Pathway Logic

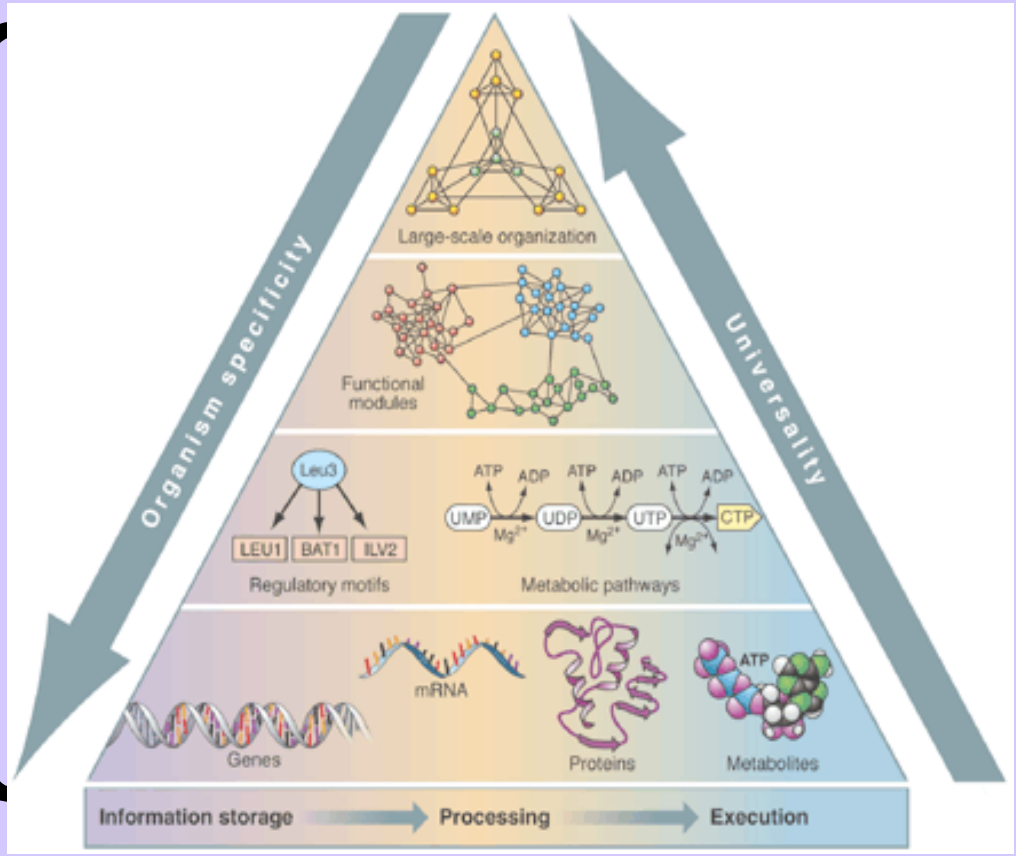
Towards a Symbolic Systems Biology

Carolyn Talcott
Merrill Knapp
Keith Laderoute
Patrick Lincoln

Systems Biology

Organisms are integrated subsystems
at multiple levels of organization

Pathway Logic:



Pathway Logic:

Problem

There is a practical need to represent very large biological networks of all kinds as models at different levels of detail/abstraction.

- 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 (~ 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 they must be

- easily modified--extended or updated
- useable by bench researchers for formulating and testing hypotheses about how signals and other changes are propagated.

Symbolic Systems Biology at SRI

- BioCyc
  - EcoCyc, MetaCyc, HumanCyc …
- BioSPICE
  - Integrated system design and modeling tools
  - Biological data warehousing
  - Database and tool interoperability
- Pathway Logic
  - Tools for symbolic modeling of biological pathways
- Hybrid Qualitative/Quantitative modeling
  - Delata/Notch signaling
  - B.subtilis sporulation

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
- Essentially the same approach has also been used to model metabolic pathways.

- These signaling and metabolic 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
  - show me all pathways leading to a specified final condition
  - model-checking---is there a pathway with certain given properties?

