

# *PA054: Formální modely v systémové biologii*

David Šafránek

2.5.2012

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



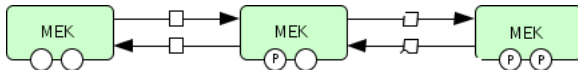
# *Obsah*

*Algebraický přístup k modelování biochemických reakcí*

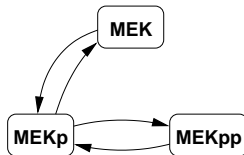
# Princip modelování

## Substráty – molekuly

- populace interagujících molekul
- každou molekulu (proteinu) lze popsat stavově
  - stav zachycuje konfiguraci vazebných míst
  - volná vazebná místa
  - vazebná místa obsazena jinou molekulou (dimerizující stavy)



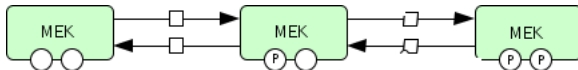
- formalizace prostřednictvím sekvenčního procesu:



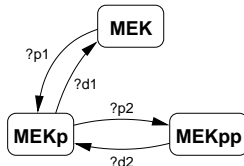
# Princip modelování

## Substráty – molekuly

- populace interagujících molekul
- každou molekulu (proteinu) lze popsat stavově
  - stav zachycuje konfiguraci vazebných míst
  - volná vazebná místa
  - vazebná místa obsazena jinou molekulou (dimerizující stavy)



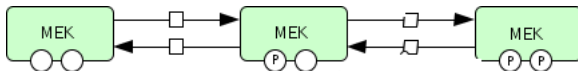
- formalizace prostřednictvím sekvenčního procesu:



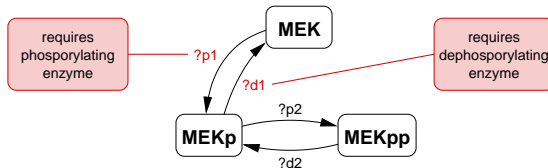
# Princip modelování

## Substráty – molekuly

- populace interagujících molekul
- každou molekulu (proteinu) lze popsat stavově
  - stav zachycuje konfiguraci vazebných míst
  - volná vazebná místa
  - vazebná místa obsazena jinou molekulou (dimerizující stavy)

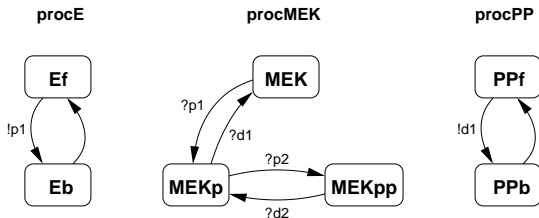


- formalizace prostřednictvím sekvenčního procesu:



# Princip modelování

## Interakce

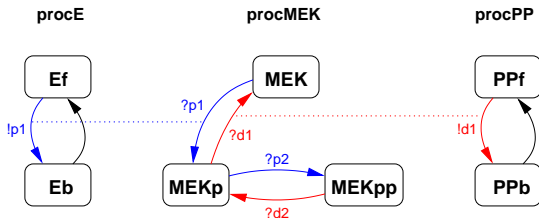


- paralelní chování procesů (molekul)
- binární synchronizace
- roztok lze modelovat jako populaci procesů:

$$procMEK | procMEK | procE | procE | procE | procPP | procPP$$

# Princip modelování

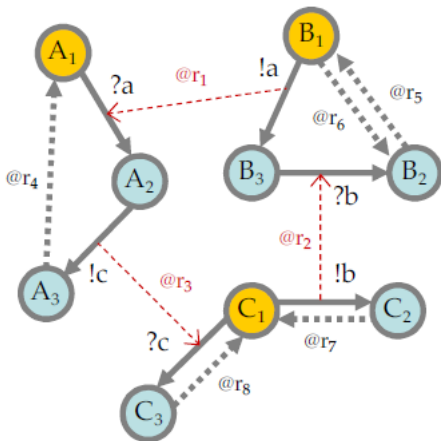
## Interakce



- paralelní chování procesů (molekul)
- binární synchronizace
- roztok lze modelovat jako populaci procesů:

