Clues to the evolution of complex signaling machinery

Bruce J. Mayer*

Raymond and Beverly Sackler Laboratory of Genetics and Molecular Medicine, Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, CT 06030

Multicellular life demands that activities such as cell proliferation, differentiation, adhesion, and motility be exquisitely controlled. Many of these activities are regulated by tyrosine phosphorylation, the covalent addition of a phosphate group to tyrosine residues in cell proteins, and the emergence of this signaling mechanism may actually have been a key enabling event in the transition to multicellularity. Two papers in this issue of PNAS take advantage of genome sequences from diverse eukaryotic lineages to address the origins of the tyrosine phosphorylation signaling machinery; they provide tantalizing insights into more general questions of how complex signaling mechanisms might have evolved (1, 2).

Protein tyrosine phosphorylation is regulated in the cell by the opposing activities of two enzymes: protein tyrosine kinases (PTKs), which transfer phosphate from ATP to substrate proteins, and protein tyrosine phosphatases (PTPs), which remove it. Functionally, the most important effect of tyrosine phosphorylation is to create high-affinity binding sites for other proteins containing small modular phosphotyrosine (pTyr)-binding domains, most notably Src homology 2 (SH2) domains (3). Thus tyrosine phosphorylation transmits downstream signals by creating new protein complexes in the cell, leading to changes in the subcellular localization or enzyme activity of the binding partners.

Despite its importance, tyrosine phosphorylation is a relatively rare modification; in eukaryotes the vast majority of protein phosphorylation is on serine and threonine side chains. These sites are phosphorylated by serine/threonine kinases and dephosphorylated by serine/threonine phosphatases, which are recognizably distinct from their pTyr-specific counterparts at the sequence level. For the most part, each class of enzymes is highly specific for its corresponding amino acid substrate, because of the very different shape and chemical reactivity of tyrosine compared with serine and threonine.

An Evolutionary Paradox

The exploitation of tyrosine phosphorylation for signaling must be a relatively recent evolutionary innovation, because simple single-celled eukaryotes contain few if any recognizable PTKs, PTPs, or pTyr-binding domains. Because the three activities (kinase, phosphatase, and binding module, or "writer, eraser, and reader" in the terminology of Lim and colleagues) act in concert to regulate signaling, this has raised the chicken-and-egg question of how the whole system could have evolved; presumably, each activity in the absence of the others would confer no selective advantage. Some have used this paradox to argue that natural selection cannot fully explain the complexity of existing signaling mechanisms. As beautifully illustrated by these two papers, however, comparative genome analysis has gone a long way toward resolving this puzzle.

Pincus *et al.* (2) used sequence similarity to identify and enumerate all of the PTKs, PTPs, and SH2 domains in diverse eukaryotic lineages. They found that these genomes could be clearly divided into two groups. One group, including all multicellular animals (metazoans), has abundant PTKs, PTPs, and SH2 domains; the second group, composed entirely of single-celled eukaryotes, lacks recognizable PTKs and has few if any PTPs or SH2s. As pointed out by the authors, this observation suggests an evolutionary "phase transition" during which the number of genes encoding all three activities expanded in concert, as expected if they began working together to generate useful signaling machinery when multicellular animals emerged. In this regard perhaps the most interesting genome, which has only recently been completed (4), is that of the unicellular choanoflagellate *Monosiga brevicollis*. Evolutionarily, choanoflagellates are thought to be the most closely related of all single-celled organisms to metazoans (5). Although all choanoflagellates have a free-living unicellular stage, many can also form colonies, reinforcing the idea that these organisms represent a transitional stage between unicellular and multicellular life. Unexpectedly, the *Monosiga* genome is brimming with tyrosine kinase machinery, in number and complexity at least as elaborate as that of any metazoan genome. This diversity is more fully explored by Manning *et al.* (1).

Despite its importance, tyrosine phosphorylation is a relatively rare modification.

A Plausible Model

What do these new analyses suggest about how the full-blown PTK–PTP–SH2 system of information transfer evolved? Important clues are provided by the genomes of simpler unicellular eukaryotes, which lack the complete set of pTyr signaling machinery. Fungi and slime molds each have one or more SH2 domains and PTPs, but no recognizable PTKs. This raises the question of what fitness advantage is provided by PTP and SH2 domains in the absence of tyrosine kinases—in other words, what use are reader and eraser in the absence of the writer? As pointed out by Pincus *et al.* (2), the likely answer lies in the multifunctional protein kinases that can, albeit inefficiently, phosphorylate tyrosine in addition to serine and threonine residues. All eukaryotes, including yeasts and plants, use the tyrosine phosphorylation of MAP kinases by multifunctional kinases to regulate diverse cell activities (6). Thus the evolution of PTPs that could efficiently dephosphorylate such sites could be favored, even in the absence of dedicated PTKs.