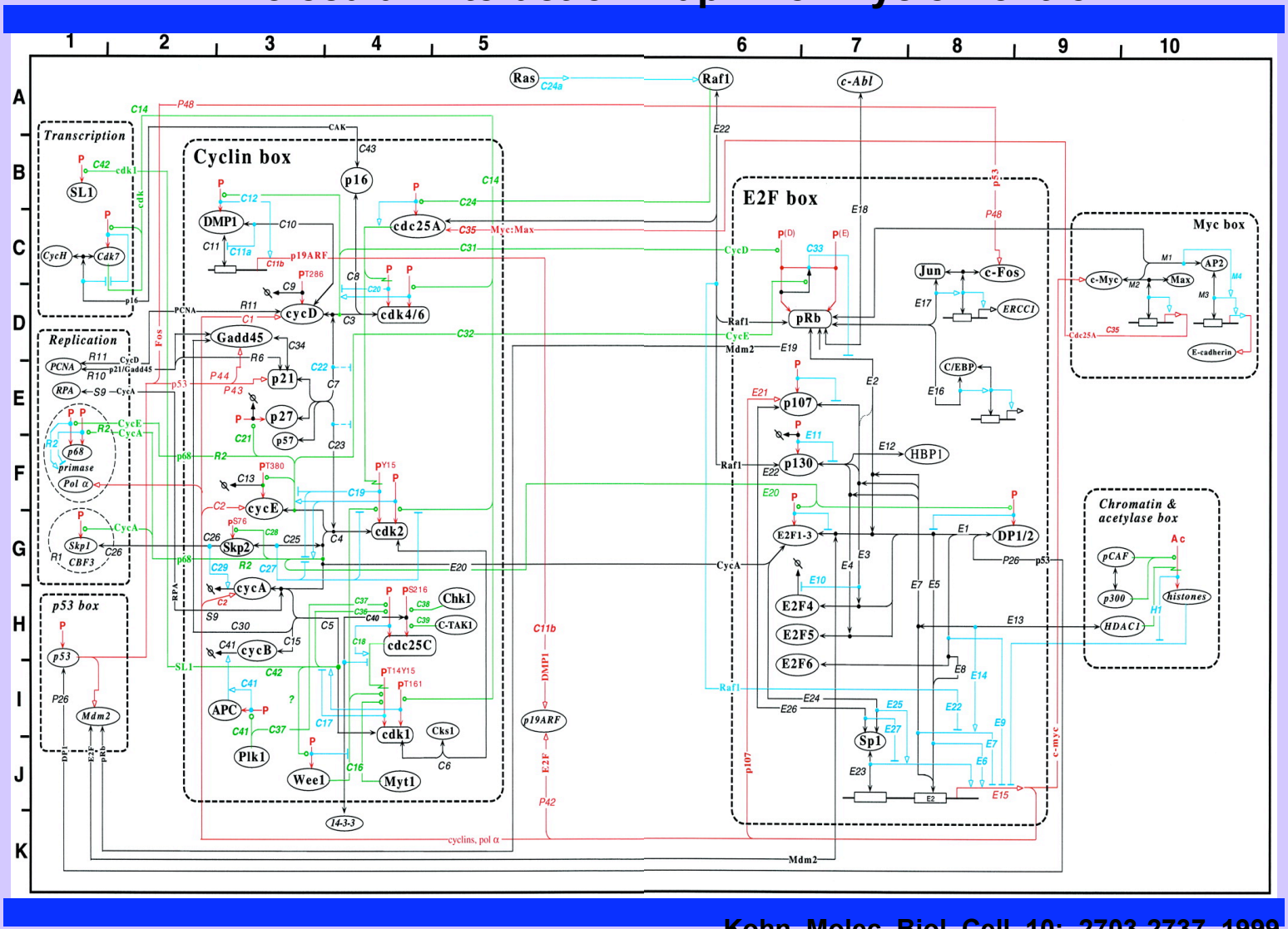
Pathway Logic Goals

- 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''.

