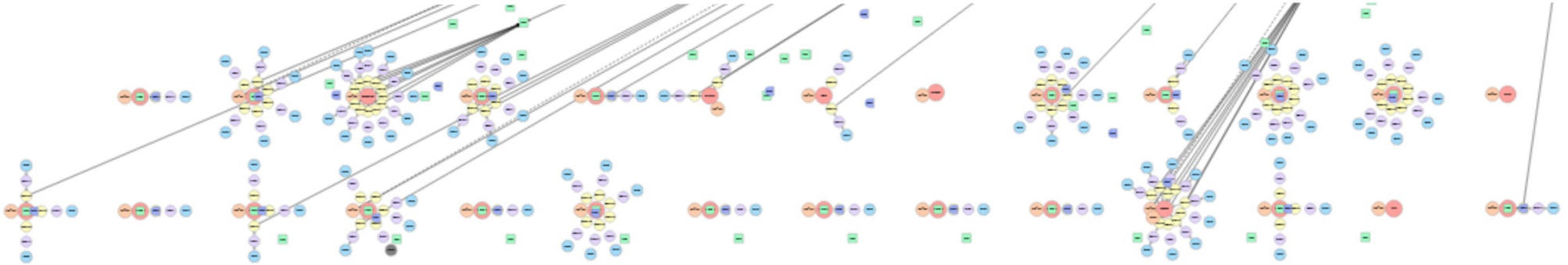


Biocuration and rule-based modelling of protein interaction networks in KAMI

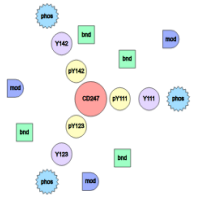
Sébastien Légaré, Eugenia Oshurko and Russ Harmer

LSB 2018 - Oxford

ENS DE LYON

July 13, 2018

Introduction



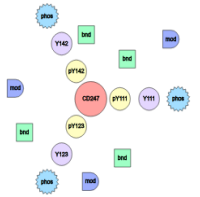
What does **K A M I** stand for?

Knowledge **A**ggregator and **M**odel **I**nstantiator

What is special about KAMI ?

- Rule-based strategy
- Incremental aggregation of large models
- Allows *a posteriori* understanding of models

Outline



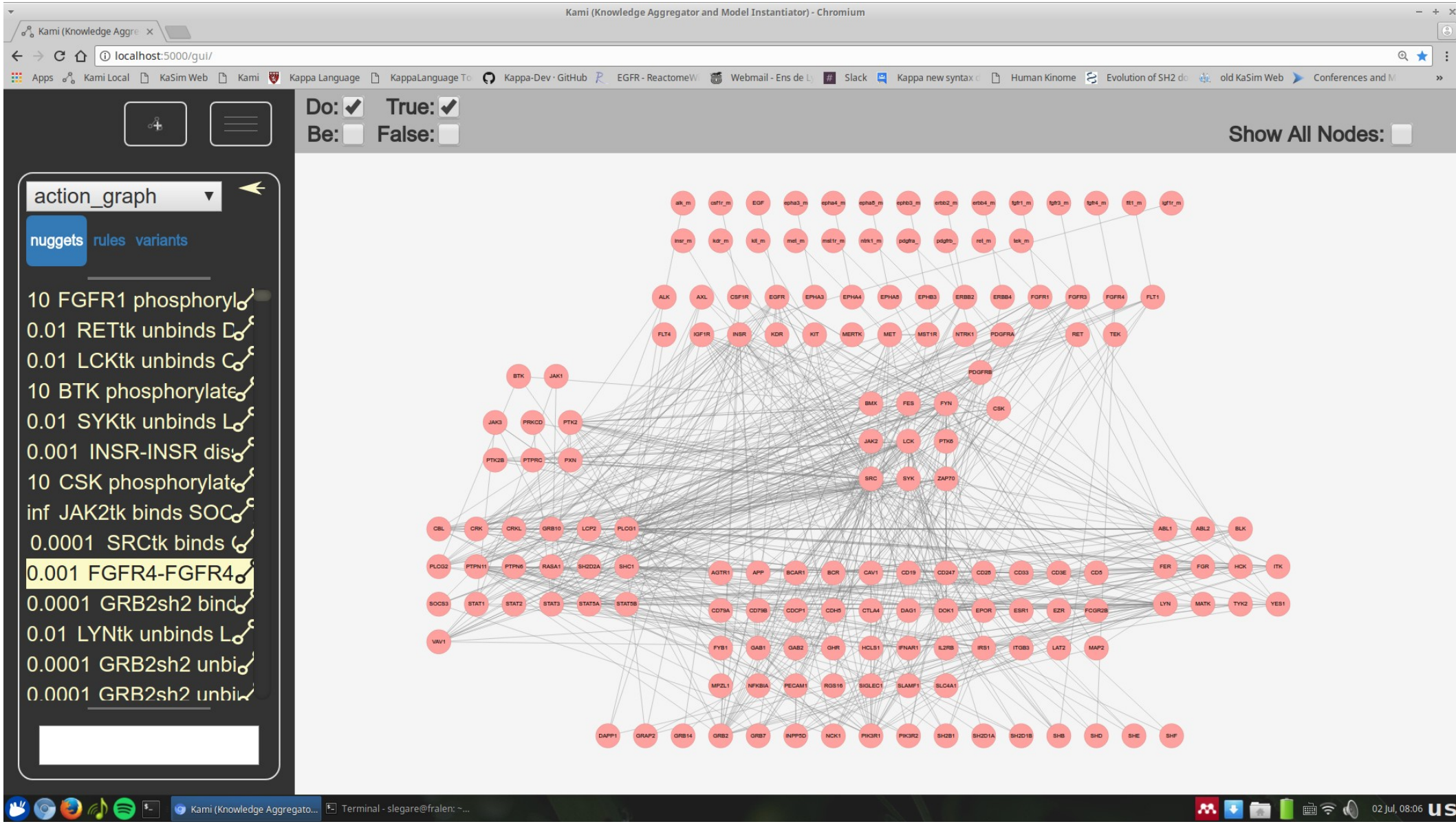
KAMI

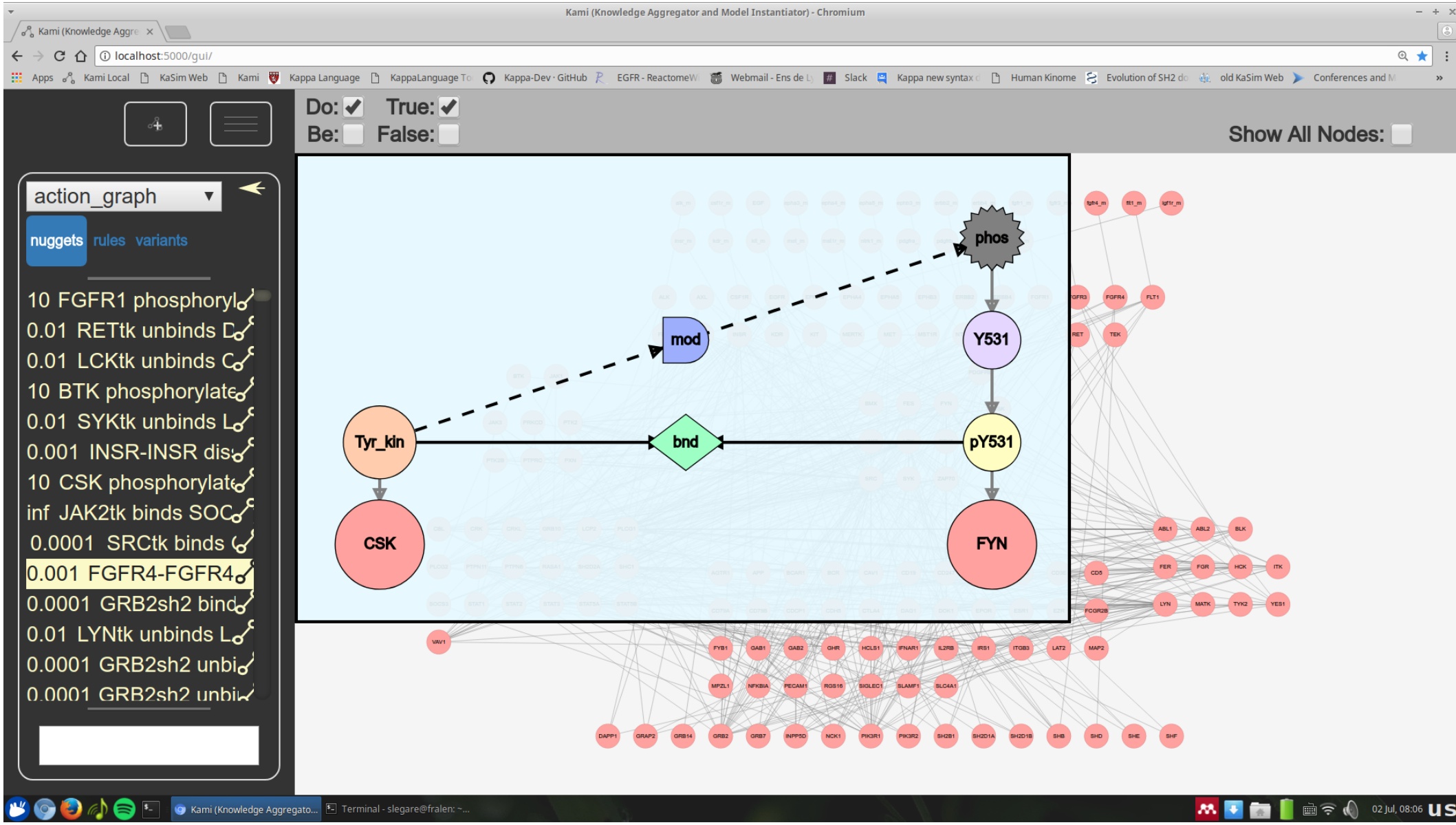
1) KAMIStudio (KAMI GUI)
2) Knowledge representation (Nugget / ActionGraph)
3) Building a model

