Pathway Logic Modeling of Protein Functional Domains in Signal Transduction

Carolyn Talcott Steven Eker
Patrick Lincoln Merrill Knapp
Keith Laderoute

<FirstName.LastName>@sri.com

About Pathway Logic

Pathway Logic is an approach to modeling biological entities and processes based on formal methods and rewriting logic. Using Pathway Logic signal transduction processes have been modeled at different levels of abstraction involving:

- the overall state of proteins, or
- protein functional domains (PFDs) and their interactions

These signaling networks can be queried using formal methods tools. For example, by choosing an initial condition and trying the following:

- execution—show me some signaling pathway
- search—show me all pathways leading to a specified final condition
- model-checking—is there a pathway with certain given properties?

Pathway Logic

There is a practical need to represent very large biological networks of all kinds as models at different levels of detail/abstraction (see Panel 5 below). For example, consider the following:

- The proteome of eukaryotic cells is at least an order of magnitude larger than the genome (very large and diverse protein networks)
- A large fraction of the genome of mammalian cells ($\approx 10\%$ of the human genome) encodes genomic regulators producing very large regulatory networks of the genome itself
- Biological networks interact as modules/subnetworks to produce high levels of physiological organization (e.g., circadian clock subnetworks are integrated with metabolic, survival, and growth subnetworks)

In silico models of such networks would be valuable but must have certain features. In particular, they must be easily modified–extended or updated–and useable by bench researchers for formulating and testing hypotheses about how signals and other changes are propagated.

Goals

Build network models that working biologists and biomedical researchers can interact with and modify.

Make formal methods tools accessible to the general biological and biomedical research community.

Enable bench researchers to generate informed hypotheses about complex biological networks. For example, a researcher should be able to ask the question "How is the network perturbed when I knockout/in gene X".