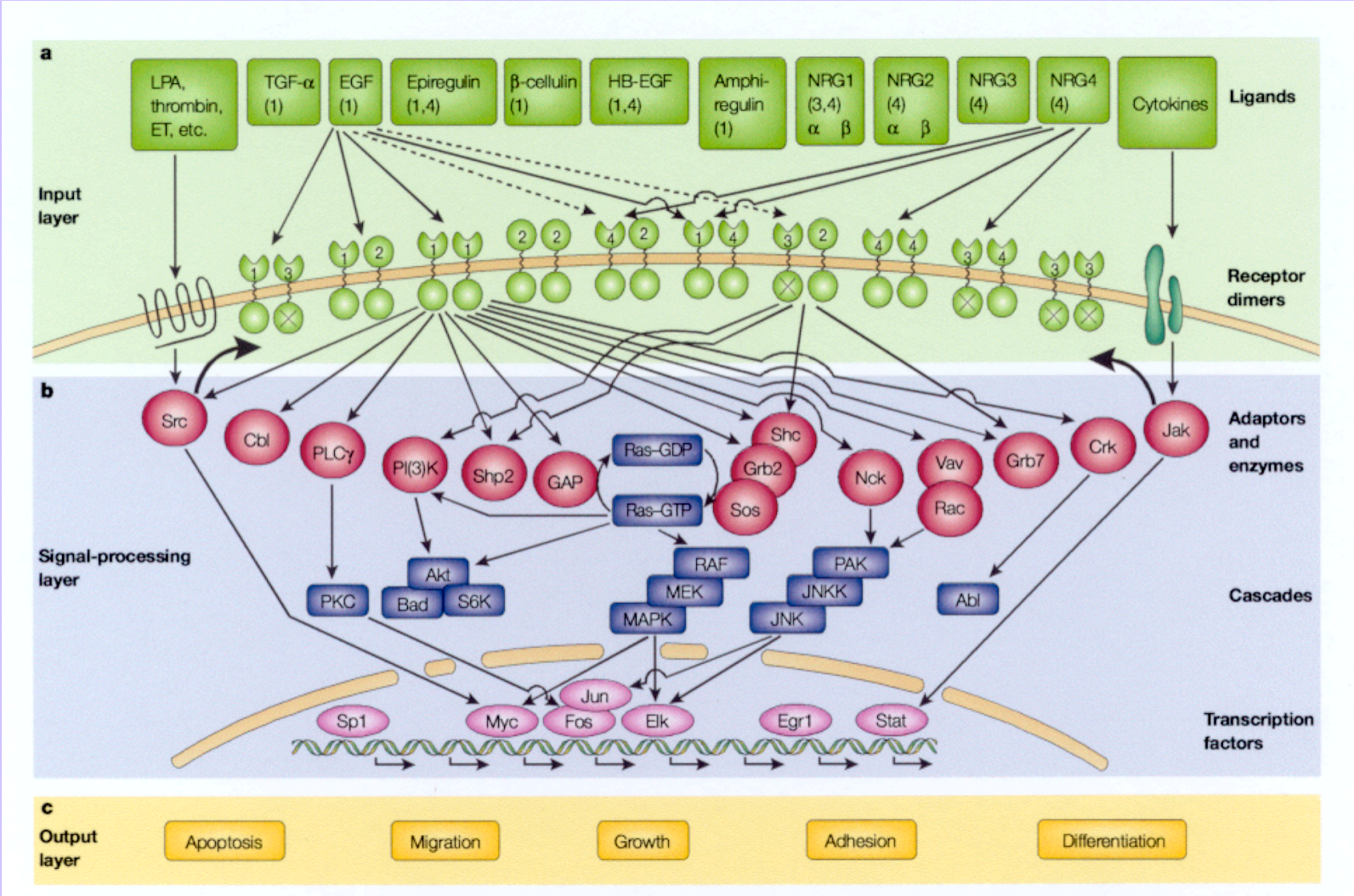
Modeling Cellular Networks: "Wiring Diagrams"