$$procMEK | procMEK | procE | procE | procE | procPP | procPP$$

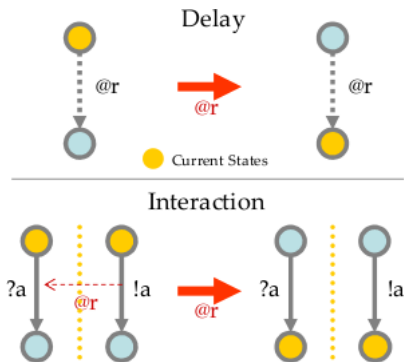
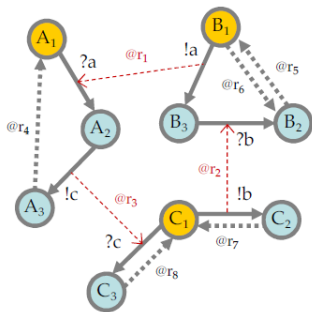
# Molekuly jako komunikující automaty

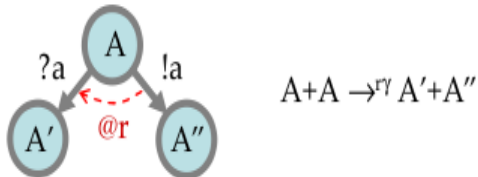
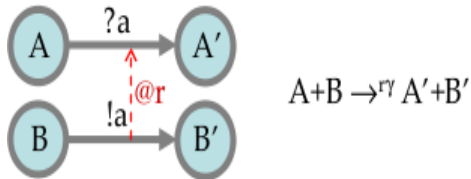
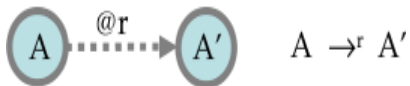


individuální přechody a synchronizace lze modelovat stochasticky  
 (provedení přechodu v čase  $t \sim \text{Exp}(r_i)$ )



# Molekuly jako komunikující automaty



*Elementární reakce jako komunikující automaty*

# *Elementární kalkulus pro popis (bio)chemických systémů*

## *Syntax*

$$E ::= 0 \dot{:} X=M, E$$

Reagents

$$M ::= 0 \dot{:} \pi;P \oplus M$$

Molecule

$$P ::= 0 \dot{:} X \mid P$$

Solution

$$\pi ::= \tau_{(r)} \dot{:} ?n_{(r)} \dot{:} !n_{(r)}$$

Interaction prefix

$$\text{CGF} ::= E, P$$

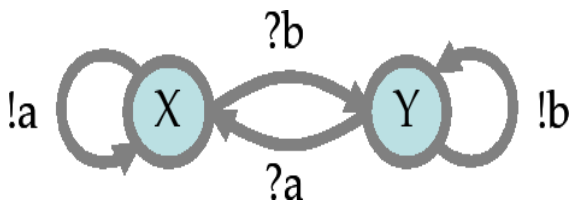
Chemical Ground Form

*Příklad*

$$X = !a_{(r)};X \oplus ?b_{(s)};Y$$

$$Y = !b_{(s)};Y \oplus ?a_{(r)};X$$

$$X \mid X \mid X \mid Y \mid Y$$



# Elementární kalkulus pro popis (bio)chemických systémů

## Sémantika

$$\begin{array}{l}
 \Rightarrow !x.P + M \xrightarrow{!x} P \\
 \Rightarrow ?x.P + M \xrightarrow{?x} P \\
 \Rightarrow \tau_r.P \xrightarrow{r} P \\
 P \xrightarrow{!x} P' \quad Q \xrightarrow{?x} Q' \Rightarrow P|Q \xrightarrow{x} P'|Q' \\
 M \xrightarrow{\alpha} P' \Rightarrow \pi.P + M \xrightarrow{\alpha} P' \\
 P \xrightarrow{\alpha} \Rightarrow P|Q \xrightarrow{\alpha} P'|Q \\
 X = P \quad P \xrightarrow{\alpha} P' \Rightarrow X \xrightarrow{\alpha} P'
 \end{array}$$

SOS pravidla definují fragment stochastického  $\pi$ -kalkulu

# Rozšíření o předávání hodnot

## Syntax

$E ::= 0 \dot{=} X(\mathbf{p})=M, E$	Reagents
$M ::= 0 \dot{=} \pi;P \oplus M$	Molecule
$P ::= 0 \dot{=} X(\mathbf{p}) \mid P$	Solution
$\pi ::= \tau_{(r)} \dot{=} ?n(\mathbf{p}) \dot{=} !n(\mathbf{p})$	Interaction prefix
$CPF ::= E, P$	Chemical Parametric Form