Multiple Levels of Abstraction

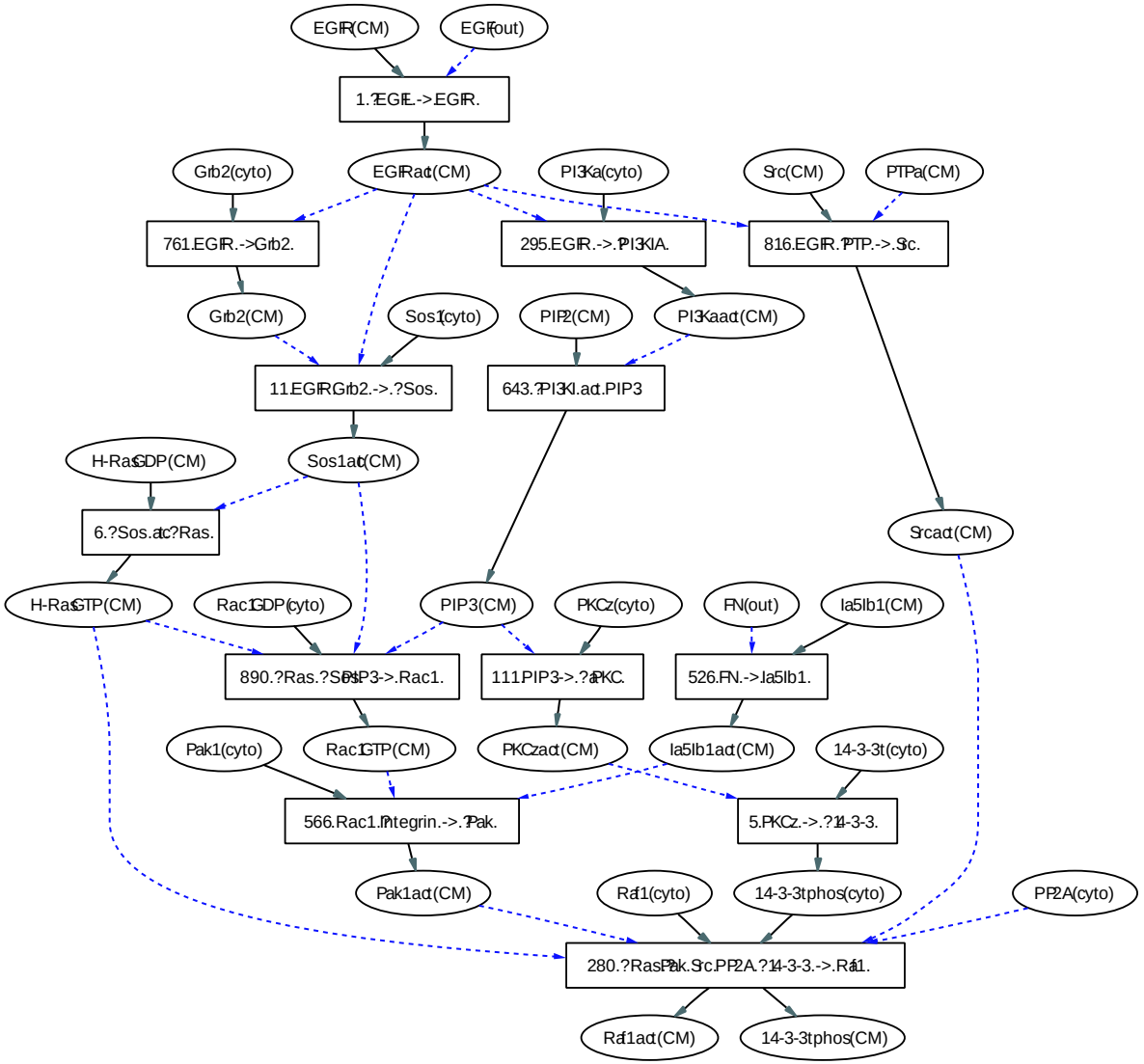
Level I

- Signaling proteins considered as abstract entities possibly annotated with an indication of their state of activation or inactivation
- Complex formation, activation, inactivation processes are modeled as atomic/discrete/single step processes

Level II

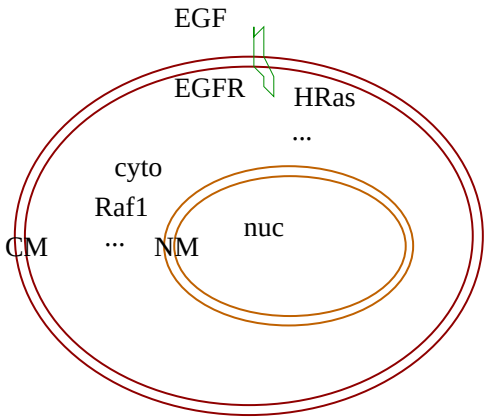
- Relevant protein functional domains, binding sites, and/or biochemical modifications are made explicit as attributes which themselves can be annotated
- Signal transduction processes elaborated to model protein-protein interactions

Pathway Logic



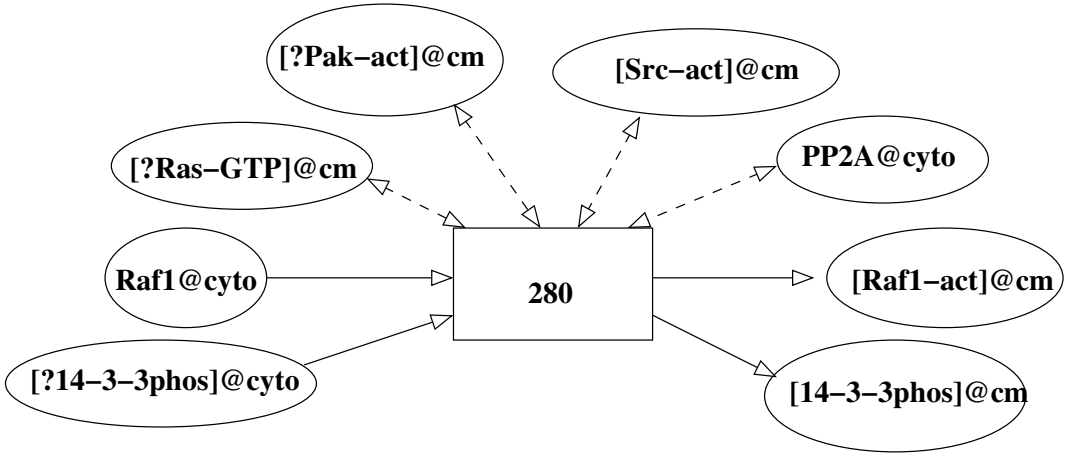
A portion of the Epidermal Growth Factor Receptor (EGFR) network leading to activated Raf1 (Level I)