A Molecular Interaction Map: Cell Cycle Control



Kohn, Molec. Biol. Cell, 10: 2703-2737, 1999

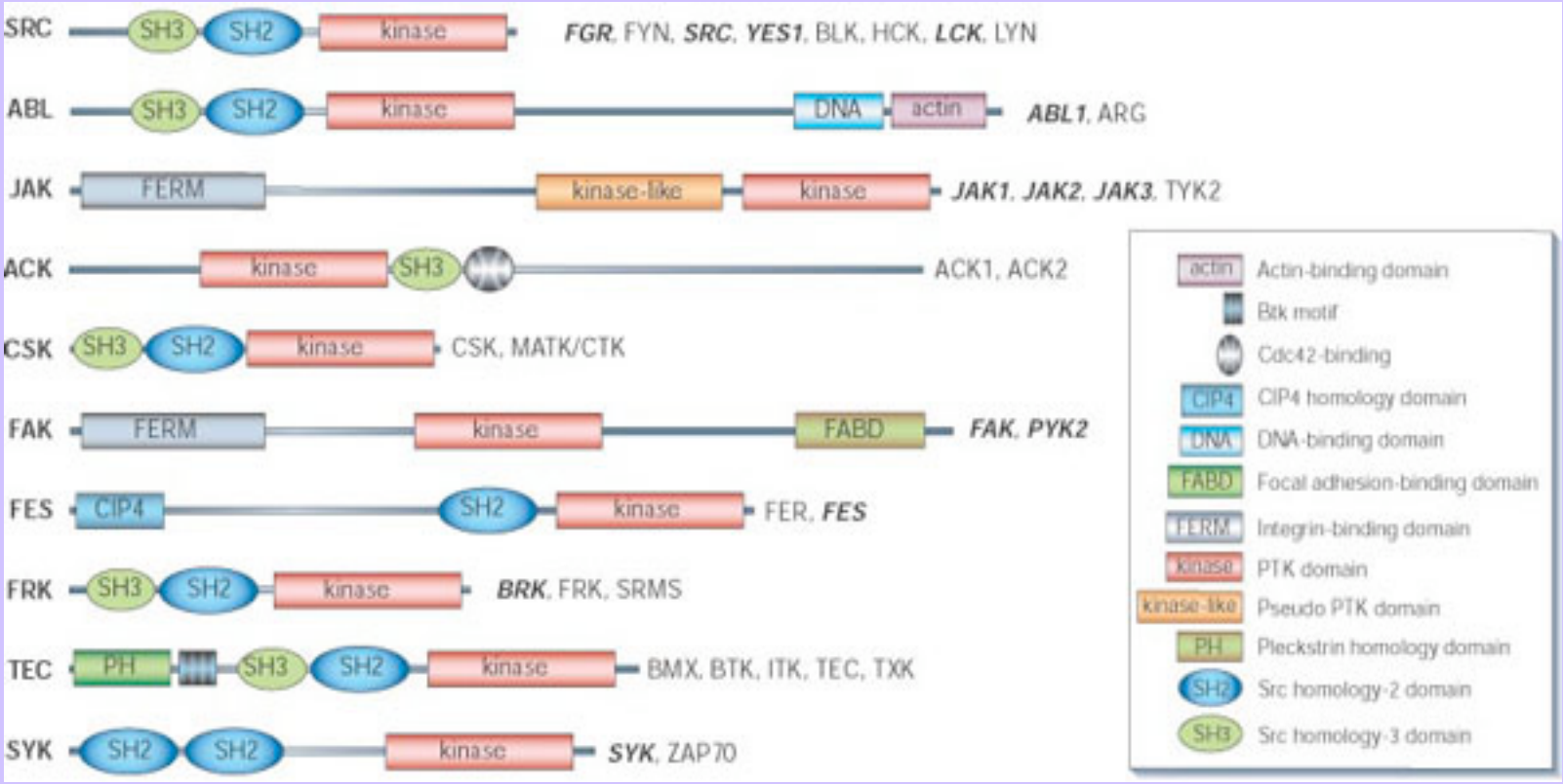
The ErbB Network



Yardon and Sliwkowski, Nat. Rev. Mol.Cell Biol. 2: 127-137, 2001

Protein Functional Domains

Concept: Signaling Proteins Are Collections of Domains or Modules
(Blume-Jensen and Hunter, Nature 411: 355-65, 2001



Some human cytoplasmic protein kinases.

Rewriting Logic
The formal technology

- Network components and interactions are symbolized for analysis in rewriting logic (using only curated data/results) using the Maude system.
- A transition in the state of a system is represented by a rewriting rule $t \rightarrow t'$ where t represents a local part of the initial state that is replaces by t' when the rule is applied.
- Execution from an initial state of interest gives an idea of model behavior.
- More possible outcomes of a networked process can be found by search.
- Hypotheses are explicitly tested by model checking techniques.