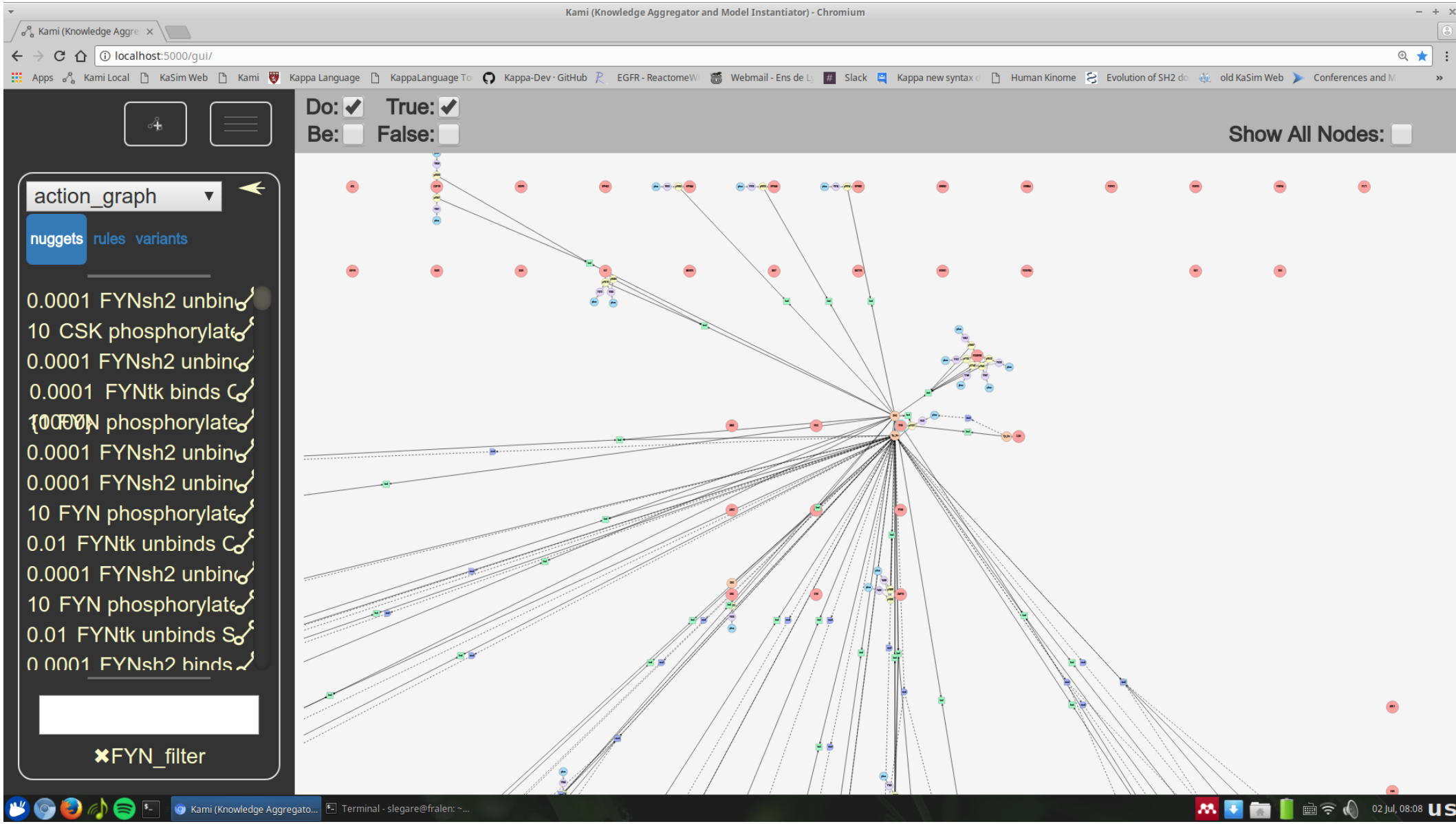
Kappa

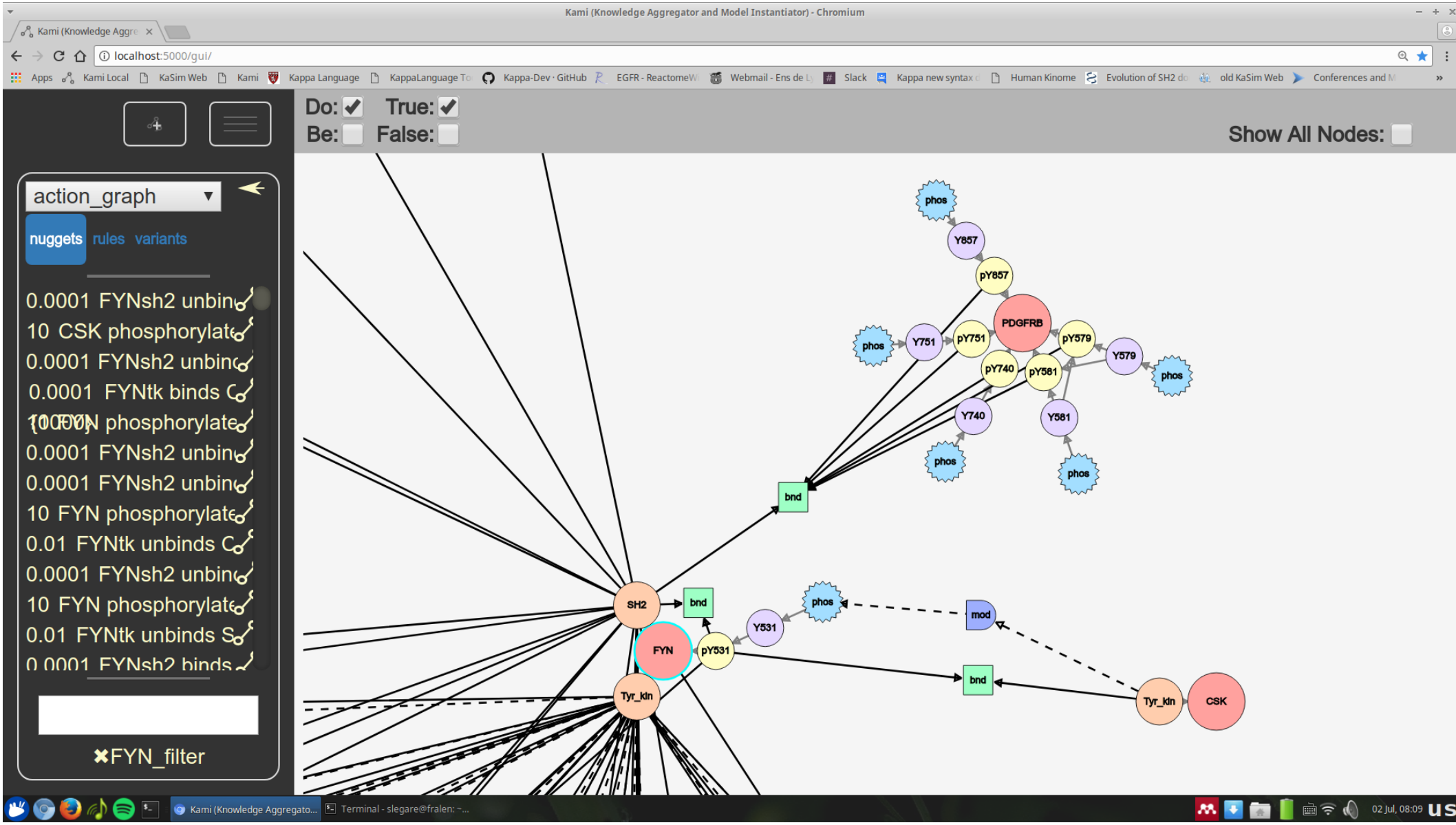
4) Rule-based executable model
5) Pathway discovery (Causality analysis)

6) The pYnet model

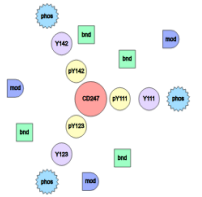

Kami (Knowledge Aggregator and Model Instantiator) - Chromium
localhost:5000/gui/
Do:
True:
Be:
False:
Show All Nodes:
action_graph
nuggets
rules
variants
10 FGFR1 phosphoryl
0.01 RETtk unbinds D
0.01 LCKtk unbinds C
10 BTK phosphorylate
0.01 SYKtk unbinds L
0.001 INSR-INSR dis
10 CSK phosphorylate
inf JAK2tk binds SOC
0.0001 SRCtk binds
0.001 FGFR4-FGFR4
0.0001 GRB2sh2 bind
0.01 LYNtk unbinds L
0.0001 GRB2sh2 unbi
0.0001 GRB2sh2 unbi
ALK
AXL
CSF1R
EGFR
EPHA3
EPHA4
EPHA5
EPHB3
ERBB2
ERBB4
FGFR1
FGFR3
FGFR4
FLT1
FLT4
IGF1R
INSR
KDR
KIT
MERTK
MET
MST1R
NTRK1
PDGFRA
RET
TEK
PDGFRB
BTK
JAK1
JAK3
PRKCD
PTK2
PTK2B
PTPRC
PXN
BMX
FES
FYN
CSK
JAK2
LCK
PTK6
SRC
SYK
ZAP70
CBL
CRK
CRKL
GRB10
LCP2
PLCG1
PLCG2
PTPN11
PTPN6
RASA1
SH2D2A
SHC1
SOCS3
STAT1
STAT2
STAT3
STAT5A
STAT5B
VAV1
ABL1
ABL2
BLK
FER
FGR
HCK
ITK
LYN
MATK
TYK2
YES1
AGTR1
APP
BCAR1
BCR
CAV1
CD19
CD247
CD28
CD33
CD3E
CD5
CD79A
CD79B
CDCP1
CDH5
CTLA4
DAG1
DOK1
EPOR
ESR1
EZR
FCGR2B
FYB1
GAB1
GAB2
GHR
HCLS1
IFNAR1
IL2RB
IRS1
ITGB3
LAT2
MAP2
MPZL1
NFKBIA
PECAM1
RGS16
SIGLEC1
SLAMF1
SLC4A1
DAPP1
GRAP2
GRB14
GRB2
GRB7
INPP5D
NCK1
PIK3R1
PIK3R2
SH2B1
SH2D1A
SH2D1B
SHB
SHD
SHE
SHF
Kami (Knowledge Aggregato...
Terminal - slegare@fralen: ~...
02 Jul, 08:06


Kami (Knowledge Aggregator and Model Instantiator) - Chromium
localhost:5000/gui/
Do:
True:
Be:
False:
Show All Nodes:
action_graph
nuggets
rules
variants
10 FGFR1 phosphoryl
0.01 RETtk unbinds D
0.01 LCKtk unbinds C
10 BTK phosphorylate
0.01 SYKtk unbinds L
0.001 INSR-INSR dis
10 CSK phosphorylate
inf JAK2tk binds SOC
0.0001 SRCtk binds
0.001 FGFR4-FGFR4
0.0001 GRB2sh2 bind
0.01 LYNtk unbinds L
0.0001 GRB2sh2 unbi
0.0001 GRB2sh2 unbi
phos
mod
Y531
Tyr_kin
bnd
pY531
CSK
FYN
Kami (Knowledge Aggregato...
Terminal - slegare@fralen: ~...
02 Jul, 08:06


Kami (Knowledge Aggregator and Model Instantiator) - Chromium
localhost:5000/gui/
Do:
True:
Be:
False:
Show All Nodes:
action_graph
nuggets
rules
variants
0.0001 FYNsh2 unbin
10 CSK phosphorylat
0.0001 FYNsh2 unbin
0.0001 FYNtk binds C
10 FYN phosphorylate
0.0001 FYNsh2 unbin
0.0001 FYNsh2 unbin
10 FYN phosphorylate
0.01 FYNtk unbinds C
0.0001 FYNsh2 unbin
10 FYN phosphorylate
0.01 FYNtk unbinds S
FYN_filter
Kami (Knowledge Aggregato...
Terminal - slegare@fralen: ~...
02 Jul, 08:08


Kami (Knowledge Aggregator and Model Instantiator) - Chromium
localhost:5000/gui/
Do:
True:
Be:
False:
Show All Nodes:
action_graph
nuggets
rules
variants
0.0001 FYNsh2 unbin
10 CSK phosphorylat
0.0001 FYNsh2 unbin
0.0001 FYNtk binds C
10 FYN phosphorylate
0.0001 FYNsh2 unbin
0.0001 FYNsh2 unbin
10 FYN phosphorylate
0.01 FYNtk unbinds C
0.0001 FYNsh2 unbin
10 FYN phosphorylate
0.01 FYNtk unbinds S
FYN_filter
phos
Y857
pY857
PDGFRB
Y751
pY751
pY579
Y579
pY740
pY581
Y740
Y581
bnd
SH2
FYN
Tyr_kin
pY531
Y531
mod
CSK
Kami (Knowledge Aggregato...
Terminal - slegare@fralen: ~...
02 Jul, 08:09

Knowledge representation



Split in two layers, the nuggets and action graph

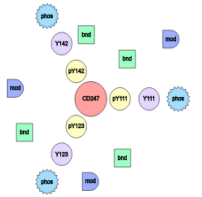
- Unambiguously specify interactions
- Limited set of symbols

Can read a model using KAMIStudio

Every nugget is independent

- Facilitates incremental aggregation

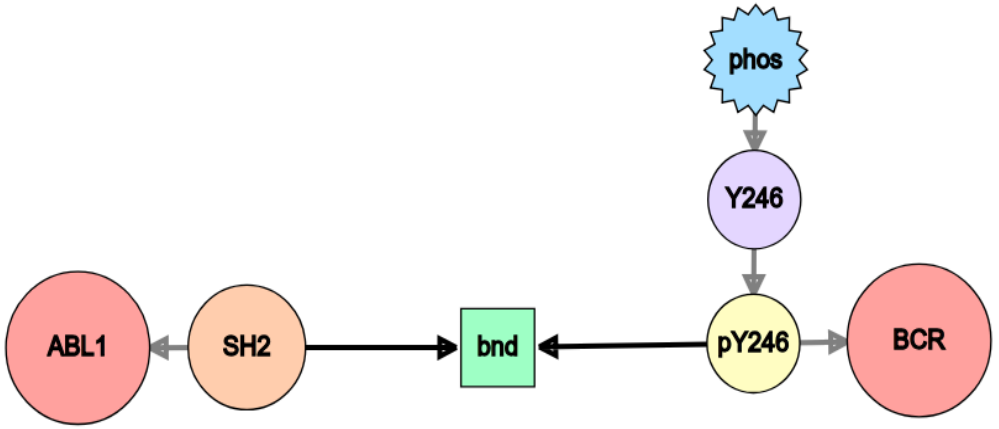
Building a KAMI model



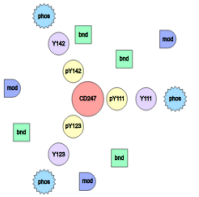
KAMI interaction (programmatic)

```
Binding(
  RegionActor(
    gene=Gene(uniprot_ac="P00519",
      hgnc_symbol="ABL1"),
    region=Region(name="SH2",
      interproid="IPR000980")
  ),
  SiteActor(
    gene=Gene(uniprot_ac="P11274",
      hgnc_symbol="BCR"),
    site=Site(name="pY246",
    residues=[Residue(aa="Y", loc=246,
    state=State("phosphorylation", True))])
  ),
  rate=0.001,
  desc="ABL1 binds BCR-Y246"
)
```

KAMI nugget (graph representation)



How does it work?