Representing Cells in Maude




```
qt = PD(Ca++ EGF FN
      {CM  | EGFR Fak Ia5Ib1 *** cm EGFRs / Epac / Integrins
             PIP2 PKA PTPa Pten Src [H-Ras - GDP]
             { 14-3-3t  Akt1 Cas Cbl Gab1 Grb2  *** cytoplasm
               Eps8 Erk1 Jnk1 Ksr1
               Mek1 Mkk3 Mkk4 Mkk6          *** cyto MKK
               Mekk1 Mekk4 Pak1 Pdk1        *** cyto MKKK
               PI3Ka                        *** cyto PI3Ks
               PKCb1 PKCd PKCz              *** cyto PKCs
               PLCg PLCe                    *** cyto PLCs
               PP2A Raf1 Shc Sos1           *** cyto
               [Rac1 - GDP]                 *** cyto Stats
        {NM  | empty {empty}}}}) .          *** nucleus
```

Level I Raf activation rule

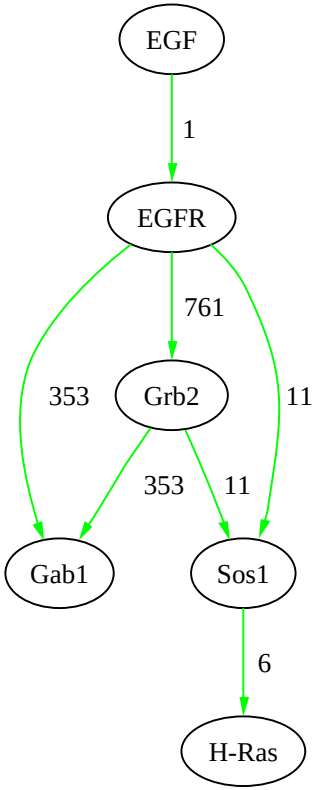



```
crl[280.?Ras.?Pak.Src.PP2A.?14-3-3.->.Raf1]:
  {CM | cm [?Ras - GTP] [?Pak - act] [Src - act]
       {cyto Raf1 [?14-3-3 - phos] PP2A }}
  =>
  {CM | cm [?Ras - GTP] [?Pak - act] [Src - act]
       [Raf1 - act] [?14-3-3 - phos]
       {cyto PP2A}}
 if ?Ras S:Soup := N-Ras K-Ras H-Ras [ metadata "21192014(R)"] .
```

