

# Pathway Logic Modeling of Protein Functional Domains in Signal Transduction

Carolyn Talcott, Steven Eker, Merrill Knapp, Patrick Lincoln, Keith Laderoute  
SRI International, 333 Ravenswood Avenue, Menlo Park CA 94025  
{firstname.lastname}@sri.com

## Abstract

We describe the application of an approach called Pathway Logic to the symbolic modeling of protein functional domains (PFDs). We show how signal transduction processes can be modeled at different levels of abstraction involving either an overall state of a protein or its PFDs and their interactions, and how the resulting signalling network can be queried using formal methods tools.

## 1. Introduction

Cells respond to changes in their environment through biochemical signaling pathways that detect and transmit information to effector molecules within different cellular compartments. Protein functional domains (PFDs) are consensus sequences within signaling molecules that recognize and bind other signaling components to make complexes. Pathway Logic is an application of techniques from formal methods to the modeling and analysis of signal transduction networks in mammalian cells [2]. These signaling network models are developed using Maude (<http://maude.cs.uiuc.edu>), a formal language based on rewriting logic. Models can be queried (analyzed) using the execution, search and model-checking tools of the Maude system. Models can also be exported in formats suitable for input to other tools for additional analysis capabilities and visualization.

## 2. Signaling Networks in Pathway Logic

Several important concepts from cell biology must be defined in Maude to model signaling processes, including: intracellular proteins, biochemicals such as second messengers, extracellular stimuli, biochemical modification of proteins, protein association, and cellular compartmentalization of proteins. These concepts are represented as algebraic terms in Maude by declaring constants naming proteins and other chemicals and operators representing modification, association, and compartmentalization. Biochem-

ical processes are expressed as rewrite rules. Each rewrite rule has the form  $t \rightarrow t'$  where  $t$  and  $t'$  represent a local part of the system state, and the rule says that when the system has a subcomponent matching  $t$ , that subcomponent can evolve to  $t'$ , possibly concurrently with changes described by rules matching other parts of the system state. A signaling network is represented by a collection of rewrite rules together with these declarations. Pathways in the network are generated automatically using tools for execution and search.

A Pathway Logic model of the EGFR network is being developed in which the rewrite rules, curated from the scientific literature, describe relevant biochemical processes in terms of overall protein state (level 1). In addition, we are developing the capability of building, displaying, and analyzing Pathway Logic models at a level of abstraction that explicitly involves PFDs and posttranslational modifications of individual signaling molecules (level 2). Here we use the recruitment and activation the ubiquitous Raf-1 serine/threonine protein kinase to illustrate the two levels of representation, since the Raf-1 system is a reasonably well established and detailed example of a signal integrator in the EGFR network [3, 1].

## 3. Modeling the Activation of Raf-1

At level 1 the activation of Raf-1 and its recruitment to the cell membrane is represented by the following rule.

```
crl [280.?Ras.?Pak.Src.PP2A.?14-3-3.->.Raf1] :
  {CM | cm [?Ras - GTP] [?Pak - act] [Src - act]
    {cyto Raf1 [?14-3-3 - phos] PP2A }}
=>
  {CM | cm [?Ras - GTP] [?Pak - act] [Src - act]
    [Raf1 - act] [?14-3-3 - phos]
    {cyto PP2A}}
  if ?Ras S:Soup := N-Ras K-Ras H-Ras .
```

The rule says that if the cell contains a Ras type protein with a GTP modification ([?Ras - GTP]), an activated Pak ([?Pak - act]), and an activated Src ([Src - act]) are on the interior side of the cell membrane (CM); and Raf-1, phosphorylated 14-3-3 and PP2A are in the cytoplasm the

Raf-1 can be activated [Raf1 - act] and recruited along with 14-3-3, leaving PP2A in the cytoplasm.