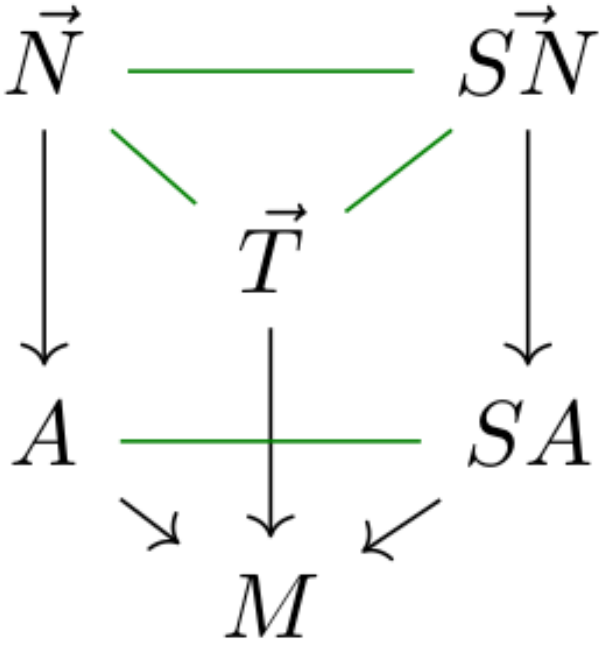


$\vec{N}$: Nugget

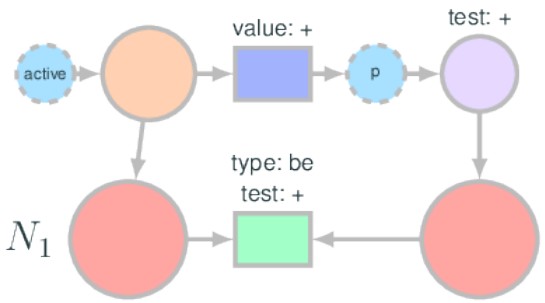
A : Action Graph

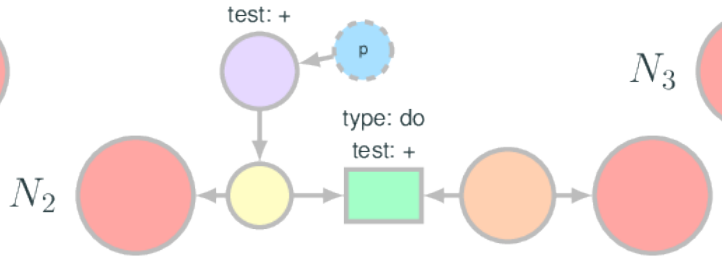
Arrow : Typing

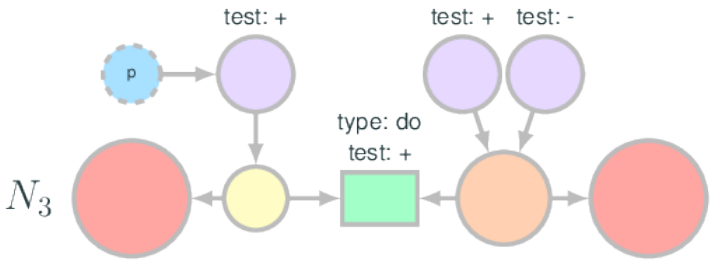
KAMI graph hierarchy

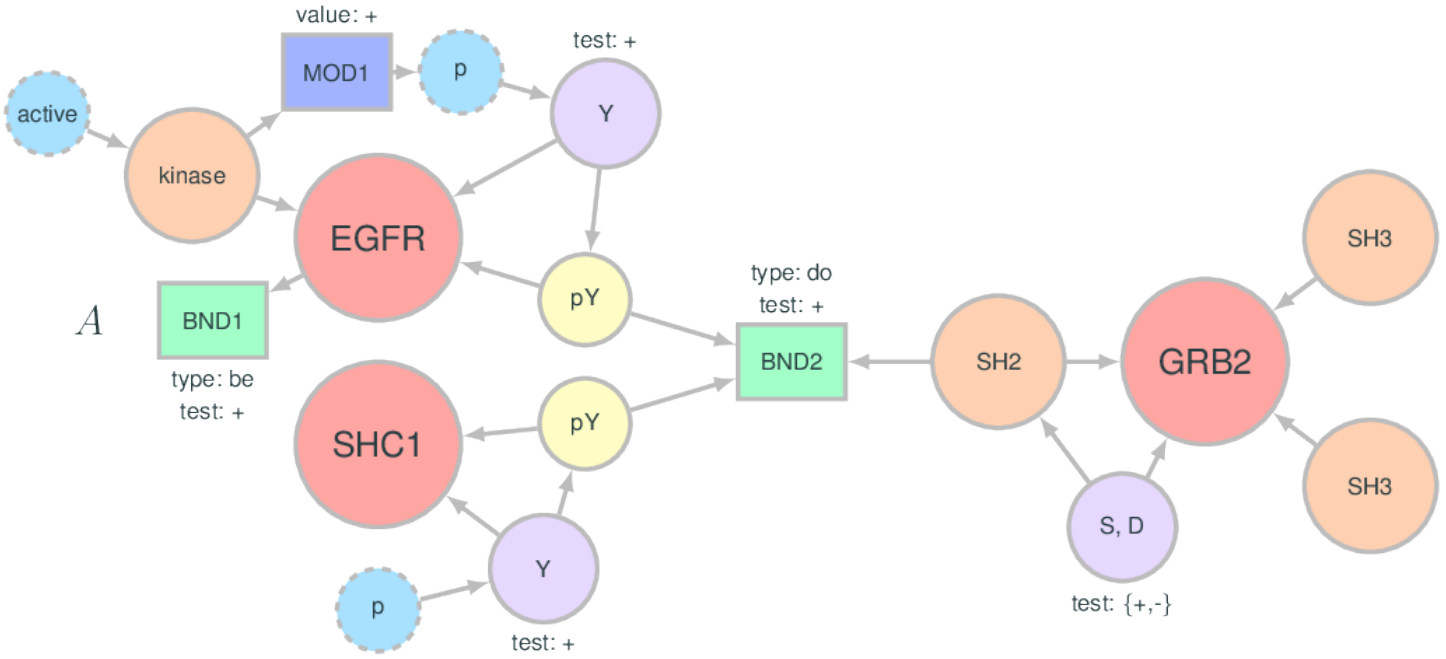


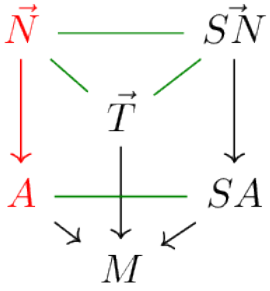
Typing


active
value: +
p
test: +
type: be
test: +
N_1

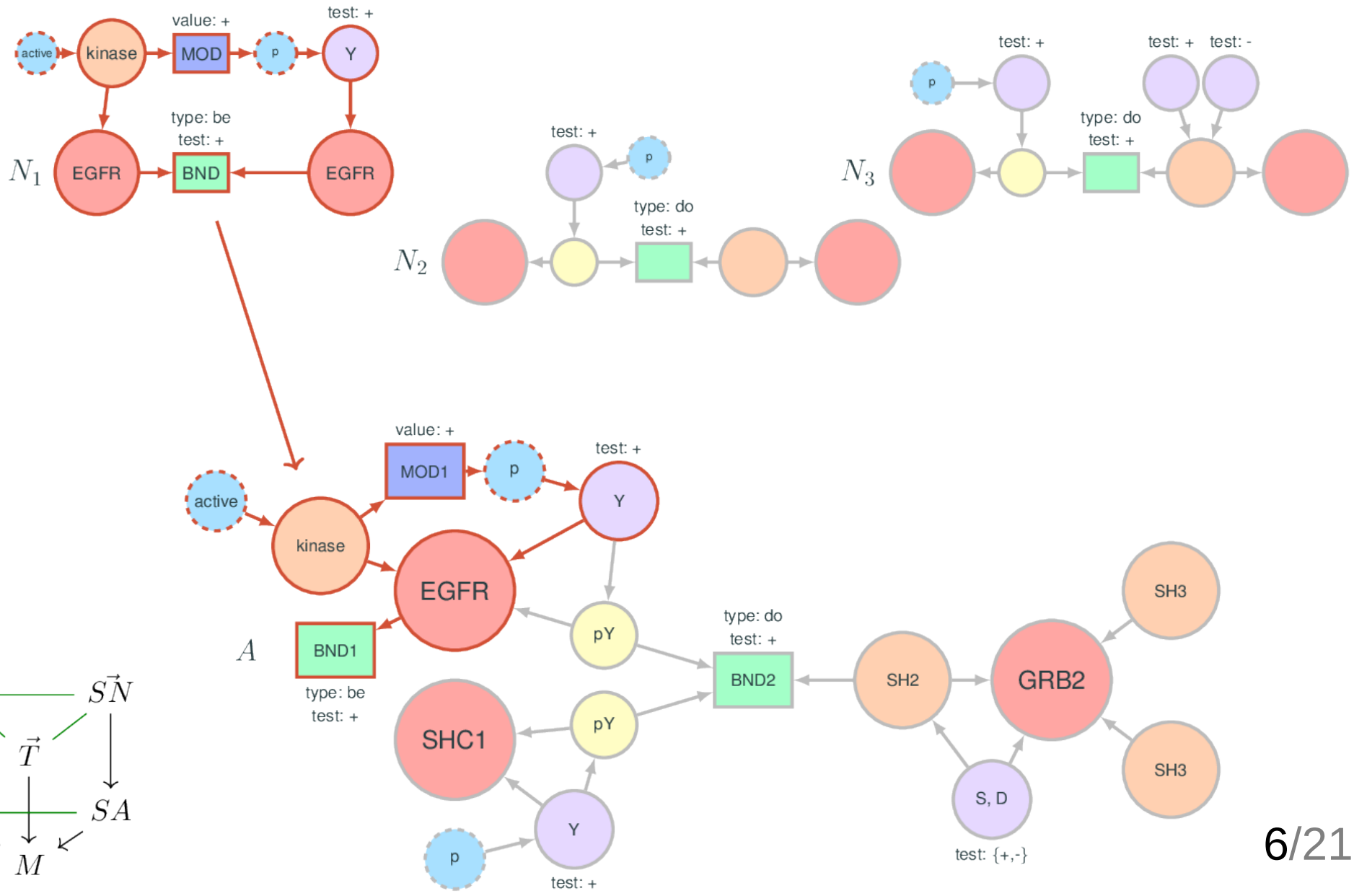

test: +
p
type: do
test: +
N_2


test: +
p
test: +
test: -
type: do
test: +
N_3

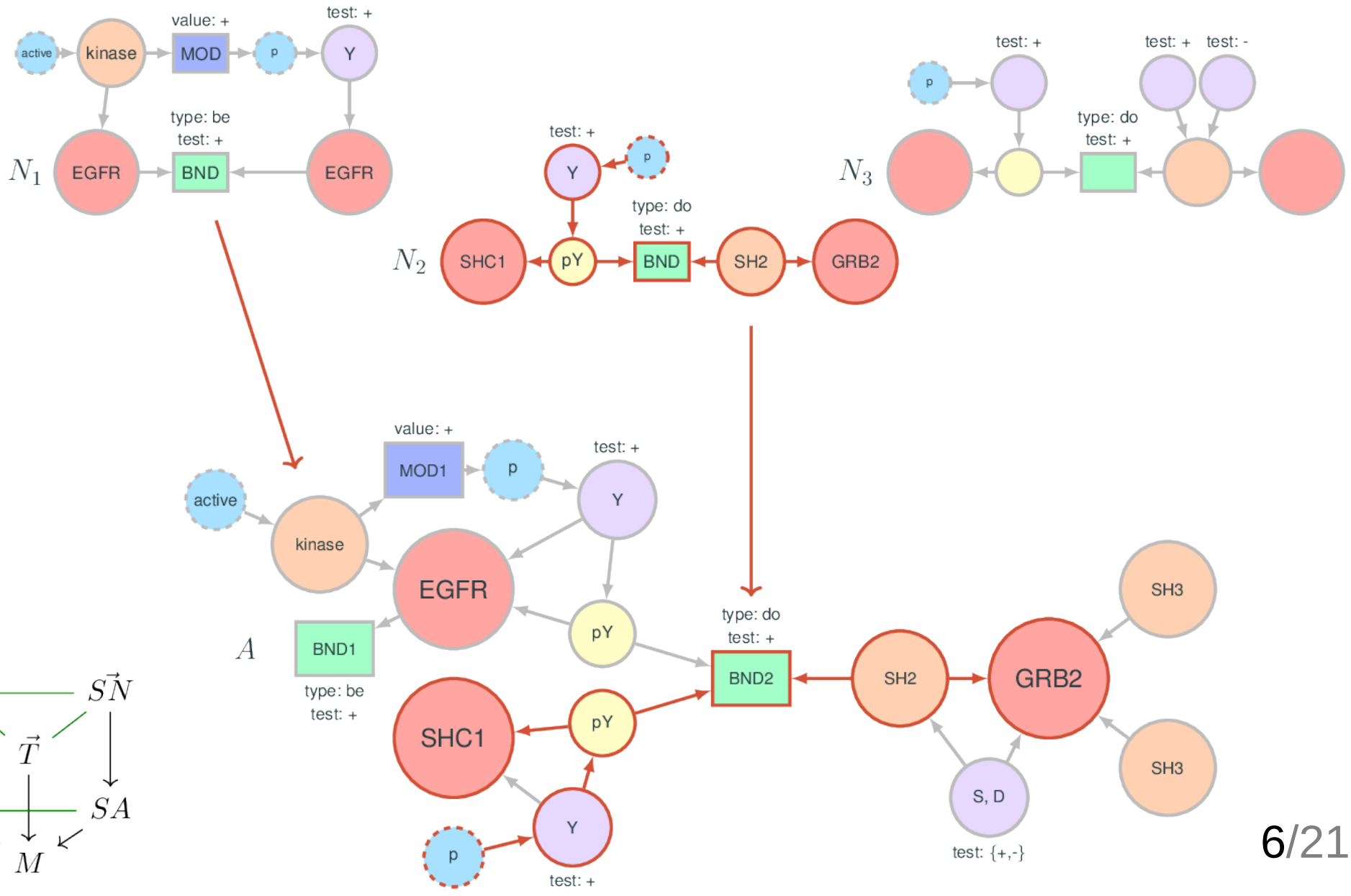

value: +
MOD1
p
test: +
Y
active
kinase
EGFR
pY
type: do
test: +
BND2
SH2
GRB2
SH3
SH3
A
BND1
type: be
test: +
SHC1
pY
S, D
test: {+,-}
Y
p
test: +


$\vec{N}$
$S\vec{N}$
$\vec{T}$
A
SA
M

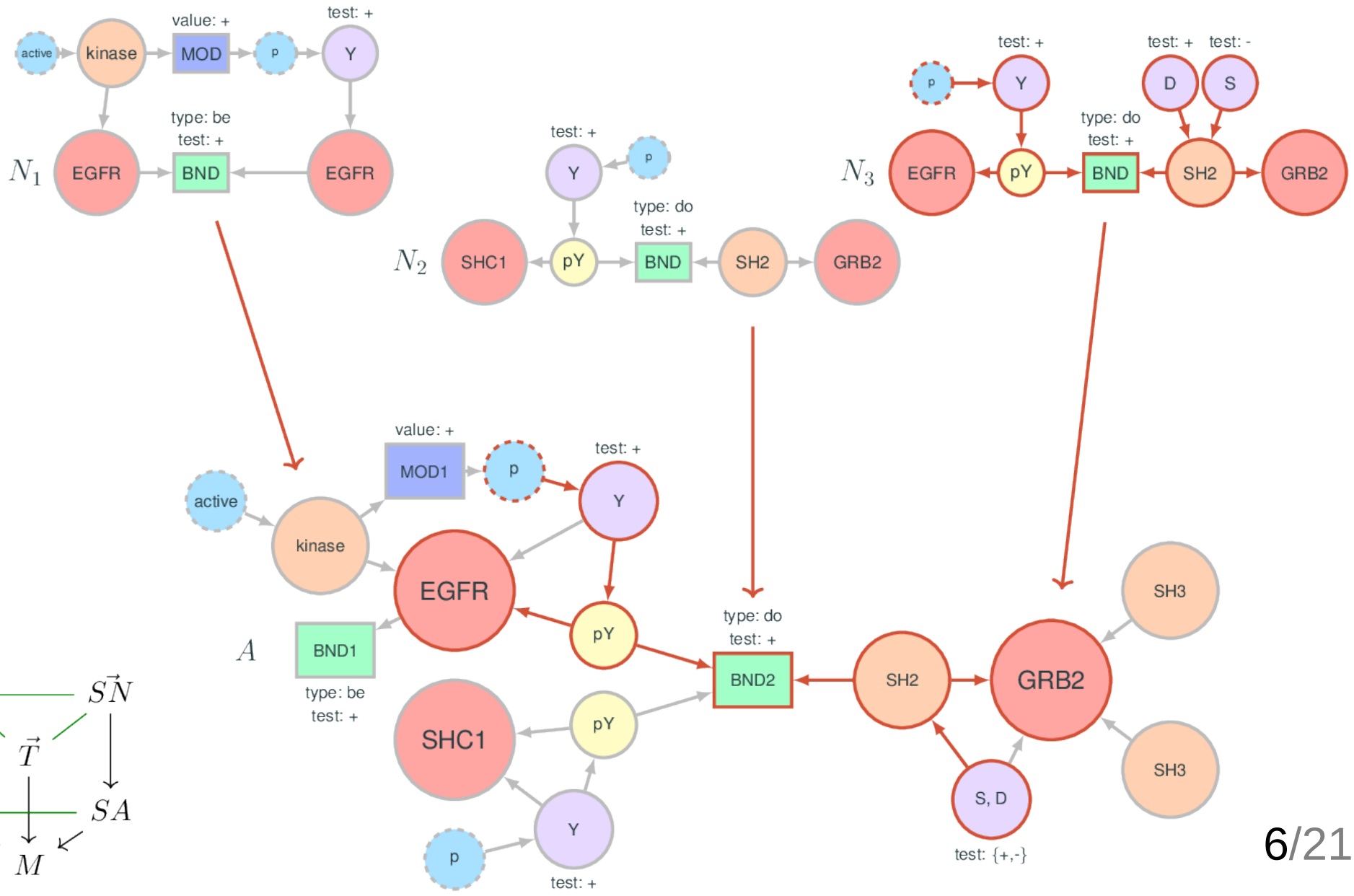
Typing


