, but allows mOrange targeted to the acidic lumen of secretory granules to report exocytosis into the neutral extracellular environment. Some pH sensitive GFPs already do the same, but mOrange works at much longer wavelengths, potentially allowing monitoring of two separate types of exocytotic events or of exocytosis in the presence of FRET from cyan to yellow fluorescent proteins.

Tandem dimer tomato (“tdTomato) has the highest overall brightness due to highest extinction and an even higher quantum yield, and is also quite photostable. The main disadvantage is that it is twice the molecular weight, because here one of the dimer interfaces is still intact. But tdTomato is not a dimerizer because we fused the N and C termini of the two sub-units together so that this dimer is satisfied inside itself as an internal combination. And by the way, if anyone objects that tomatoes are not fruits, you are thinking like a grocer rather than a biologist.

The closest successor to mRFP1 is “mCherry”. It has a higher extinction, faster maturing, and much greater indifference to fusion. (mRFP1 sometimes lost considerable brightness when other proteins were fused to its N-terminus.) Most importantly, mCherry is much more photo-stable, withstanding about an order of magnitude more photon dose than mRFP1 before bleaching to 50%.

We also discovered a variant, “mGrape2”, which proved photoisomerizable in a manner somewhat similar...
to “Dronpa” mentioned by Dr. Miyawaki earlier in this symposium. However, mGrape2 needs blue rather than violet light for photosomerization, and then changes from orange to red emission. Furthermore, over several hours in the dark, it spontaneously reverts back to its original color, rather than requiring a second wavelength of light to push it backwards.

Finally, mPlum is our longest wavelength fluorescent protein to date. Longer wavelengths are of interest because mammalian tissue becomes more transparent at longer wavelengths, especially beyond 600 nm. Near-infrared wavelengths would be best of all. The emission of mRFP1 peaks around 580 nm, whereas mPlum’s maximum is at 648 nm, more than halfway to the true infrared (700 nm or beyond). mPlum is particularly interesting because it was evolved by a new procedure that I think may have more general applicability. Up to now, our mutations have been made by this rather laborious procedure of making libraries of mutations in vitro, and then putting that DNA back into, say, bacteria, growing up the protein, selecting the phenotype, picking the best new property, whether it is a brighter fluorescence or a longer wavelength or whatever, and then growing up the colonies, sequencing the DNA, and figuring out which mutation actually made the best change. Then we have to think and design our next set of mutations. The entire cycle is quite laborious. We thought that a system in which cells themselves would make mutations for us in situ would be very valuable as a complementary alternative.

Until now, if one wants to make mutations within mammalian cells, typically one would shine ultraviolet light on them, irradiate them with X-rays, or apply chemical mutagens like nitrous acid or ethyl methanesulfonate (EMS). These methods of in situ mutagenesis are very random and do not target the gene you care about. They hit all genes equally, so that if you mutate at a high rate, you destroy the viability of the cells, just like strong ultraviolet or X-ray energy will sterilize the tissue if you use too high an energy density. What we really want is a means to focus the mutations just on say, mRFP1 and leave the housekeeping genes of the cell intact.

As I am not an expert molecular geneticist, I asked more knowledgeable friends “Is there some special genetic system where you can put some cis-acting elements around one particular gene to make that gene mutate rapidly?”, but they did not know of such a system. Then I looked a little more carefully in the literature and eventually realized that we all “know” how to do such directed mutation, except that for most of us it is an unconscious process that only our B lymphocytes understand how to accomplish, even if our scientific minds do not.

B lymphocytes use a process called somatic hypermutation (SHM) at the last stage of refining an antibody. These cells have figured out a special system for increasing the mutation rate 1 million-fold over the background, mostly targeted to the immunoglobulin locus. It has certain biological rules; it is not perfectly random. The mutation process starts with an enzyme called activation induced deaminase (AID), which attacks cytosine. This primary alteration is then amplified or accentuated by error-prone DNA repair. The fine details of this process are under intensive investigation in immunology labs, although I am not an expert on it. Our interest was to exploit this procedure, and it was crucial to know that it is actually not totally specific for immunoglobulins. You can get other genes to mutate if you transcribe them at a high level. However, what other people had done with non-immunoglobulin genes was to take GFP, our familiar workhorse, make a point mutation so that it is now unable to express, put this into B lymphocytes and show that at some low frequency the B lymphocytes repair the stop codon, which is simply fixing an artificial mistake without generating anything novel and useful.