EGF - EGFR Network

A little sample of Pathway Logic modeling

EGF/EGFR experiments

Activation of a transcription factor (cJun cFos) following binding of extracellular epidermal growth factor (EGF) to its receptor (EGFR)

```
ops q1 q1x : -> Dish .
eq q1 = PD(EGF {CM | EGFR Pak1 PIP2 nWasp [H-Ras - GDP]
                     {Akt1  Gab1 Grb2 Gsk3 Eps8 Erk1
                      Mek1 Mekk1 Mekk4 Mkk4 Mkk3 Mlk2
                      Jnk1 p38a p70s6k Pdk1 PI3Ka PKCb1
                      Raf1 Rsk1 Shc Sos [Cdc42 - GDP]
                      {NM | empty {cJun cFos }}}}) .

eq q1x = q1 - < PI3Ka , cyto >
```

Model Checking

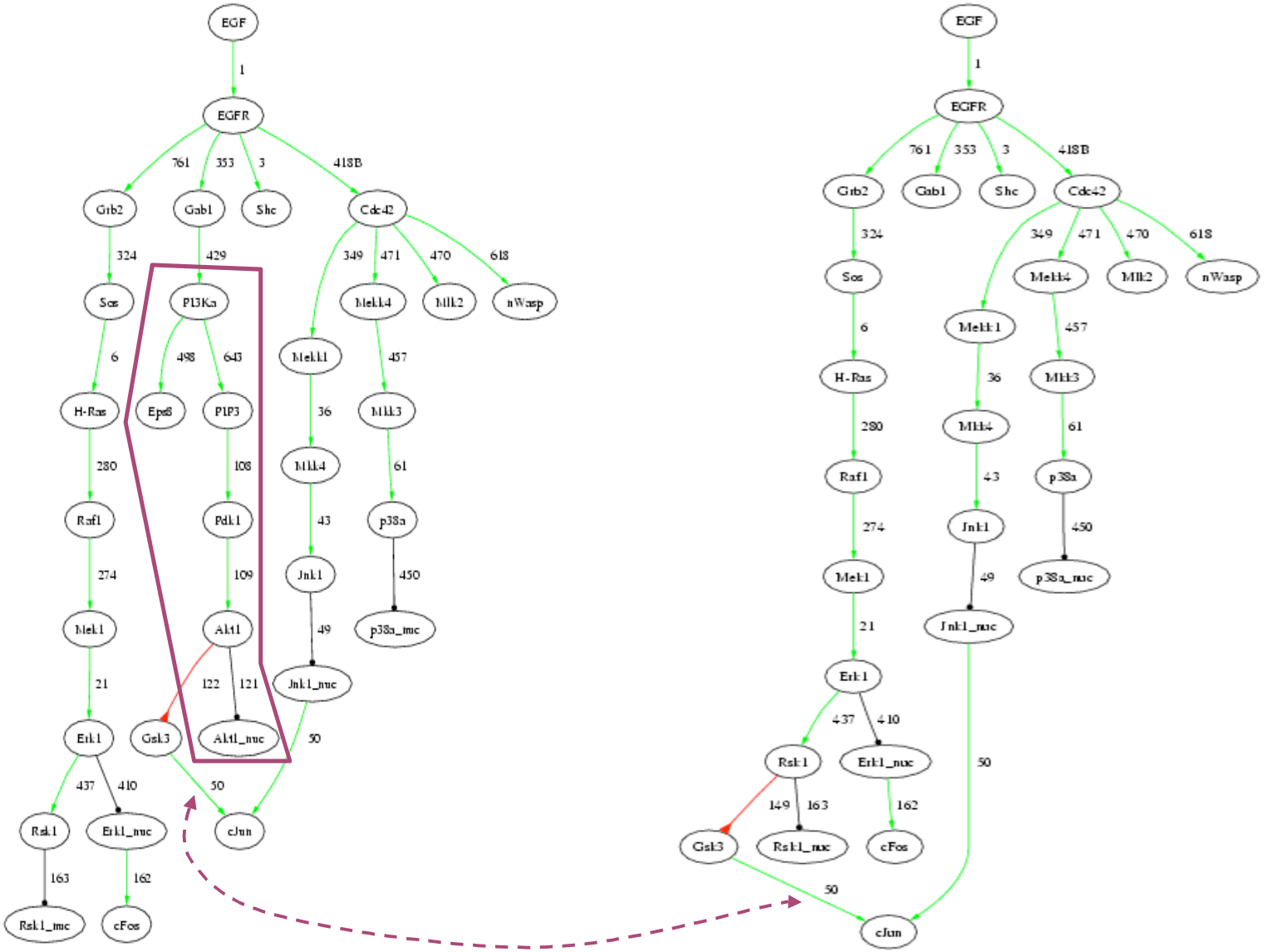
```
subsort  Dish < State .
eq PD(out:Soup
      {CM | cm:Soup
           {cyto:Soup
            {NM | nm:Soup
                  {nuc:Soup
                   [cJun - act] [cFos - act] }}}})
    |= prop1 = true .

eq findPath(S:State,P:Prop)
         = getPath(P:Prop, S:State |= ~ <> P:Prop) .

red  findPath(q1,prop1) .
red  findPath(q1x,prop1) .
```