active
kinase
value: +
MOD
p
test: +
Y
N_1
EGFR
type: be
test: +
BND
EGFR
test: +
p
type: do
test: +
N_2
test: +
test: +
test: -
p
type: do
test: +
N_3
value: +
MOD1
p
test: +
Y
active
kinase
EGFR
pY
BND1
type: be
test: +
A
type: do
test: +
BND2
SH2
GRB2
SH3
SH3
S, D
test: {+,-}
SHC1
pY
Y
p
test: +
$\vec{N}$
$S\vec{N}$
$\vec{T}$
A
SA
M

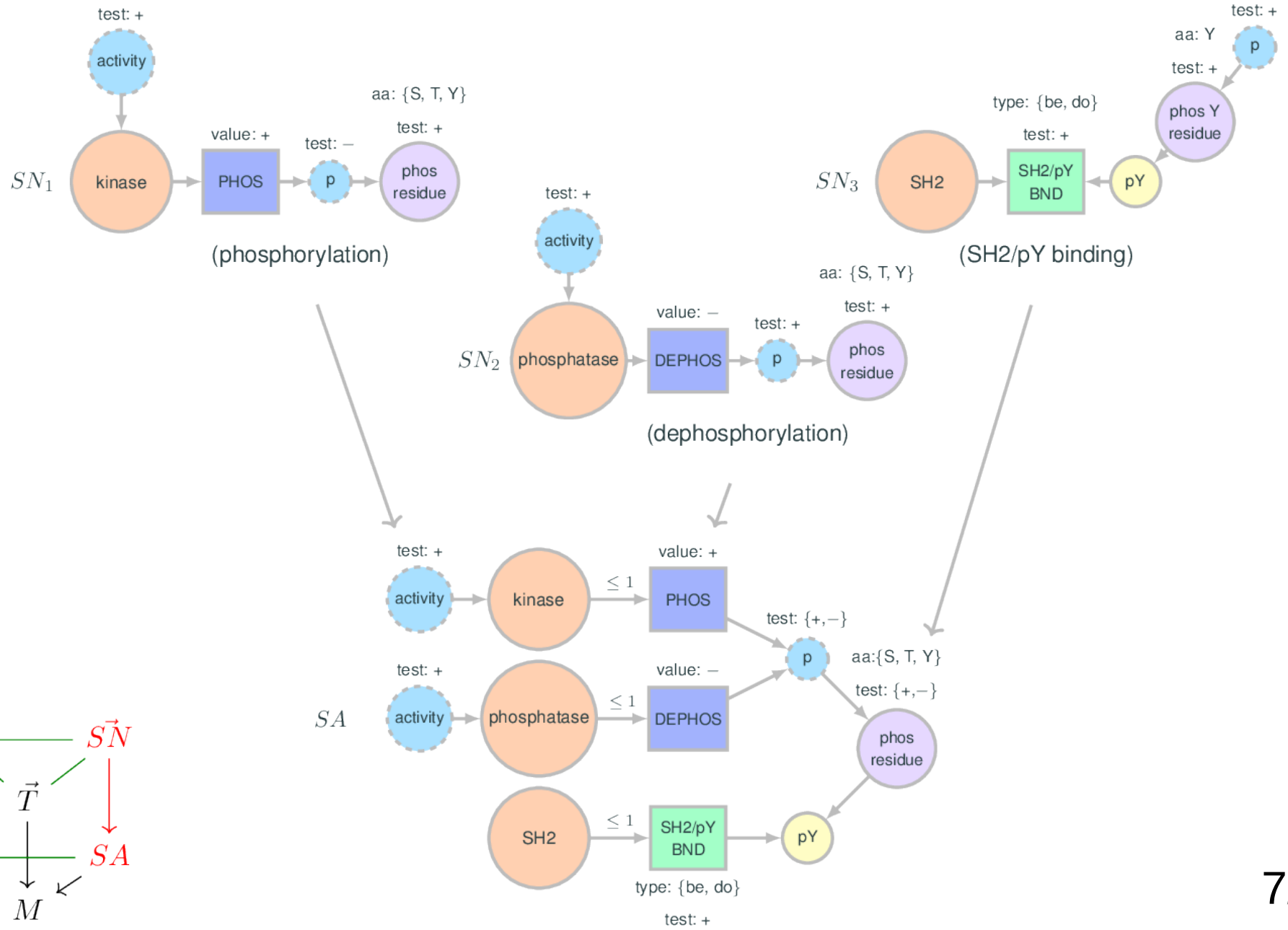
Typing

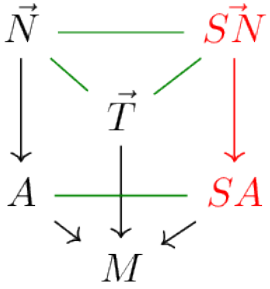

active
kinase
value: +
MOD
p
test: +
Y
N_1
EGFR
type: be
test: +
BND
EGFR
test: +
p
Y
type: do
test: +
N_2
SHC1
pY
BND
SH2
GRB2
N_3
test: +
p
type: do
test: +
test: +
test: -
value: +
MOD1
p
test: +
Y
active
kinase
EGFR
pY
A
BND1
type: be
test: +
type: do
test: +
BND2
SH2
GRB2
SH3
SH3
SHC1
pY
S, D
test: {+,-}
Y
p
test: +
$\vec{N}$
$S\vec{N}$
$\vec{T}$
A
SA
M