We therefore decided to take mRFP1 and put it into Ramos cells, a line of B lymphocytes that is known to conduct SHM, under the control of a tetracycline promoter so that we could turn the mutational process on and off via the level of transcription. The basic procedure was then simply to grow up millions of cells, then switch on transcription with the tetracycline analog doxycycline to make lots of mutations. Each cell on average probably makes one mutation — a different mutation — somewhere in the mRFP1 gene. We used fluorescence activated cell sorting to select around 5% of the cells with the highest ratio of far-red to near-red fluorescences and reasonable brightness, rather like examining a class of medical students and picking those top 5% with the highest grades. Then we simply grow the cells up again, add doxycycline again and sort again for the best 5%. This simple cycle of growing and sorting is much less laborious and can thus be done twice a week, not once a month.

The fluorescence wavelengths initially did not improve much from generation to generation (Fig. 3), which made us fear that we were pursuing a hopeless cause. Fortunately, however, we kept going, and in round 10 or so things really began to improve markedly. By round 16 or 20, there were some major wavelength improvements, which subsequently leveled off, after which we decided to write the paper. Curiously, the excitation spectrum first shifted to the right and reached its longest wavelengths at about round 10 in a protein we call “mRaspberry”, and then shifted backwards in the final evolution to round 23 even as the emission maximum got longer and longer. Therefore the Stokes shift between the emission and the excitation peaks became particularly large. The best clone out of round 23 is called “mPlum”. It is about 8 to 10-fold more photostable than the starting mRFP1. Over the 23 generations of SHM, it accumulated
7 mutations, which we would not have chosen rationally. The rational strategy to increase emission wavelengths would be to mimic the way we made yellow fluorescent protein (YFP) out of GFP, which is to stack a tyrosine over the chromophore. This strategy produced mGrape, whereas SHM found a different and more effective strategy. Thus exogenously introduced non-immunoglobulin genes can undergo useful mutations distributed throughout the sequence.

At each of the seven positions mutated, you could imagine 18 other amino acids other than the wild type and the mutation picked already by SHM. Are any of these 18 better than SHM’s choice? So far, never has that been the case, which implies that SHM has already done a fairly exhaustive search there. Another proof of how much sequence space SHM has explored comes from examining a genetic element left over from the virus. This is a “long terminal repeat (LTR)”, which is past the end of the protein-encoding region and therefore not under any phenotypic selection pressure. Almost 13% of the LTR has been altered, again reflecting the extensive mutagenesis.

The first use of the immune system in biotechnology was to generate polyclonal antibodies. Then Milstein and Kohler invented monoclonal antibodies, which finally made antibodies into molecularly defined, reproducible reagents. Schultz and Lerner and many other people have worked on catalytic antibodies, which act somewhat like enzymes to mediate chemical reactions. But even so, all of these are still antibodies. Now we see that the immune system can evolve proteins that have no relationship to antibodies as long as you have a high throughput single cell screen applied to a somatically hypermutating cell line. I believe that if you want to engineer a new protein such as a receptor with mutant properties or a new enzyme, as long as you have a screen that works at the single cell level and leaves the best performers alive, you have a good chance of using this rich source of diversity as an efficient means of generating the property you want.

Now I switch focus from fully genetically encodable systems to a hybrid combining the advantages of molecular biology and organic chemistry. The virtue of the molecular biology is that it gives you very accurate targeting through genetically encodable elements, but the advantage of organic chemistry is that we can put in un-
usual properties with a much smaller structure than you need with a fluorescent protein. All fluorescent proteins that generate visible chromophores from their own internal amino acids are at least 220 or so amino acids. Nobody has yet found a significantly smaller one yet. Here we can reduce the genetically targeted portion ultimately down to as small as six, and we can bring in properties that no organism has ever yet had to evolve. For example, no organism has ever been interested in helping us with magnetic resonance imaging (MRI). Incorporate a gadolinium, why would any creature evolve that? MRI was only invented two or so decades ago. So, we still need the flexibility of organic synthesis.