# Rozšíření o předáváníí hodnot

## Sémantika

$$\begin{array}{l}
 P \xrightarrow{!x(n)} P' \quad Q \xrightarrow{?x(m)} Q' \\
 M \xrightarrow{\alpha} P' \\
 P \xrightarrow{\alpha} \\
 X(m) = P \quad P_{\{n/m\}} \xrightarrow{\alpha} P' \Rightarrow X(n) \xrightarrow{\alpha} P'
 \end{array}
 \Rightarrow !x(n).P + M \xrightarrow{!x(n)} P$$

$$\begin{array}{l}
 \Rightarrow ?x(m).P + M \xrightarrow{?x(m)} P_{\{n/m\}} \\
 \Rightarrow \tau_r.P \xrightarrow{r} P \\
 \Rightarrow P|Q \xrightarrow{x} P'|Q' \\
 \Rightarrow \pi.P + M \xrightarrow{\alpha} P' \\
 \Rightarrow P|Q \xrightarrow{\alpha} P'|Q
 \end{array}$$

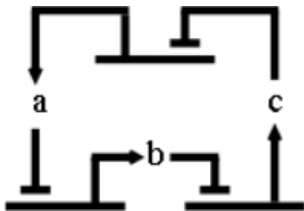
SOS pravidla definují fragment stochastického  $\pi$ -kalkulu

# Příklad modelování genetické regulační sítě



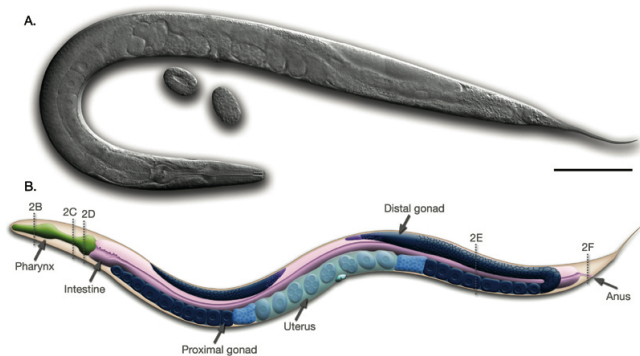
$$\begin{aligned}
 \text{Gene}(a, b) &= \tau_t.(\text{Gene}(a, b)|\text{Protein}(b)) + ?a.\text{Blocked}(a, b) \\
 \text{Blocked}(a, b) &= \tau_u.\text{Gene}(a, b) \\
 \text{Protein}(b) &= !b.\text{Protein}(b) + \tau_d.0
 \end{aligned}$$



*Příklad modelování genetické regulační sítě*
$$\text{Gene}(c, a) | \text{Gene}(a, b) | \text{Gene}(b, c)$$

# High-level modelování

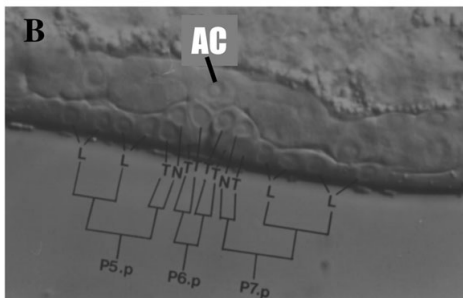
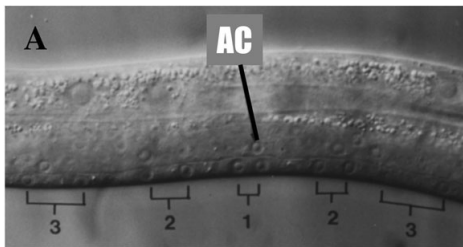
*C. elegans*