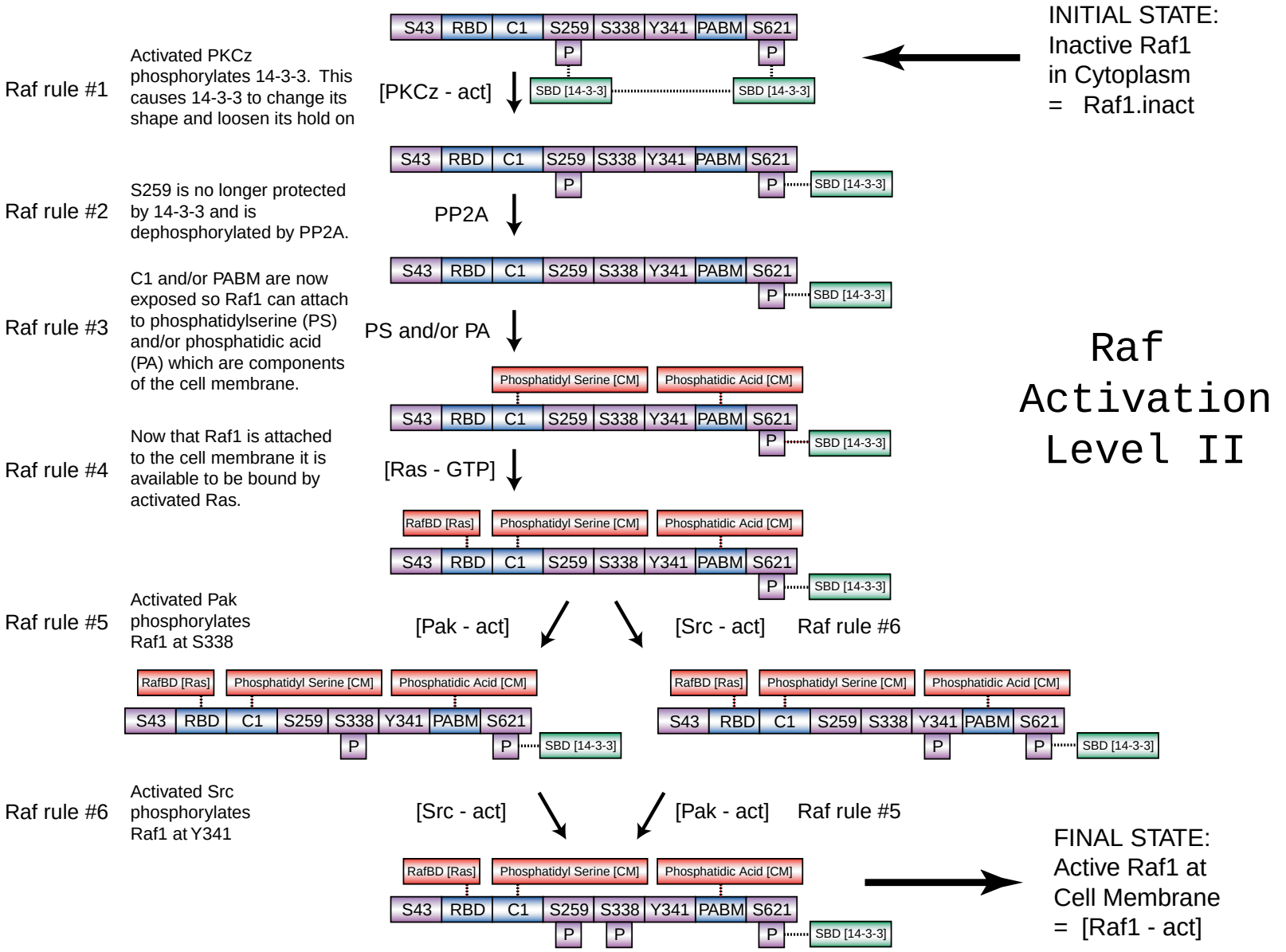
Analysis Level I

Can Raf-1 be activated starting from initial state qt ? Yes

```
eq PD(out:Soup
      {CM | cm:Soup [H-Ras - GTP]
          {cyto:Soup}})
   |= propt = true .

red findPath(qt0,propt).
```