Model Checking q1 -- simple paths

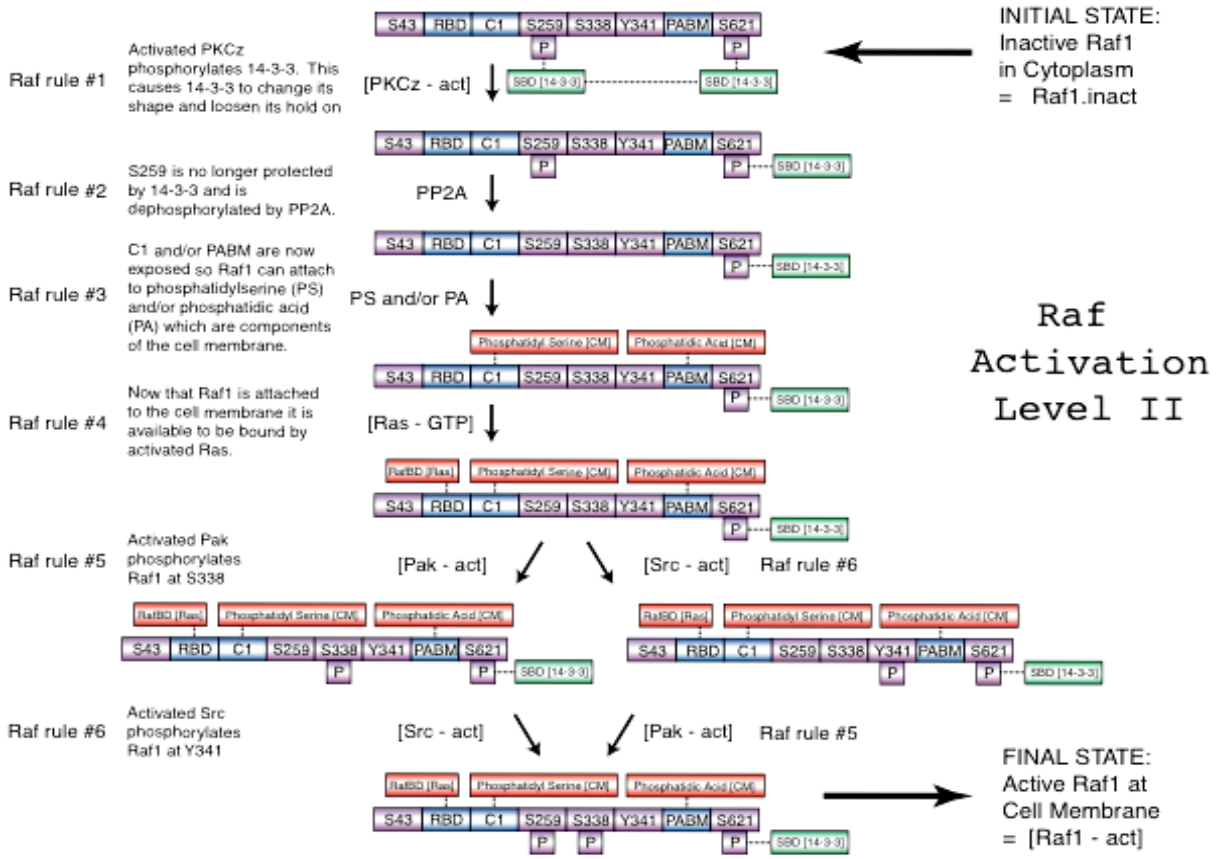
```
•Maude> red findPath(q1,prop1) .
•result SimplePath: spath('1.EGFLike.binds.EGFR '353M.EGFR.act.Gab1
'418B.EGFR.act.Cdc42 '349.Cdc42.act.Mekk1 '471.Cdc42.act.Mekk4
'470.Cdc42.act.Mlk '3.EGFR.act.Shc '429.Gab1.act.PI3K '498.PI3K.act.Eps8
'411.Shc.act.Grb2 '324.Shc.Grb2.act.Sos '618.Cdc42.act.Wasp '6.Sos.act.Ras
'280.Ras.act.Raf '274.Raf1.act.Mek1/2 '643.PI3KI.mets.PIP2.to.PIP3
'108.PIP3.act.Pdk1 '109.PIP3.Pdk1.act.Akt '122.Akt.deact.Gsk3    '121.Akt.to.nuc
'21.Mek1.act.Erk1/2 '437.Erk.act.Rsk '410.Erk1/2.to.nuc '163.Rsk.to.nuc
'36.Mekk1.act.Mkk4/7 '457.Mekk4.act.Mkk3/6 '61.Mkk3/6.act.p38 '450.p38.to.nuc.
'43.Mkk4/7.act.Jnk '49.Jnk.to.nuc '50B.Jnk.Gsk3.act.cJun '162.Erk.Rsk.act.cFos,
•PD(EGF {CM | Grb2 Pak1 PIP3 PKCb1 [Cdc42 - gtp] [EGFR - act] [Eps8 - act] [Gab1 -
act] [Pdk1 - act] [PI3Ka - act] [Raf1 - act] [H-Ras - gtp] [Shc - act] [Sos - act]
[nWasp -    act]{p70s6k [Gsk3 - deact] [Mek1 - act] [Mekk1 - act] [Mekk4 - act]