Typing

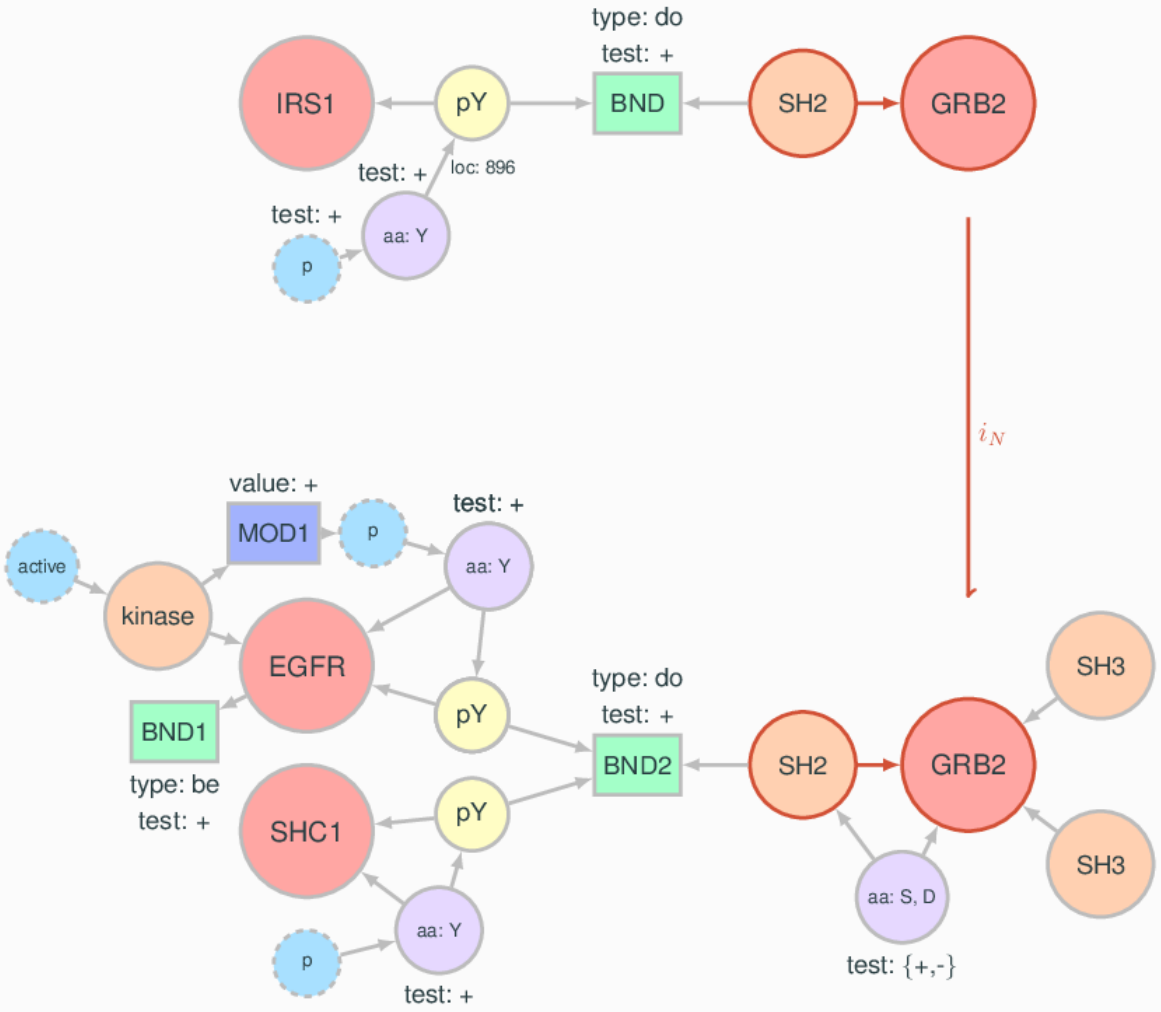

active
kinase
value: +
MOD
p
test: +
Y
N_1
EGFR
type: be
test: +
BND
EGFR
test: +
Y
p
N_2
SHC1
pY
type: do
test: +
BND
SH2
GRB2
test: +
p
Y
test: +
test: -
D
S
N_3
EGFR
pY
type: do
test: +
BND
SH2
GRB2
value: +
MOD1
p
test: +
Y
active
kinase
EGFR
pY
A
BND1
type: be
test: +
type: do
test: +
BND2
SH2
GRB2
SH3
SH3
SHC1
pY
S, D
test: {+,-}
Y
p
test: +
$\vec{N}$
$S\vec{N}$
$\vec{T}$
A
SA
M