IntroFig1

# *High-level modelování*

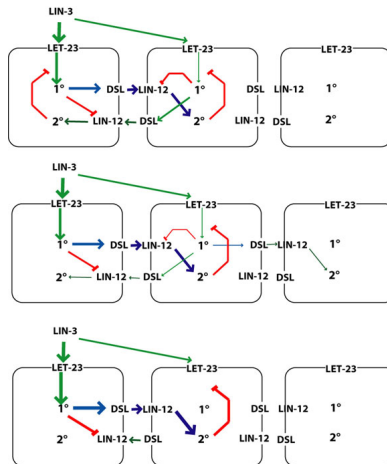
*C. elegans: vývoj vulvy*





# High-level modelování

*C. elegans*: možné varianty signálních drah

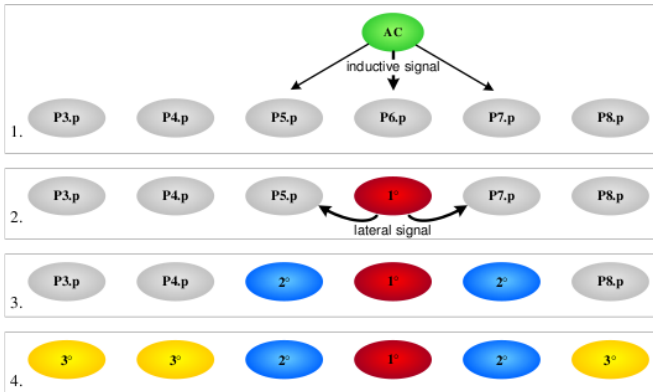


*LIN3* ... signál od řídicí buňky (anchor cell, AC)

*LIN12* ... laterální mezibuněčný signál

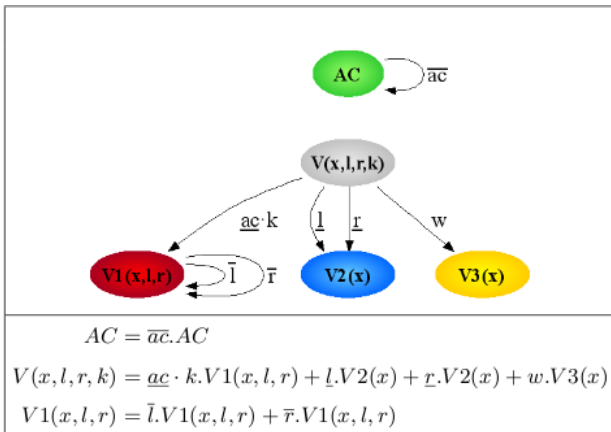
# High-level modelování – interagující buňky

## Vývoj vulvy *C. elegans*



# High-level modelování – interagující buňky

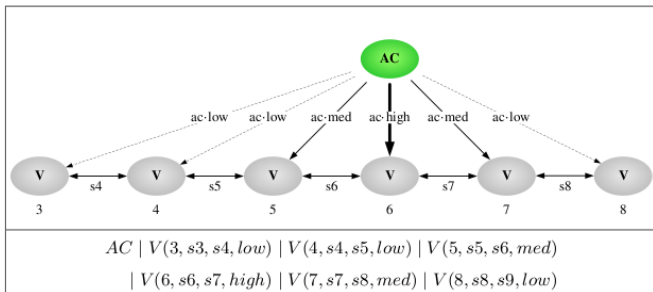
## Vývoj vulvy *C. elegans*



$x$  ... č. buňky,  $l, r$  ... kanály laterálního signálu,  
 $k$  ... vzdálenost od AC (frekvence interakce s AC)

# High-level modelování – interagující buňky

## Vývoj vulvy *C. elegans*

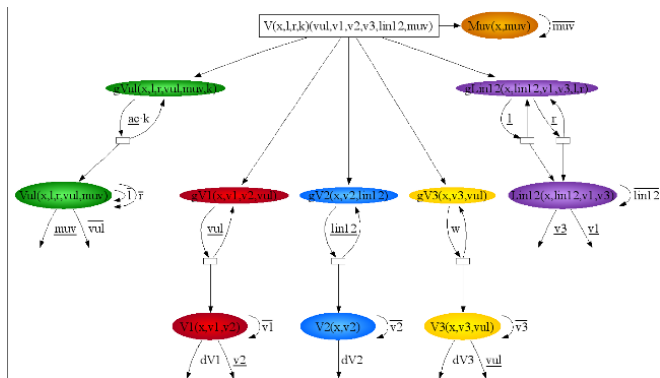


Andrew Phillips, A Visual Process Calculus for Biology, in Symbolic Systems Biology: Theory and Methods, Jones and Bartlett Publishers, 2010.