[Mkk3 - act] [Mkk4 - act] [Mlk2 - act] {NM | empty{[Akt1 - act] [cFos - act] [cJun
- act] [Erk1 - act] [Jnk1 - act] [p38a - act] [Rsk1 - act]}}}}))
```

```
•Maude> red findPath(q1x,prop1) .
•result SimplePath: spath('1.EGFLike.act.EGFR '353M.EGFR.act.Gab1
'418B.EGFR.act.Cdc42 '349.Cdc42.act.Mekk1 '471.Cdc42.act.Mekk4 '470.Cdc42.act.Mlk
'3.EGFR.act.Shc '411.Shc.act.Grb2 '324.Shc.Grb2.act.Sos '618.Cdc42.act.Wasp
'6.Sos.act.Ras '280.Ras.act.Raf '274.Raf1.act.Mek1/2 '21.Mek1/2.act.Erk1/2
'437.Erk.act.Rsk '410.Erk1/2.to.nuc '36.Mekk1.act.Mkk4/7 '457.Mekk4.act.Mkk3/6
'61.Mkk3/6.act.p38 '450.p38.to.nuc. '43.Mkk4/7.act.Jnk '49.Jnk.to.nuc
'149.Rsk.iphos.Gsk3 '50B.Jnk.Gsk3.act.cJun '163.Rsk.to.nuc '162.Erk.Rsk.act.cFos,
•PD(EGF {CM | Grb2 Pak1 PIP2 [Cdc42 - GTP] [EGFR - act] [Gab1 - act] [Raf1 - act]
[H-Ras- GTP] [Shc - act] [Sos - act] [nWasp - act] {Akt1 Eps8 Pdk1 p70s6k PKCb1
[Gsk3 - deact] [Mek1 - act] [Mekk1 - act] [Mekk4 - act] [Mkk3 - act] [Mkk4 - act]
[Mlk2 - act]  {NM | empty {[cFos - act] [cJun - act] [Erk1 - act] [Jnk1 - act]
[p38a - act] [Rsk1 - act]}}}}))
```