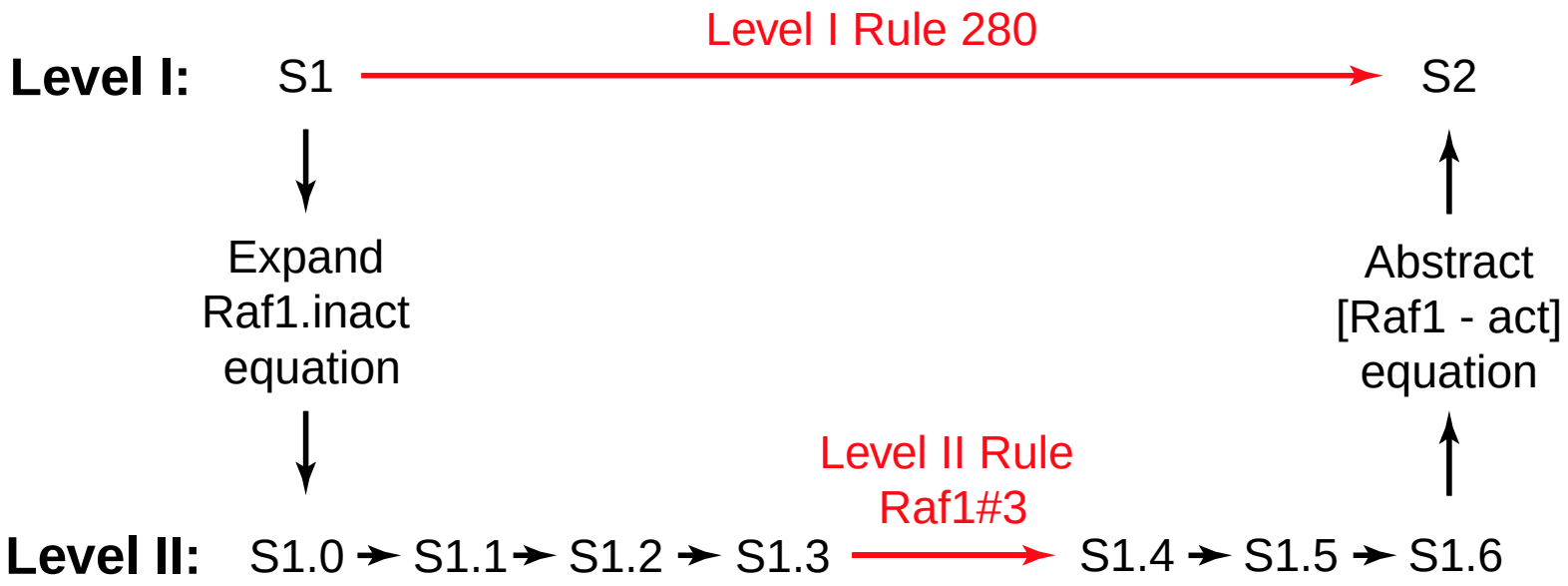
Pathway Logic



Level II representation of proteins and signaling

```
rl[Raf1#3.PS.PA]:
 {CM | cm PS PA
   {cyto [Raf1 | (S 43), (S 259), (S 338), (Y 341),
                 (S 621 - phos - bound), C1, PABM, raf:Atts]
       [14-3-3a | (SBD - bound), (DMD - bound), 1a:Atts]
       [14-3-3b | SBD, (DMD - bound), (T 141 - phos)]
       e((14-3-3a, DMD), (14-3-3b,DMD))
       e((Raf1, (S 621)), (14-3-3a,SBD))}}
  =>
 {CM | cm PS PA
       [Raf1 | (S 43), (S 259), (S 338), (Y 341),
               (S 621 - phos - bound), (C1 - bound), (PABM - bound), raf:Atts]
       [14-3-3a | (SBD - bound), (DMD - bound), 1a:Atts]
       [14-3-3b | SBD, (DMD - bound), (T 141 - phos)]
       e((14-3-3a,DMD), (14-3-3b,DMD))
       e((Raf1, (S 621)), (14-3-3a,SBD))
       e((Raf1, C1), b(PS)) e((Raf1,PABM), b(PA))
      {cyto}} .
  [ metadata "21278045(R-20)  20379031(D) for PA  99426181(D) for PS"] .
```

Connecting levels I and II



Using the level II model

Initial state:

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

Can Raf-1 in a cell described by qraf be activated? Yes

```
eq PD( out:Soup {CM | cm:Soup [Raf1 - act] {cyto:Soup}} )
     |= praf0 = true .
Maude> red findPath(qraf,praf0) .
> SimplePath:  spath(Raf1#1.PKCz Raf1#2.PP2A Raf1#3.PS.PA Raf1#4.Ras
      Raf1#5.S338phos Raf1#6.Y341phos Raf1#7.Raf1.is.act,
      PD({CM | PA PS [Pak1 - act] [PKCz - act]
               [H-Ras - GTP] [Src - act] [Raf1 - act]
               {14-3-3b PP2A 14-3-3a}}))
```