Semantic nuggets

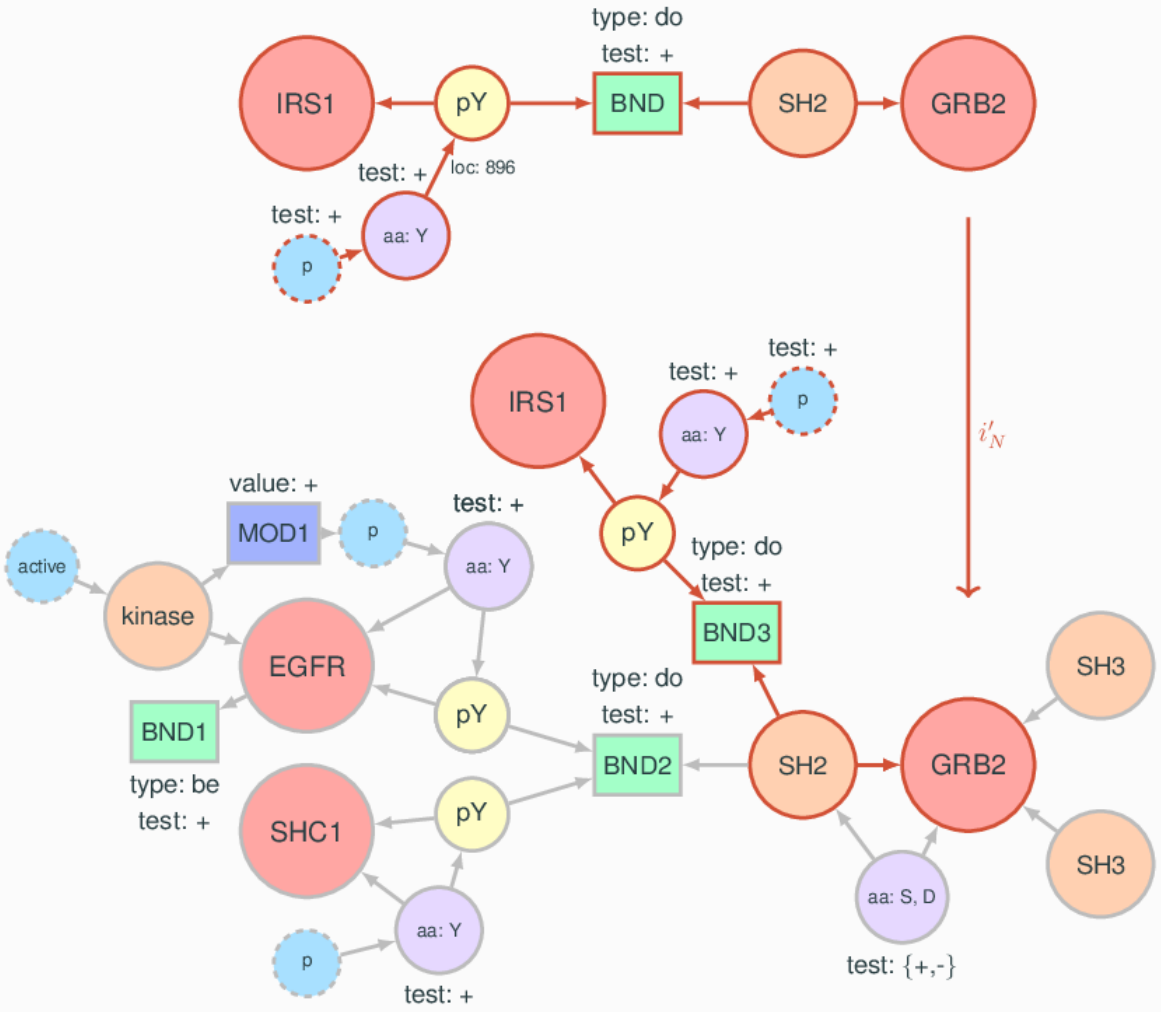




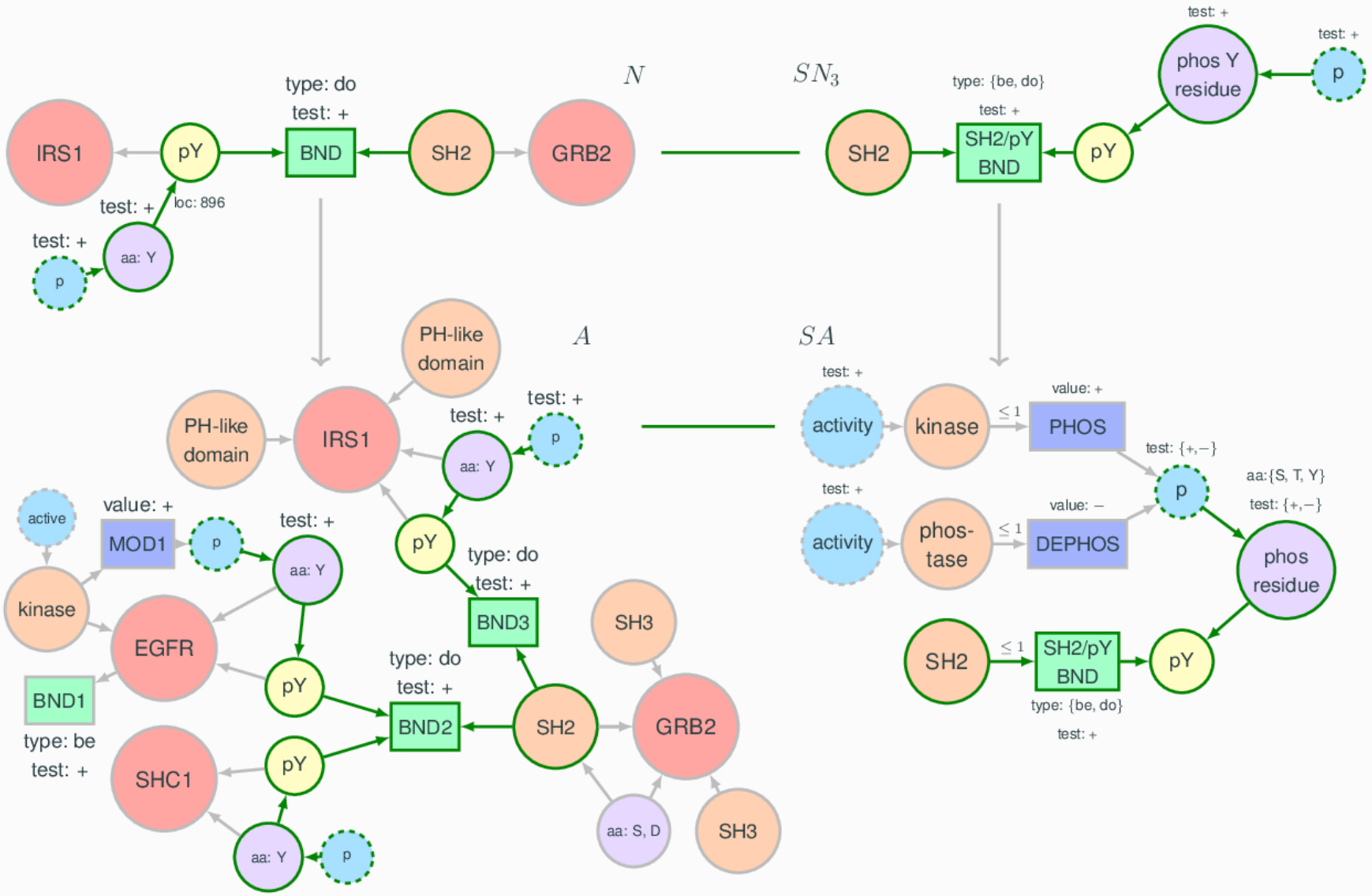
Knowledge aggregation



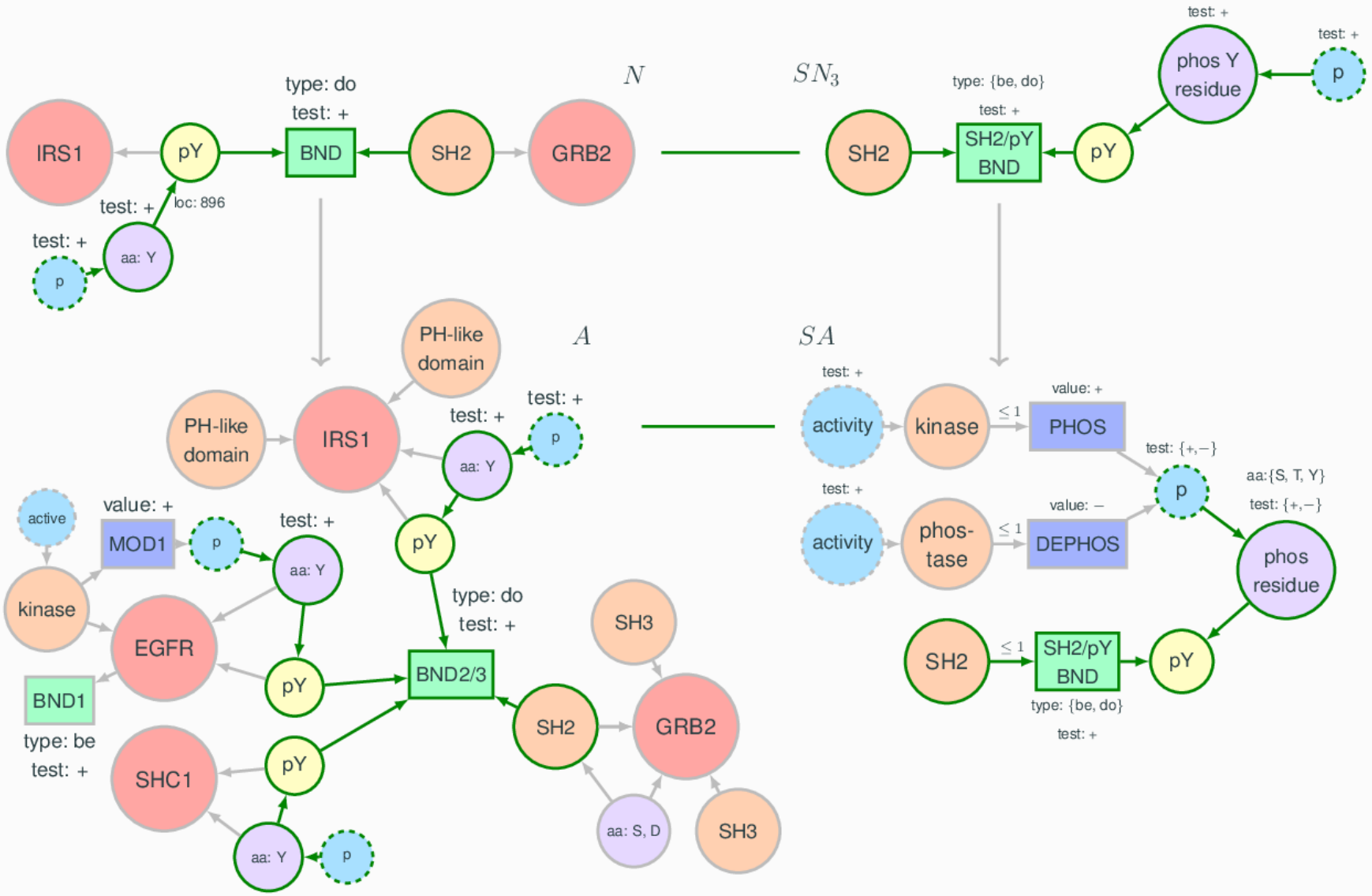
Knowledge aggregation



Knowledge aggregation (continued)



Knowledge aggregation (continued)

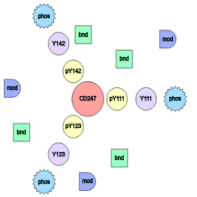


Biocuration with KAMI

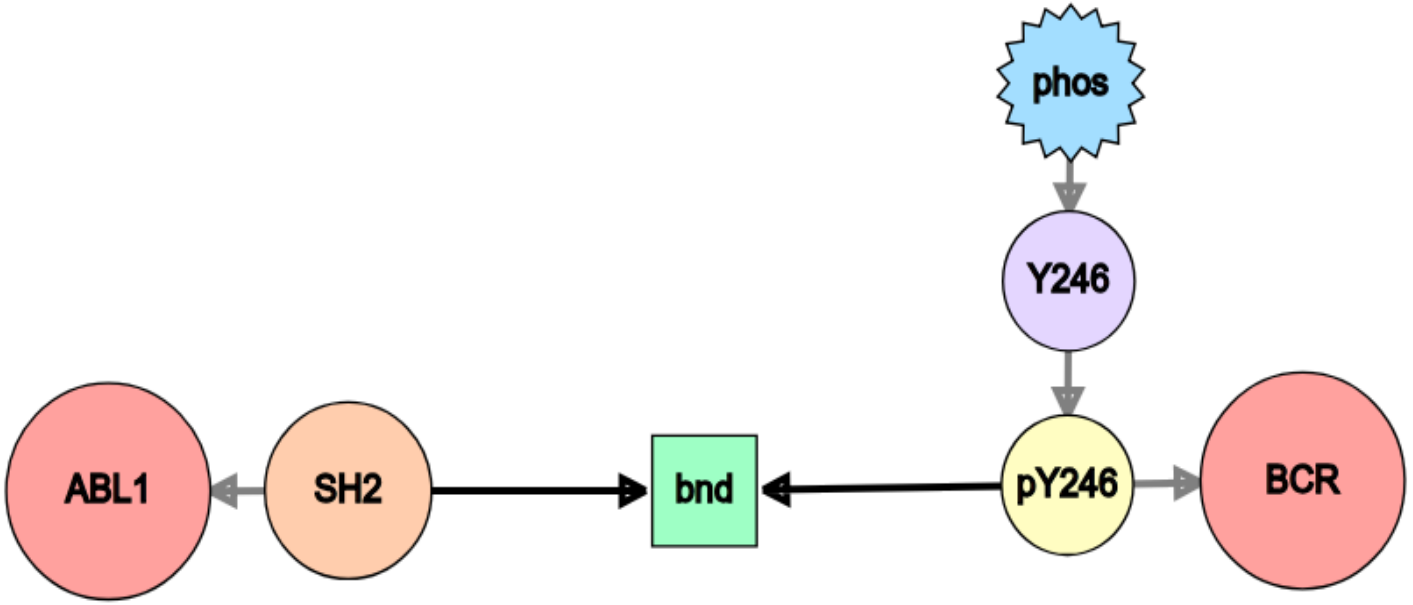
Allows users to easily add new data

- Detects elements already present in action graph
- Semantic checks
- Completes interaction if more detailed
- Ignores new data if it already exists

Kappa rule-based model

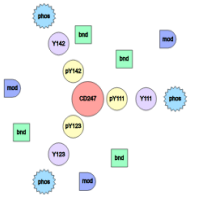


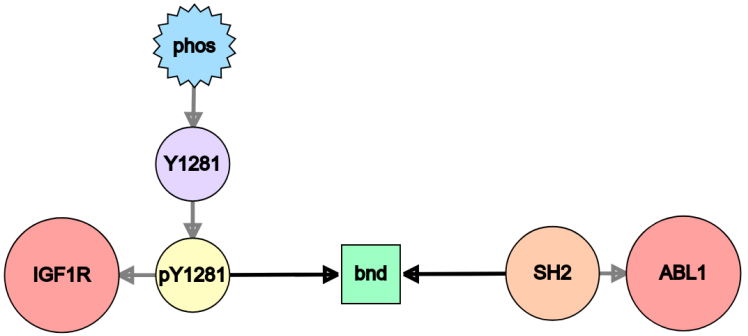
KAMI nugget

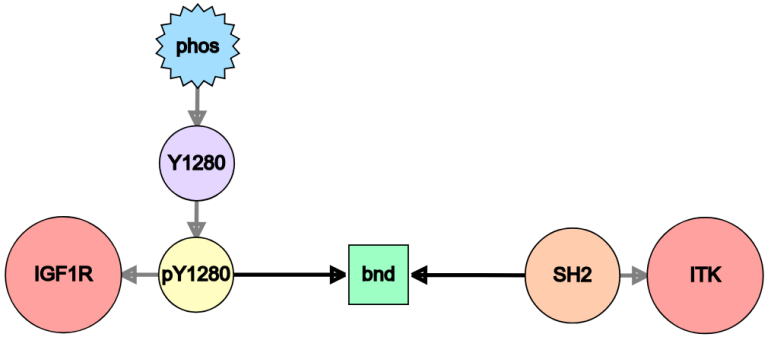


Kappa rule **ABL1(SH2[./1]), BCR(pY246[./1] Y246_phos{True})**

1 nugget ≈ 1 rule (info from AG)



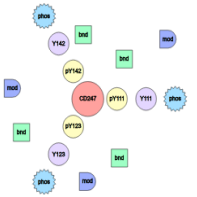




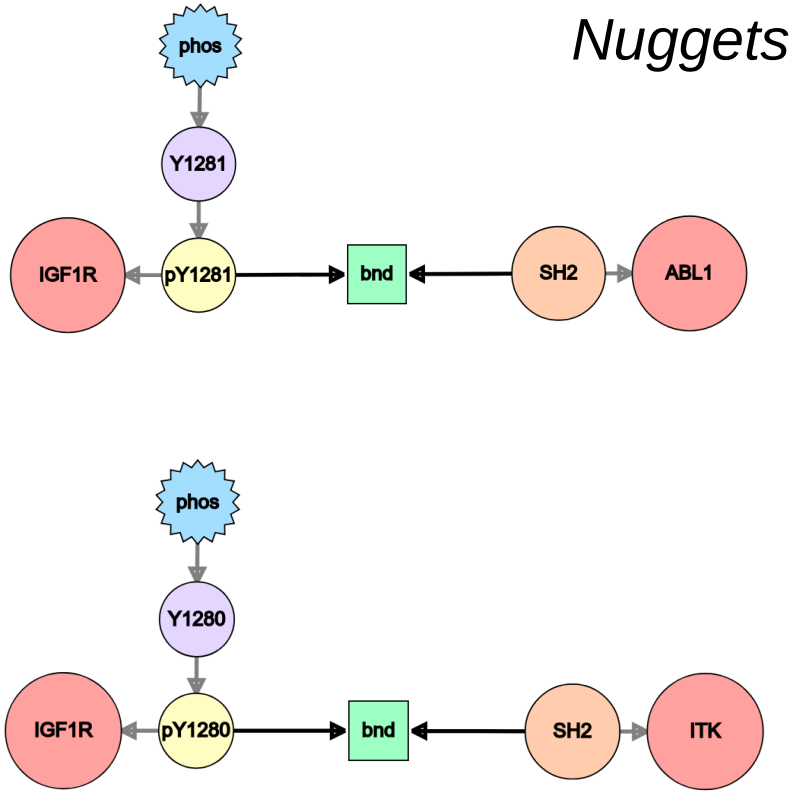

```
IFGR1(pY1281[./1] Y1281_phos{True}), ABL1(SH2[./1])

IFGR1(pY1280[./1] Y1280_phos{True}), ITK(SH2[./1])
```