Roadmaps for q1,q1x runs

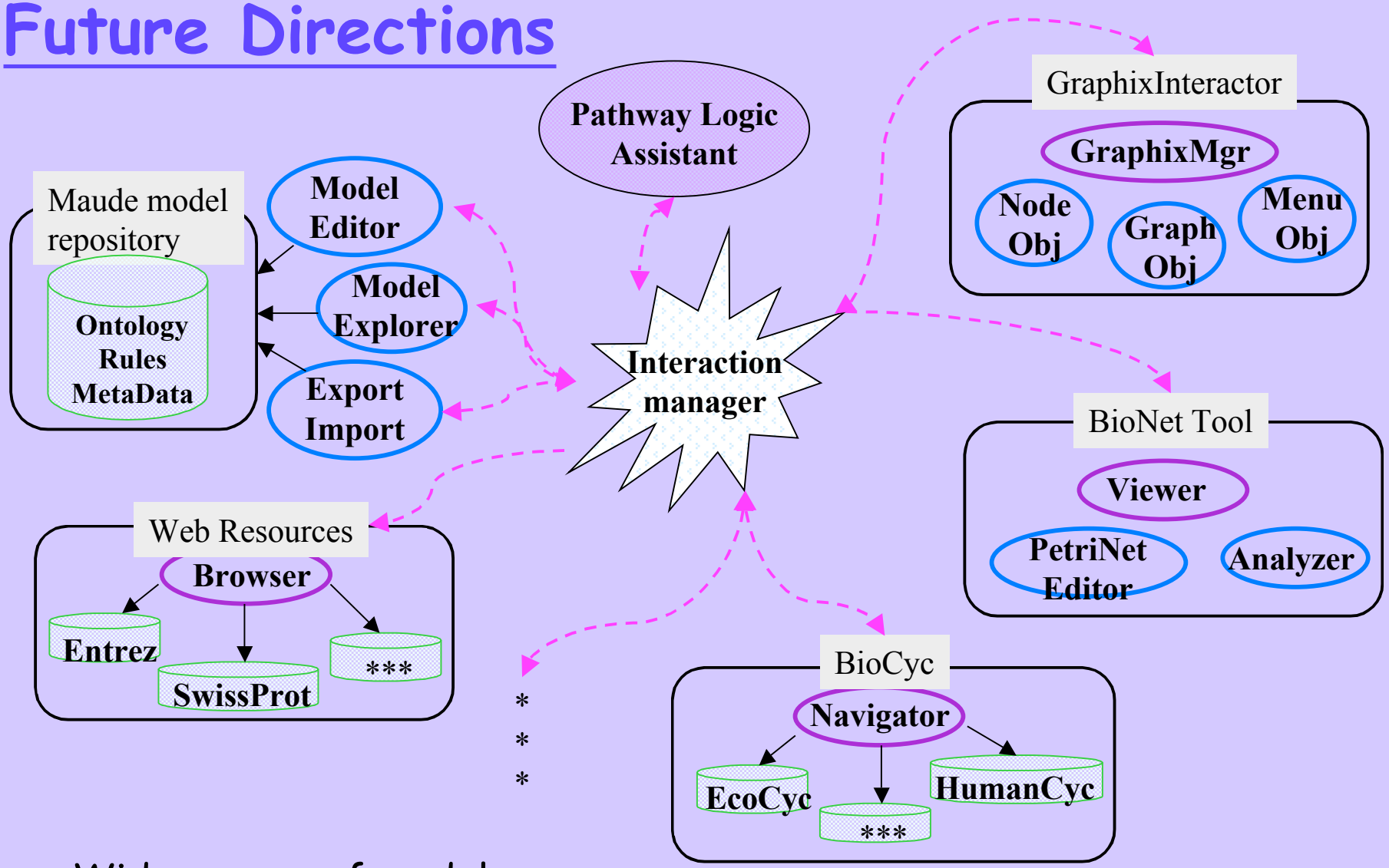


Raf Domains Diagram



Raf Activation Level II

Future Directions



- Wider range of models
- Combining signaling and metabolic networks
- Database and tool interoperation, better interaction

Pathway Logic Workbench Demo

- Gycolysis signaling pathway
- Gycolysis metabolic pathway