Validating the level II model

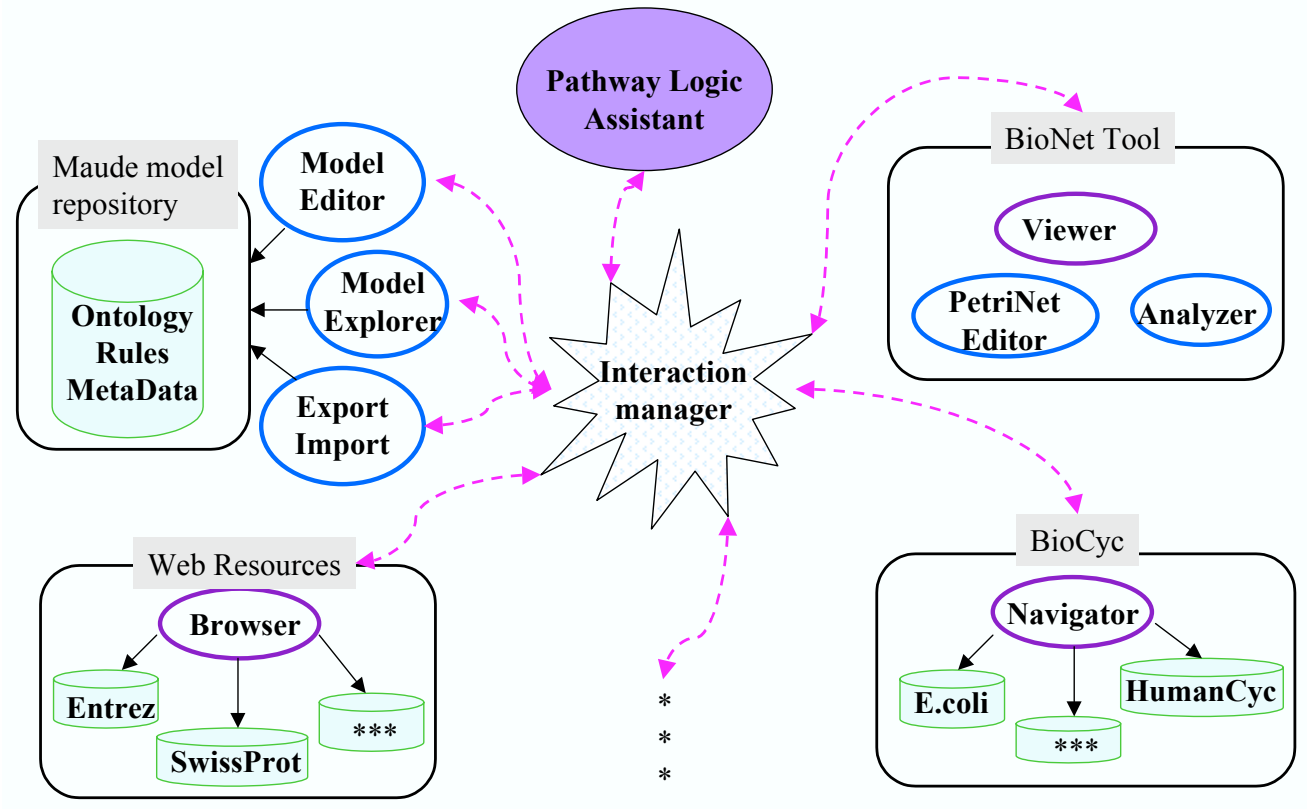
Can Raf1 bind to PS and PA at the cell membrane? Yes

```
eq PD( out:Soup {CM | cm:Soup
      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))
      {cyto:Soup}} ) |= praf1 = true .

Maude> red findPath(qraf,praf1) .
result: SimplePath spath(Raf1#1.PKCz Raf1#2.PP2A Raf1#3.PS.PA PD(...))
```

Can Raf-1 be bound to both 14-3-3s and PS and PA? No

```
eq 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}} )

Maude> red findPath(qraf,praf2) .
result: SimplePath (noPath).SimplePath
```

Future Directions



Pathway Logic Workbench

- Bigger models
- Curation from wider range of experimental data
- Metabolic networks

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. Eker and et. al. Pathway logic: Symbolic analysis of biological signaling. In *Proceedings of the Pacific Symposium on Biocomputing*, pages 400–412, January 2002.

[3] S. Eker, M. Knapp, K. Laderoute, P. Lincoln, and C. Talcott. Pathway logic: Executable models of biological networks. In *Fourth International Workshop on Rewriting Logic and Its Applications*, 2002.

[4] K. W. Kohn. Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Mol Biol Cell*, 10:2703–2734, 1999.

[5] W. Kolch. Meaningful relationships: The regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. *Biochem J.*, 351:289–305, 2000.

[6] Maude Website: `http://maude.cs.uiuc.edu`