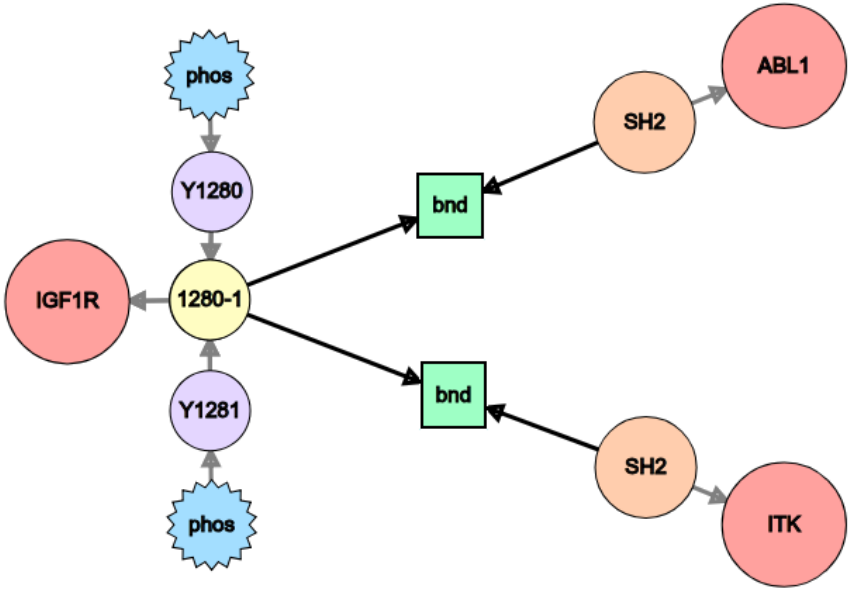
1 nugget ≈ 1 rule (info from AG)



Nuggets



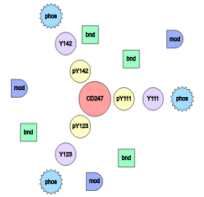
Action Graph

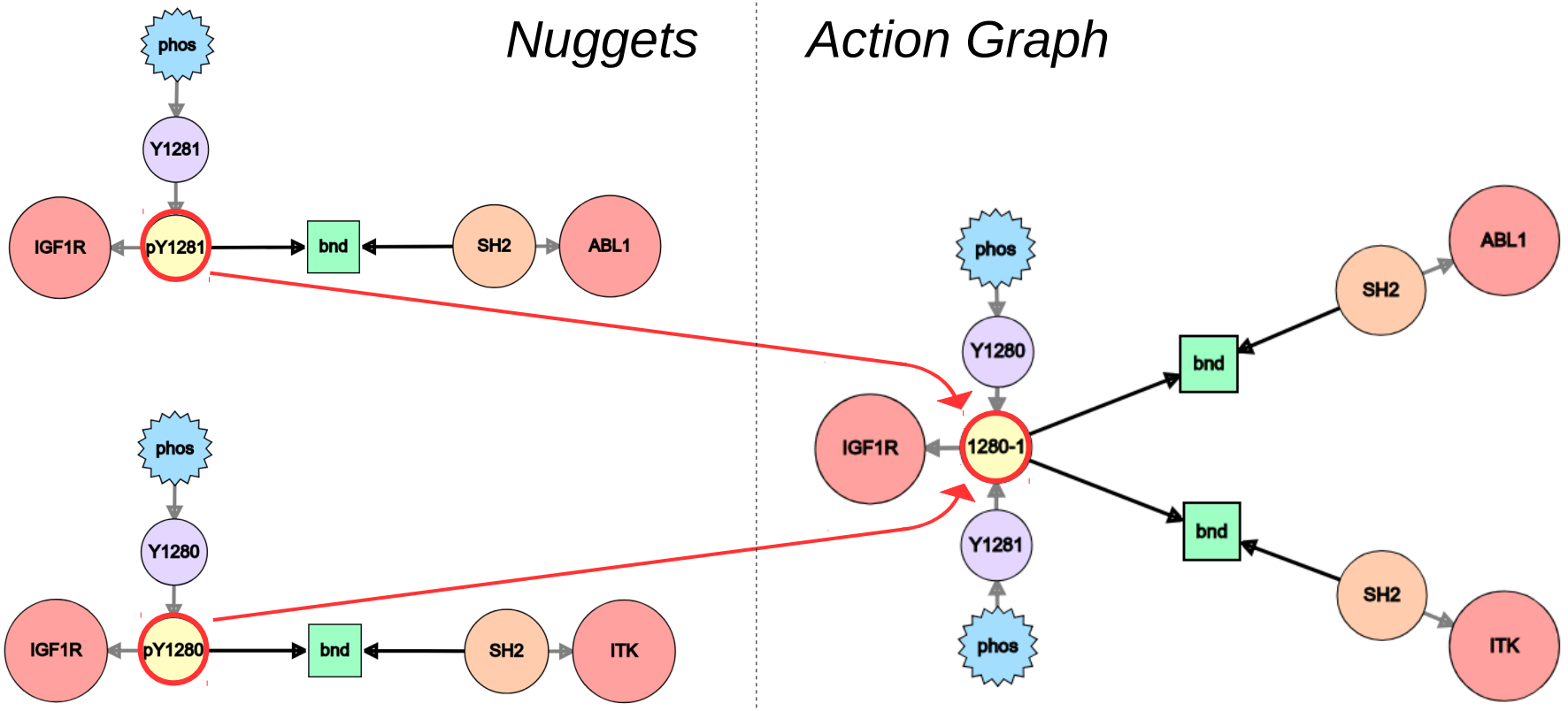



```
IFGR1(pY1281[./1] Y1281_phos{True}), ABL1(SH2[./1])

IFGR1(pY1280[./1] Y1280_phos{True}), ITK(SH2[./1])
```

1 nugget ≈ 1 rule (info from AG)



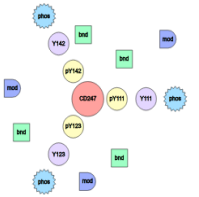


IFGR1(**pY1280-1[./1]** Y1281_phos{True}), ABL1(SH2[./1])

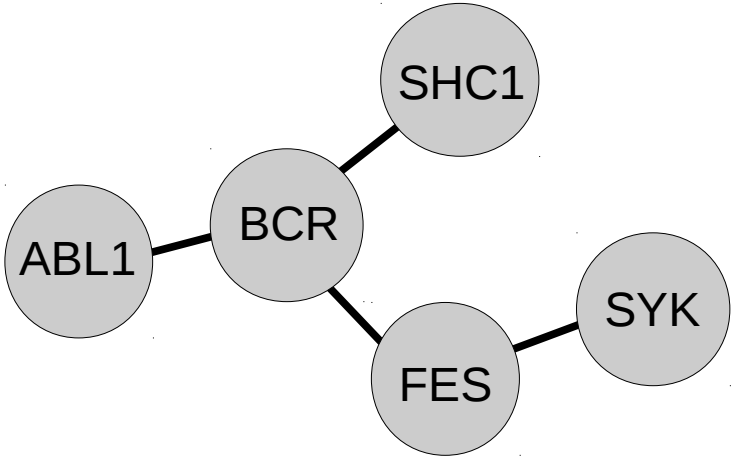
||

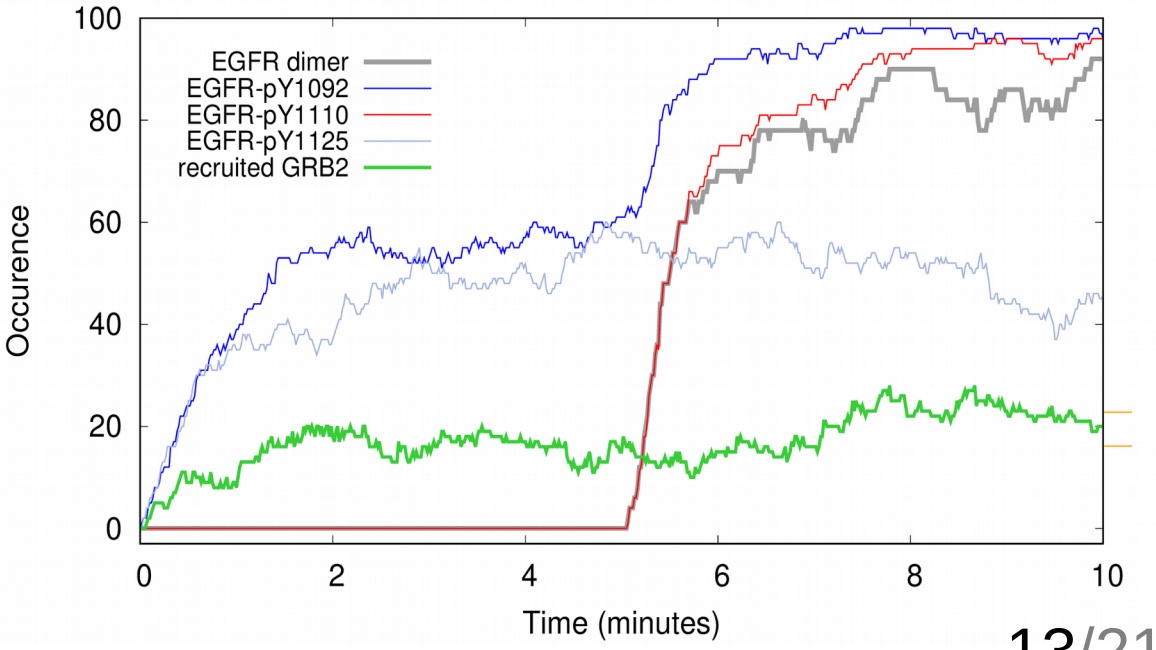
IFGR1(**pY1280-1[./1]** Y1280_phos{True}), ITK(SH2[./1])

Simulations with KaSim

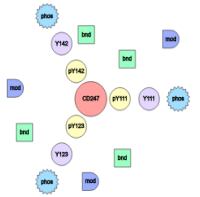


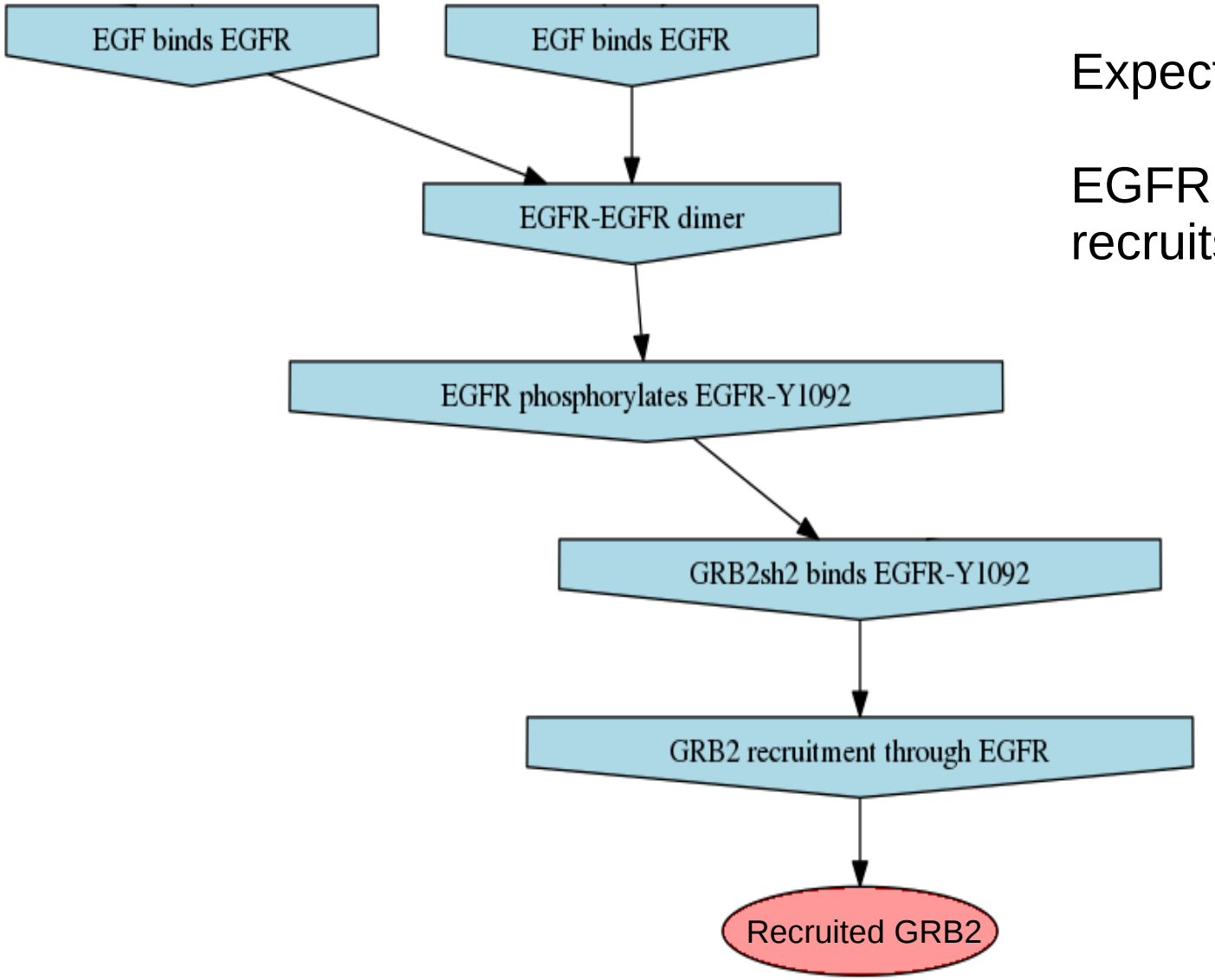
- Rule-based
- Deals with combinatorial complexity
- Quantitative
- Stochastic
- No spatial dimension
- Analysis system dynamics