Albert Griffin was a brave organic chemistry student in my lab who came up with an initial solution (Fig. 4) based on fluorescein, on which he put two arsenics by a simple transmetallation to create a molecule called “FlAsH.” We picked arsenic because it has affinity for pairs of thiols inside cells, but a single pair is not specific enough. By putting two arsenics in, we create selectivity for a quadruple group of cysteines as provided in the motif Cys-Cys-Pro-Gly-Cys-Cys, where the Pro-Gly probably form a hairpin turn to allow the four cysteines to cluster together. The biarsenical derivative of fluorescein (“FlAsH”) is membrane permeant. Its effective dissociation constant from the tetracysteine turned out to be on the order of a few picomolar. As an extra bonus, FlAsH is non-fluorescent when bound to an antidote that keeps it away from ordinary pairs of cysteines, probably due to photo-induced electron transfer, the same quenching mechanism Dr. Urano discussed earlier in this symposium. When bound to the tetracysteine, FlAsH becomes brightly fluorescent, probably because steric constraints prevent the coplanarity necessary for photo-induced electron transfer. We now have new colors such as ReAsH which is isosteric to the fluorescein, but red instead of green. FlAsH and ReAsH have to be made by organic synthesis of course, so most biologists will not want to make them, but they are now commercially available from Invitrogen. Unfortunately the company insisted on changing their names to “Lumio Green” and “Lumio Red” for reasons I do not understand.

Fig. 5 summarizes an example where the small size of this genetically targeted fluorescent label is crucial. This comes from recent work from Carsten Hoffman in Martin Lohse’s lab. They have been looking at conformational changes in G-protein coupled receptors, the family with seven transmembrane segments, which are so important in many signaling pathways. They originally put YFP at the C-terminus and CFP into intracellular loop 3 of the adenosine receptor. The resulting hybrid molecule responded to adenosine with a small but rapid loss of FRET, within 80-100 milliseconds, which they could show was the conformational change in the G-protein.
coupled receptor upon binding with adenosine. Unfortunately, the inclusion of the CFP (>230 amino acids) almost completely destroyed the ability of this receptor to activate downstream signaling. The native adenosine receptor will stimulate adenylyl cyclase when it binds adenosine, but the hybrid molecule with two bulky fluorescent proteins has lost this important ability. But when Hoffman replaced the CFP with this much smaller insert, just six amino acids (CCPGCC) in intracellular loop 3, then labeled that motif in live cells with FlAsH, then he observed an even bigger FRET response to adenosine and restored completely normal activation of adenylyl cyclase. In this case, smaller is better.

I will skip past our work with Mark Ellisman’s group on connexin subunits and mention that various other groups, including that of Robert Malenka, have started applying the ability of FlAsH and ReAsH to sequentially label the same protein, or the same type of protein. Here they put in the tetracysteine motif near, but not quite at, the C-terminus of the AMPA form of the glutamate receptor. This type of glutamate receptor in the brain is of intense interest to neurobiologists because it is believed to mediate several forms of synaptic plasticity, so that when you have more AMPA type glutamate receptors in the membrane, your synapse is more sensitive, and when you take AMPA receptors away, then your synapse becomes less sensitive. It becomes of great interest to know the local trafficking of the AMPA receptor — in particular, can neuronal dendrites synthesize fresh AMPA receptors on their own without calling upon the cell nucleus for help?

Normally biologists believe that to increase the production of new proteins, the cell must send a signal to the nucleus to make new messenger ribonucleic acid, which then encodes the synthesis of new protein. But in a neuron with thousands of different synapses, which must be independently regulated, it is very cumbersome if they all have to talk to the nucleus, because how does the nucleus know how to send the protein to one particular synapse out of the thousands? People have thus been very interested in the possibility that individual synapses, or at least parts of neuronal dendrites, have autonomous control of protein synthesis. To test this idea, the Malenka group (Ju et al) inserted a tetracysteine motif into the AMPA receptor, then labeled it with ReAsH, the red dye, so that all the copies of the AMPA receptor made

\[\text{Fig. 5 FLAsH-tetracysteine insertion into GPCR is better than YFP insertion} \]
\[\text{Bigger FRET response, no perturbation of coupling to downstream effector: Smaller is better!} \]
\[\text{Modified from Figs.1a, 1b, 2e, and 3a of Hoffmann C, et al: Nature Methods 2005; 2: 171-176} \]
up to that point would turn red (Fig. 6). They removed unbound ReAsH, then immediately used a glass microneedle to cut a dendrite off the rest of the cell. About 30% of the time, this portion of the neuron would continue to live for a few hours, even though separated from the main cell body and nucleus. Then they could ask; “Can this isolated dendrite make any fresh protein on its own?” By just waiting a few hours, then labeling with FlAsH, any fresh protein made after the surgery would label green, whereas old protein made before ReAsH withdrawal would stay red. Thus, if green synapses can be seen, fresh protein synthesis must have occurred in this isolated portion. And indeed, out here there is green dye in the synaptic spots, as you can see from this region at a higher magnification. This is true for both GluR1 and GluR2, both types of the AMPA receptor. An important control, to make sure that the dyes do not wash off, is to add cycloheximide during this chase time to inhibit all fresh protein synthesis. Only old red protein without any additional green is visible, as expected, showing that the ReAsH saturated all the old protein and did not wash away.