We also have a few hints on the evolution of pTyr binding domains. *S. cerevisiae* contains a single recognizable SH2 domain, in the transcription elongation factor SPT6. This domain does not appear to bind to pTyr, however, but instead binds to the serine-phosphorylated tail of RNA polymerase (7). It is likely that chance mutation of such a primordial domain conferred the ability to bind specifically to tyrosine-phosphorylated sites, thus enabling entirely novel signaling connections based on tyrosine phosphorylation.

Genomes of other simple eukaryotes provide a glimpse into the next steps in the process. The slime mold *Dictyostelium discoideum* lacks dedicated PTKs, but it does sport a fairly rich repertoire of SH2-containing proteins (at least 13). Here we see a few recognizable orthologs of familiar metazoan SH2-containing signaling proteins: STAT (a transcriptional activator) and Cbl (a ubiquitin ligase that tags proteins for proteolysis). A crystal struc-

Author contributions: B.J.M. wrote the paper.

The author declares no conflict of interest.

See companion articles on pages 9674 and 9680.

*E-mail: bmayer@neuron.uchc.edu.

ture of the *Dictyostelium* STAT protein shows that its SH2 domain can bind to tyrosine-phosphorylated sites (8). So by the time of the divergence of the metazoan and *Dictyostelium* lineages, SH2 domains had apparently evolved the ability to bind pTyr, and they had been put to work in useful ways (regulating transcription, regulating protein stability). The stage was now set for the evolution of enzymes that could efficiently and specifically phosphorylate tyrosine.

In some common ancestor of *Monosiga* and metazoans, it is likely that this fateful event occurred via mutation of an existing serine/threonine kinase or multifunctional kinase. Probably losing the ability to phosphorylate serine and threonine was at least as important as being able to phosphorylate tyrosine more efficiently, because this would open up a new, untapped signaling system orthogonal to existing systems. New connections could be explored without direct (and presumably often detrimental) effects on existing regulatory pathways based on serine/threonine phosphorylation. Some of these innovations likely made it easier for cells to communicate with each other and cooperate to form tissues, organs, and ultimately complex multicellular animals. Of course this new signaling bandwidth was neither necessary nor sufficient for multicellularity—*Monosiga*, after all, clings to its single-celled lifestyle despite its highly elaborate pTyr signaling machinery, and multicellular plants get on quite well without most components of tyrosine kinase signaling (9)—but it certainly must have helped.

Insight from the Monosiga Machinery

Manning *et al*. (1) provide a closer look at the remarkably numerous and diverse pTyr signaling proteins in *Monosiga*. By careful sequence analysis, they count a minimum of 128 PTKs, 39 PTPs, and 123 SH2 domain-containing proteins, more than found in any other organism, including humans (the exact numbers of genes identified by the two groups differ considerably, reflecting different stringencies in the search methodologies used, as well as the inherent challenges in extracting gene family members from genomic sequences). Remarkably, these domains and activities are associated with other signaling domains in many unique combinations not seen in metazoans, strongly suggesting that the two lineages diverged before most of the pTyr machinery had evolved. Thus comparing the pTyr signaling components of *Monosiga* and metazoans is particularly interesting for two reasons: the common elements tell us what must have evolved first, whereas those that evolved after the divergence provide insight into the constraints on the system.

On the first point, although no clear orthologs are seen among the many transmembrane receptor tyrosine kinases found in the two lineages, there are clear orthologs for several nonreceptor PTKs [cytosolic TKs, or CTKs in the terminology of Manning *et al.* (1)]. Thus it is likely that these were the first PTKs to evolve, before the divergence of *Monosiga* and metazoans, also recently suggested by Miyata and colleagues (10). The CTKs common to the two lineages all contain an N-terminal SH2 domain, which allows them to interact with tyrosine-phosphorylated proteins, and most also have membrane-anchoring sites. In metazoans, CTKs transduce signals from the environment by associating with transmembrane proteins such as cytokine and adhesion receptors. Membrane association keeps the CTKs near the transmembrane receptors, and SH2 domains facilitate binding and processive phosphorylation of substrates (11). The early emergence of CTKs that could bind to tyrosine-phosphorylated proteins and to membranes may have helped pave the way for the explosion of new combinations seen in both metazoans and choanoflagellates.