Causality Analysis (KaStor)

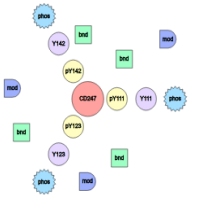


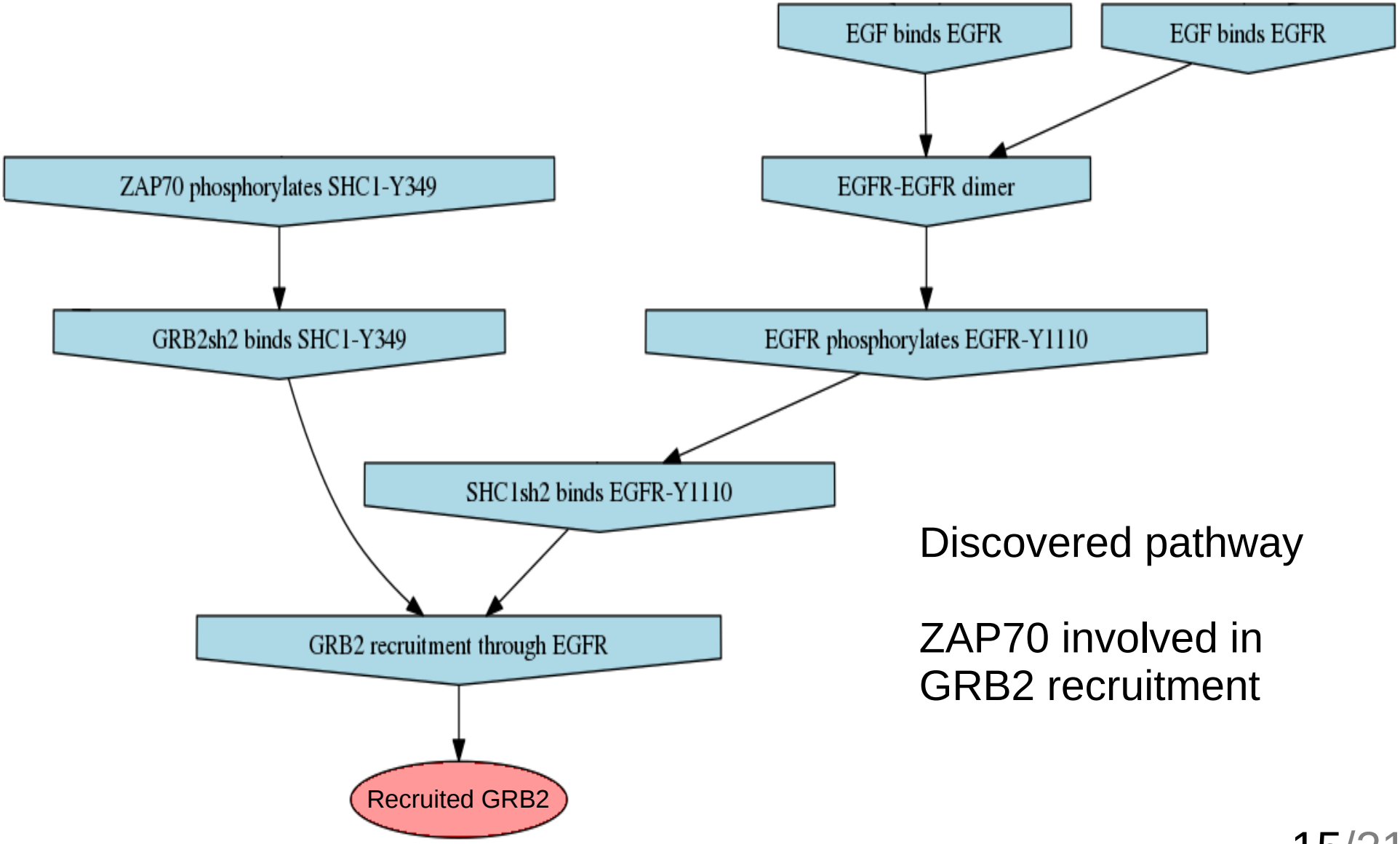


Expected pathway

EGFR dimerizes and recruits GRB2

Causality Analysis (KaStor)

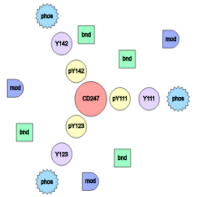




Discovered pathway

ZAP70 involved in GRB2 recruitment

The pYnet model



Cell signaling

- Tyrosine phosphorylation
- SH2 domain bindings

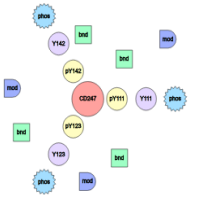
900 interactions extracted from

- PhosphoSite
- Phospho.ELM
- NCI Pathway Interaction Database

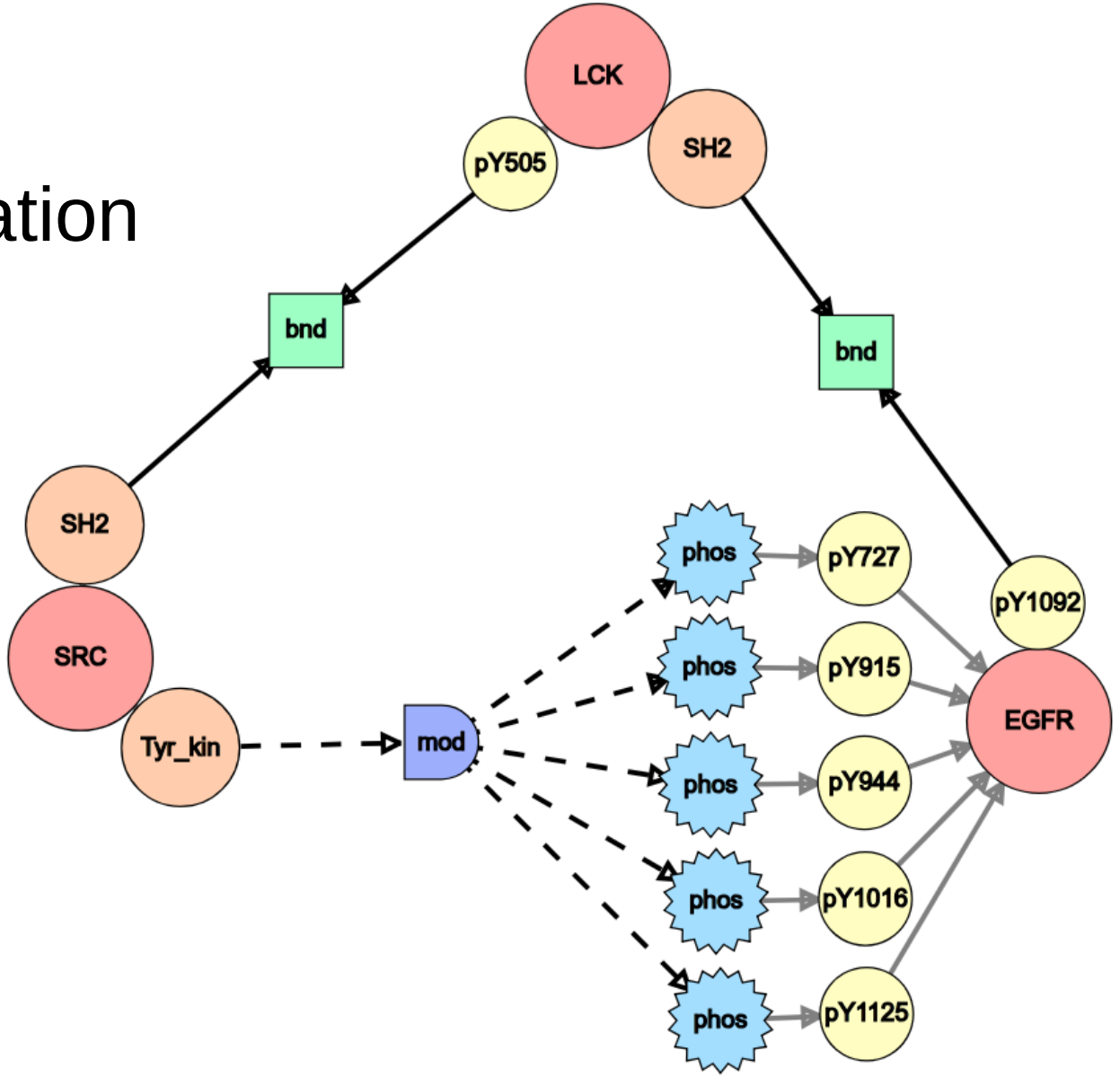
Well suited to showcase rule-based modelling

- Combinatorial complexity
- Large
- Scaffolding

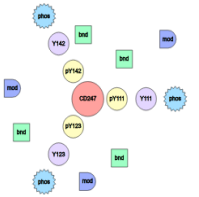
Important combinatorial compl.



Processive phosphorylation



Summary



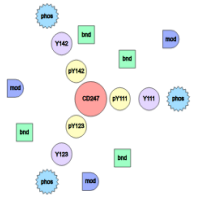
KAMI allows

- Representation of individual interactions
- Aggregation into an interaction network

KAMI works with Kappa to

- Produce dynamic simulations
- Discover pathway using causality analysis

Summary

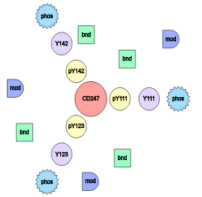


Fundamentally different approach to modelling

- No need to know exactly where new data fits
- Can just “smash” interactions together

- No need to explicitly build the pathways (bias)
- Can discover the pathways through analysis

Using KAMI and Kappa



In development

KAMI: `github.com/Kappa-Dev/KAMI`

KAMIStudio: `github.com/Kappa-Dev/KAMIStudio`

Graph rewriting

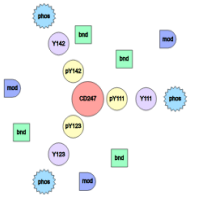
ReGraph: `github.com/Kappa-Dev/ReGraph`

Kappa

KaSim: `github.com/Kappa-Dev/KaSim.git`

Web Site: `kappalanguage.org`

Thanks



Russ Harmer

Eugenia Oshurko

ENS DE LYON

Jean Krivine

INSTITUT DE RECHERCHE EN INFORMATIQUE FONDAMENTALE

Jérôme Feret

Vincent Danos

Yves-Stan Le Cornec

1794 ENS ÉCOLE NORMALE SUPÉRIEURE

Walter Fontana

Pierre Boutillier

Hector Medina

Ioana Cristescu

VE RI TAS HARVARD MEDICAL SCHOOL