For comparison, they also applied this technique to the calcium/calmodulin kinase II (CaMKII) beta subunit, which unlike the alpha subunit is generally accepted to behave like a standard protein requiring the participation of the nucleus. Now there is fresh green CaMKII beta made afterwards in the cell body, but not in the peripheral dendrite beyond the site of the cut. Even more interestingly, it is believed that the 3’ untranslated region of the mRNA for GluR1 has the special targeting sequence that says “store me near synapses”. If you delete the 3’ untranslated region from the mRNA, GluR1 now looks like CaMKII beta, i.e. the cell body has both new green and old red copies whereas the isolated dendrite has only old red protein. Thus the sequential labeling technique can clearly distinguish dendrite-autonomous from nucleus-dependent protein synthesis.

Another important application of FlAsH and ReAsH is their ability to photogenerate singlet oxygen, which can locally and acutely inactivate proteins. Singlet oxygen is a very reactive excited state of O2 that attacks methionines, tryptophans, and histidines very close to the site of generation, and thereby has a high likelihood of inactivating the nearest proteins. For example, Oded Tour monitored gap junctions formed from tetracysteine

Fig. 6 Glutamate receptors can be newly synthesized in isolated neuronal dendrites reproduced from Fig.4 of Ju W, et al: Nat Neurosci 2004; 7: 244-253
tagged ReAsH-stained connexins.\(^2\) Initially there was good coupling through gap junctions, as shown by a high electrical coupling ratio. Applying modest levels of light did very little, but applying stronger illumination, but still only with a standard xenon lamp rather a laser, was enough to inactivate the gap junctions in just 20 or 25 seconds. There is no other pharmacological way to kill gap junctions.

This same local inactivation works on voltage-operated calcium channels (genes supplied by my brother Richard Tsien, Stanford Univ.) similarly tagged with tetracysteine motifs and labeled with ReAsH.\(^3\) Each depolarization evokes a large calcium current, which then slowly inactivates in a well-known manner. The peak amplitude is stable in the dark. Then if Dr. Tour then applies light for just five seconds, he immediately loses about 50\% of the current. Once again, that reduced amplitude is stable in the dark. Each time he illuminates for 5 sec, he loses about half the remaining current. So one can titrate the residual activity with different cumulative exposures to light. Control experiments show that other channels in the membrane remain unaffected during this procedure.

I have to say that those were not the most interesting biological uses. The most interesting use I would say so far has come from Gray Davis’ lab\(^4\) who applied it to the famous synaptic protein synaptotagmin, which is accepted to be important for exocytosis. The question Davis wanted to answer is whether synaptotagmin is separately important for endocytosis? One cannot answer this question with a genetic deletion because if you knock out synaptotagmin you have no exocytosis, so you cannot even measure endocytosis. They needed a way to acutely kill synaptotagmin only after exocytosis had already occurred. They expressed synaptotagmin with a tetracysteine tag in Drosophila embryos in which endogenous synaptotagmin had been deleted. After dye-labeling the synaptotagmin and stimulating exocytosis, they illuminated the synapses and showed that endocytosis was depressed compared to controls in which either the dye-labeling or illumination were omitted. This result argues that synaptotagmin has a direct role in endocytosis as well as exocytosis.

Genetically targeted photoinactivation is potentially an important general technique for rapidly and locally inactivating an existing protein. The more widespread methods for gene knockouts and RNA interference only remove the nucleic acid source of the protein. The existing protein just decays with its normal half-life, which can be hours to days, during which time there is a large chance for other proteins either to step in and compensate, or for an avalanche of new biological effects to occur based on the initial deletion. Such secondary effects make interpretation much more difficult. Moreover, none of those techniques has any spatial resolution. If one has or can develop a small molecule inhibitor of the protein of interest, one can acutely and suddenly eliminate its function, but such inhibitors are usually laborious and costly to develop for a new protein, requiring the resources of a pharmaceutical company. Genetically targeted photoinactivation merely requires appending the tetracysteine tag, labeling with the dye, and illuminating. One limitation is that this procedure only eliminates the exogenous, tagged copies unless they associate with endogenous untagged copies, in which case the singlet oxygen might inactivate the latter as well to give a dominant negative effect. Otherwise, one should knock out the endogenous copies while replacing them with tagged copies to keep the function temporarily normal until the latter are photoinactivated with the desired spatial and temporal control.