SH2 domain-containing proteins are the other class of pTyr signaling proteins that share common domain architectures in *Monosiga* and metazoans; at least 15 orthologous groups are identified by Manning *et al*. (1). Thus by the metazoan–*Monosiga* divergence, much of the "reader" machinery had already evolved into its mature form, while most of the PTKs and PTPs had not. This makes some sense, as it is the ability to interpret (respond to) a signal in new ways that is most likely to provide selective advantage to the organism. Once a few PTKs were in place, the key innovation was evolution of a diverse class of binding proteins that could respond exclusively to tyrosine phosphorylation.

Thus at the point where the *Monosiga* and metazoan lineages diverged, the common ancestor had a well elaborated set of SH2-containing signaling proteins, a limited set of cytosolic PTKs, and a few functional PTPs. Since that divergence, each lineage has had many hundreds of millions of years to explore the useful configurations that can be built by using these and other functional domains as building blocks. To those familiar with the pTyr signaling repertoire of typical metazoans (which differs surprisingly little from worms to humans), the new domain combinations in *Monosiga* are bizarre and fascinating. One feels a bit like an explorer in a land populated with entirely unknown creatures—very different from those back home, but composed of an essentially similar set of parts.

Amid the variety of new combinations, Manning *et al.* (1) point out a number of interesting commonalities. For example, both lineages encode numerous transmembrane receptors with PTK and PTP domains, although sequence comparison suggests none of these are derived from common ancestors. Clearly, tyrosine kinase and phosphatase activities that can be directly regulated by extracellular ligands are quite useful and evolved independently a number of times. Both lineages also contain several nonorthologous proteins with SH2 domains linked to domains that activate or inactivate small GTPases such as Ras and Rho. The ability to couple preexisting GTPase signaling pathways to the newly developing tyrosine phosphorylation-based machinery was apparently advantageous and therefore exploited independently in both lineages.

These two papers provide a plausible evolutionary path for building the pTyr signaling machinery, and new insight into the constraints (the needs of the organism, the available building blocks) that shaped its evolution. More importantly, perhaps, they help move our discussion of the origin of complex signaling mechanisms out of the realm of philosophical musing toward one of hard data and hypothesis-driven experimentation.

1. Manning G, Young SL, Miller WT, Zhai Y (2008) The protist, *Monosiga brevicollis*, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan. *Proc Natl Acad Sci USA* 105:9674–9679.
2. Pincus D, Letunic I, Bork P, Lim WA (2008) Evolution of the phospho-tyrosine signaling machinery in premetazoan lineages. *Proc Natl Acad Sci USA* 105:9680–9684.
3. Liu BA, *et al*. (2006) The human and mouse complement of SH2 domain proteins—establishing the boundaries of phosphotyrosine signaling. *Mol Cell* 22:851–868.
4. King N, *et al*. (2008) The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature* 451:783–788.
5. King N (2004) The unicellular ancestry of animal development. *Dev Cell* 7:313–325.
6. Widmann C, Gibson S, Jarpe MB, Johnson GL (1999) Mitogen-activated protein kinase: Conservation of a three-kinase module from yeast to human. *Physiol Rev* 79:143–180.
7. Yoh SM, Cho H, Pickle L, Evans RM, Jones KA (2007) The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes Dev* 21:160–174.
8. Soler-Lopez M, *et al*. (2004) Structure of an activated *Dictyostelium* STAT in its DNA-unbound form. *Mol Cell* 13:791–804.
9. Wang H, Chevalier D, Larue C, Cho SK, Walker JC (February 20, 2007) The protein phosphatases and protein kinases of *Arabidopsis thaliana*. *The Arabidopsis Book*, eds Somerville CR, Meyerowitz EM (Am Soc Plant Biologists, Rockville, MD), pp 1–38, 10.1199/tab.0106.
10. Suga H, *et al*. (2008) Ancient divergence of animal protein tyrosine kinase genes demonstrated by a gene family tree including choanoflagellate genes. *FEBS Lett* 582:815–818.
11. Shokat KM (1995) Tyrosine kinases: Modular signaling enzymes with tunable specificities. *Chem Biol* 2:509–514.