In the level 2 representation the activation of Raf-1 represented by rule 280 above becomes several rules in which structural features of some of the proteins, including Raf-1, are annotated with information about relevant PFDs and binding sites. An example we show the rule in which Src phosphorylates partially activated Raf-1 at Tyrosine 341.

```

r1[Raf1#6.Y341phos]:
{CM | cm PS PA [?Slk - act]
 [?Ras | GTPbound, (RafBD - bound)]
 [Raf1 | (S 43), (S 259), (Y 341), (C1 - bound),
 (S 621 - phos - bound), (PABM - bound),
 (RBD - bound), raf1:Atts]
 [14-3-3a | (SBD - bound), (DMD - bound), 1a:Atts]
 [14-3-3b | SBD, (DMD - bound), (T 141 - phos)]
 e((Raf1, (S 621)), (14-3-3a, SBD))
 e((Raf1, C1), b(PS)) e((Raf1, PABM), b(PA))
 e((14-3-3a, DMD), (14-3-3b, DMD))
 e((Raf1, RBD), (?Ras, RafBD))
 {cyto}}
=>
{CM | cm PS PA [?Slk - act]
 [?Ras | GTPbound, (RafBD - bound)]
 [Raf1 | (S 43), (S 259), (Y 341 - phos),
 (S 621 - phos - bound), (PABM - bound),
 (C1 - bound), (RBD - bound), raf1:Atts]
 [14-3-3a | (SBD - bound), (DMD - bound), 1a:Atts]
 [14-3-3b | SBD, (DMD - bound), (T 141 - phos)]
 e((Raf1, (S 621)), (14-3-3a, SBD))
 e((Raf1, C1), b(PS)) e((Raf1, PABM), b(PA))
 e((14-3-3a, DMD), (14-3-3b, DMD))
 e((Raf1, RBD), (?Ras, RafBD))
 {cyto}} .

```

On the lefthand side Raf-1 is associated with a dimer of 14-3-3 scaffold/adaptor proteins through binding of phosphorylated serine 621 ((S 621 -phos - bound)) in Raf-1 to the serine binding domain (SBD - bound) in the 14-3-3 dimer. This binding is represented by the edge e((Raf1, (S 621)), (14-3-3a, SBD)). In addition Raf1 is bound to Ras, PS, and PA. Notice that the rule mixes level 1 and level 2 notation, only using level 2 detail where relevant.

Level 2 rules for Raf-1 are connected to level 1 by a rule that converts the level 1 representation of inactivated Raf-1 to its level 2 representation, and a rule that converts the level 2 complex representing activated Raf-1 to its level 1 representation.

#### 4. Using the Pathway Logic model

To test the rules for Raf-1 activation at the PFD level we define an initial cell state containing inactive Raf-1 and the postulate necessary ingredients to activate it.

```

eq qraf =
  PD({CM | PS PA [Pak1 - act] [PKCz - act]
 [Src - act] [H-Ras - GTP]
 [Raf1.inact 14-3-3a 14-3-3b PP2A]}) .

```

The form PD( . . . ) represents a cell in a Petri dish, possibly with some external signalling compounds.

For example “can Raf-1 can be activated?” is answered by using the default execution strategy (applying the rules to generate a pathway), and the answer is yes. Other pathways may be possible, using search we discover that there are two ways to activate Raf-1, but they differ only in the order of application of independent rules (in general we might discover quite different pathways). This is good, but we might worry that the rules also generate impossible states, for example a state in which Raf-1 is bound to both 14-3-3’s in the dimer as well as being bound to PS and PA. To check that this can not happen, we can search for a cell state matching a pattern such as

```

PD( out:Soup {CM | cm:Soup [H-Ras - GTP]
 e((14-3-3a, DMD), (14-3-3b, DMD))
 e((Raf1, (S 621)), (14-3-3a, SBD))
 e((Raf1, (S 259)), (14-3-3b, SBD))
 e((Raf1, C1), b(PS))
 e((Raf1, PABM), b(PA))
 {cyto:Soup}})

```

and indeed Maude confirms that such a state is not reachable.

## 5. Conclusions

Pathway Logic is an example of how formal modeling techniques can be used to develop a new science of symbolic systems biology. We believe that this computational science will provide researchers with powerful tools to facilitate the understanding of complex biological systems and accelerate the design of experiments to test hypotheses about their functions in vivo. In particular we are interested in formalizing models that biologists use to think about signalling pathways in familiar terms that allow the biologist to compute with them and ask questions about possible consequences. Here we have illustrated our approach using the biochemistry of signaling involving the ubiquitous Raf-1 serine/threonine protein kinase.

## References

- [1] A. S. Dhillon and W. Kolch. Untying the regulation of the Raf-1 kinase. *Arch. Biochem Biophys*, 404:3–9, 2002.
- [2] S. E. et. al. Pathway logic: Executable models of biological networks. In *Fourth International Workshop on Rewriting Logic and Its Applications (WRLA'2002)*, volume 71 of *Electronic Notes in Theoretical Computer Science*. Elsevier, 2002. <http://www.elsevier.nl/locate/entcs/volume71.html>.
- [3] W. Kolch. Meaningful relationships: The regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. *Biochem J.*, 351:289–305, 2000.