A final comment on the tetracysteine motif is that the original sequences have been greatly improved by a high throughput library screen done by Brent Martin.\(^5\) We now recommend a 12 amino acid sequence FLNCCPGCCMEP where the six amino acids surrounding the core CPGCC have been combinatorially optimized. The new sequence gives a better quantum yield and much better affinity of binding, enabling use of less biarsenical dye and more stringent conditions to wash away unbound and nonspecifically bound dye. This reduces the background staining that up to now has been the most severe limitation of the tetracysteine-biarsenical system.

My last topic is a more clinically oriented application.\(^4\) Despite the power of genetically encoded indicators or targeted labels, they will not be readily applicable to live people because we cannot make transgenic humans. There are high hopes for gene therapy of course, but it is clearly taking a lot longer than most of us had expected. To help patients as opposed to mice, it would be desirable to develop general strategies to target synthetic molecules to diseased tissue without using gene transfer (Fig. 7). The synthetic molecules should ideally encompass both diagnostic reporters and therapeutic molecules. A few existing ligands for endogenous receptors are already in use for in vivo imaging, such as dopamine receptor ligands discussed earlier in this symposium, but they are hard to generalize and crucially lack amplification. Moreover, each endogenous receptor can bind at most one copy of the ligand. Likewise each copy of an endogenous tumor-associated antigen can bind at most one copy of antibody. Antibodies also are big molecules that have a hard time penetrating deep into solid tumors and are slow to clear from the rest of the body.\(^6\) I would prefer an amplifying mechanism whereby each pathological target molecule can cause the accumulation of a large number of reporter or therapeutic molecules, whose size we can control from small to large and which are not restricted to fluorescence as their only readout. Such a mechanism might enable personalized medicine, in which imaging and therapy could work synergistically.
Many other labs had shown that certain polycationic peptides, initially derived from the Antennapedia homeobox protein or Tat protein of human immunodeficiency virus,\textsuperscript{26, 27} can drag normally impermeant cargoes into live cells. Such cargoes can range in size from isotopes or small dye molecules\textsuperscript{28, 29} to nanoparticles of many tens of nm diameter.\textsuperscript{30} The polycationic sequence can be as simple as a string of arginines,\textsuperscript{31, 32} which I like because then I do not have to remember complex sequences. The arginines can just as well be D-amino acids, arguing that the mechanism is probably not mediated by receptors of the usual sort. The positively charged arginines first stick to the outside of the cell, which is negatively charged because of phospholipid headgroups or proteoglycans. Probably these complexes then become endocytosed,\textsuperscript{33, 34} whereupon a fraction of them somehow escape from the endosomes and end up in both the cytosol and the nucleus. The detailed mechanism of cell penetration remains obscure and controversial. Furthermore, it was not known how to control this uptake to make it specific for particular types of cells, especially for tumor cells compared to normal.

We first needed a way to generally inhibit the uptake, then an additional mechanism to remove the inhibition in selected locations (Fig. 7). What could stop the polyarginine peptides from working? We decided to fuse them to a matching string of negative charges made of glutamates, in the hope that the positive and negative charges would fold back on each other, leaving too few positive charges to stick to the plasma membrane.\textsuperscript{34} When Dr. Tao Jiang made and tested such composite peptides, we were gratified to find that uptake is indeed largely blocked. However, this inhibition can be reversed by cleaving the covalent linkage between the polyarginines and polyglutamates, for example with a protease. In the absence of that covalent linkage, electrostatic attraction alone is not enough to hold the opposite charges together, so the two halves eventually drift apart, freeing the polyarginine and its still-attached cargo to stick to the outside of the cell and eventually be taken up. Therefore, adhesion and intracellular uptake of a wide range of potential cargoes is triggered by local extracellular protease activity, which in turn is known to be elevated in a wide variety of disease states, especially metastatic tumors.

We have some evidence that the peptide does form a hairpin, as I imagined in the cartoon of Fig. 7. Melinda Roy and Patricia Jennings, expert collaborators for nuclear magnetic resonance (NMR) spectroscopy, found many nuclear Overhauser couplings, especially between the glutamates and the arginines. Such couplings indicate that the glutamates and arginines are kissing each other in a hairpin loop.

Furthermore, the intact synthetic peptide including a row of D-glutamates, a cleavable linker containing L-amino acids, a row of D-arginines, and a far-red fluorescent dye (Cy5) to monitor where the arginines go, is
almost completely unable to enter cells in culture (Fig. 8). However, the linker is chosen to be a good substrate for matrix metalloproteinase-2 (MMP-2), which is a form of collagenase that many cancer cells secrete to enable them to chew their way through the normal extracellular matrix to invade fresh tissue. If we expose the peptide to active MMP-2, the linker is cleaved, jettisoning the inhibitory glutamates. Now the arginines carrying the Cy5 marker enter the cell very well, probably even reaching some nucleoli inside the nucleus. We have many examples of other peptide sequences whose uptake into cultured cells is similarly dependent on disease-associated proteases, but can we get this mechanism to work in a tumor in a live animal? In the traditional model for cancer research, human tumor cells are transplanted into immunocompromised mice. In this case Emilia Olson injected HT1080 fibrosarcoma cells in the armpits of nude mice where they grew up into tumors. The peptide was then injected into the tail vein. After an hour, imaging of far-red fluorescence through the skin of the live animal shows that the tumor typically lights up about threefold more brightly than the surrounding tissue or noninjected contralateral site (Fig. 9). The same contrast can be seen after sacrificing the animal and sectioning the recovered tumor. We believe that this contrast is dependent on cleavage, because we can compare a scrambled peptide where we merely permute the PLGLAG into LALGPG. This is not a substrate anymore for MMP-2, even though it has the same amino acids, molecular weight, and hydrophobicity. Therefore it is not much accumulated into the tumor. Tumor contrast thus correlates to protease susceptibility.

Most recently, we have started to work on cancers from real patients even though we are not yet ready for the very expensive and difficult procedure to initiate clinical trials injecting our probes into people. Instead, we found a surgeon, Dr. Quyen Nguyen, who is regularly taking head and neck tumors out of patients, together with the normal tissue surrounding the tumor that must also be removed to try to ensure getting all the tumor. Since this tissue would be discarded anyway, it is easy to get patients’ consent to test our molecules on these freshly excised samples, which we cut into slices and incubate with the cleavable peptide or scrambled peptide. Under optimum conditions, the tumor can be up to almost eightfold brighter than the surrounding normal tissue (Fig. 10), whereas the scrambled peptide shows negligible contrast. In some cases we can even see that
the brightest staining is in the most dedifferentiated part of the tumor, whereas a redifferentiated tumor portion does not stain as well, shown here by a keratin pearl, a diagnostic sign for a squamous cell carcinoma that is trying to become normal tissue again. Thus we have some preliminary hint that this approach can work on clinical specimens. Of course the ultimate goal is to image not the tumor that has just been removed, but any tumor that is left in the patient, which will require real clinical trials and elaborate regulatory approval.

We believe that this new mechanism for cleavage-activated uptake should deliver contrast agents for all modes of imaging, even though the results I showed so far are limited to fluorescence. However, others have shown that tat or polyarginine peptides will work for radioactive or magnetic reporters as well. We also hope to exploit radiation sensitizers like boron-10 or gadolinium-157. The latter is particularly attractive because it should be both be visible by MRI and usable for neutron capture therapy. The linkers can be cleaved by a variety of proteases or even by reduction of a disulphide bond, which could potentially recognize hypoxic, partially necrotic tissue. We are not dependent on antibodies. This mechanism allows enzymatic amplification. Protease activities are mechanistically important and interesting. The immediate challenges for the future are to increase the contrast yet further and to demonstrate in vivo non-optical imaging. My last scientific comment is that this mechanism is an extracellular analog of FRET, in that the polyarginine is like a donor, whereas the polyglutamate is like a quencher. When they are together, the natural tendency of the polyarginine to enter cells is quenched. The difference from FRET is that this mechanism is not restricted to fluorescence, but should also allow radioactive, magnetic, and therapeutic cargoes to be concentrated in diseased tissue.

These diverse examples of molecular engineering can be summarized as follows: We have new long-wavelength fluorescent proteins. Somatic hypermutation can redirect the genius of the immune system. Tetracysteine/biarsenical labeling has some unique advantages and applications. We have glimmerings of a new approach towards cancer imaging and therapy, which is something I have always wanted to develop before my career ends, particularly because my father and a nephew both died from cancer.
References