# High-level modelování – interagující buňky

Vývoj vulvy *C. elegans* – podrobný model



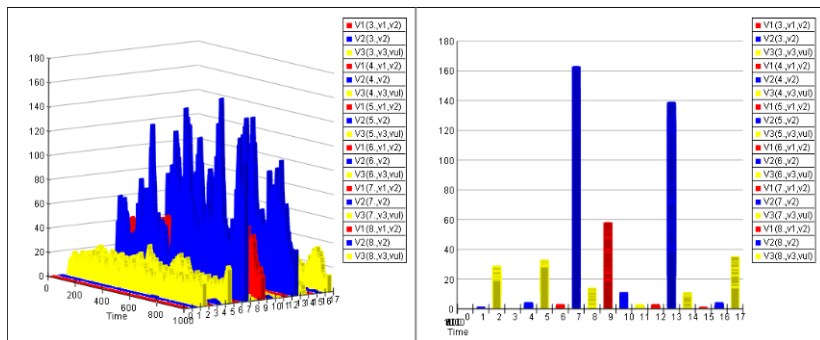
# High-level modelování – interagující buňky

## Vývoj vulvy *C. elegans* – podrobný model

$$\begin{aligned}
 V(x, l, r, k) &= (vul, v1, v3, v3, lin12, muv) \\
 &\quad (gVul(x, l, r, vul, muv, k) | gLin12(x, lin12, v1, v3, l, r) \\
 &\quad | gV1(x, v1, v2, vul) | gV2(x, v2, lin12) | gV3(x, v3, vul) | Muv(x, muv)) \\
 gVul(x, l, r, vul, muv, k) &= \underline{ac} \cdot k.(gVul(x, l, r, vul, muv, k) | Vul(x, l, r, vul, muv)) \\
 gLin12(x, lin12, v1, v3, l, r) &= \underline{l}.(gLin12(x, lin12, v1, v3, l, r) | Lin12(x, lin12, v1, v3) \\
 &\quad + \underline{r}.(gLin12(x, lin12, v1, v3, l, r) | Lin12(x, lin12, v1, v3)) \\
 gV1(x, v1, v2, vul) &= \underline{vul}.(gV1(x, v1, v2, vul) | V1(x, v1, v2)) \\
 gV2(x, v2, lin12) &= \underline{lin12}.(gV2(x, v2, lin12) | V2(x, v2)) \\
 gV3(x, v3, vul) &= \underline{w}.(gV3(x, v3, vul) | V3(x, v3, vul)) \\
 Vul(x, l, r, vul, muv) &= \bar{l}.Vul(x, l, r, vul, muv) + \bar{r}.Vul(x, l, r, vul, muv) + \bar{vul} + \underline{muv} \\
 Lin12(x, lin12, v1, v3) &= \bar{lin12}.Lin12(x, lin12, v1, v3) + \underline{v3} + \underline{v1} \\
 Muv(x, muv) &= \bar{muv}.Muv(x, muv) \\
 V1(x, v1, v2) &= \bar{v1}.V1(x, v1, v2) + \underline{v2} + dV1 \\
 V2(x, v2) &= \bar{v2}.V2(x, v2) + dV2 \\
 V3(x, v3, vul) &= \bar{v3}.V3(x, v3, vul) + \underline{vul} + dV3
 \end{aligned}$$

# High-level modelování – simulace

## Vývoj vulvy *C. elegans*

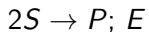


## *Kalkuly pro biologické systémy*

- Stochastic  $\pi$ -calculus (SPiM, BioSPI)
- BioPEPA
- $\kappa$ -calculus
- Brane-Calculus, BetaBinders, BlenX
- ...

*BioPEPA*

- na rozdíl od  $\pi$ -based kalkulů je BioPEPA orientovaný blíže SBML
- rozlišení není až na úroveň molekul, ale na úroveň substrátů a reakcí  
⇒ proces není molekula, ale substrát
- výhodou je kompatibilita sémantik s SBML (srovnej s Petriho sítěmi)
- možnost plného modelování netriviální kinetiky (např. regulace, enzymová kinetika)

*BioPEPA – příklad*

$$S = (\alpha, 2) \downarrow S$$

$$E = (\alpha, 1) \oplus E$$

$$P = (\alpha, 1) \uparrow P$$

$$(S(I_{S_0}) \bowtie_{\alpha} (E(I_{E_0}))) \bowtie_{\alpha} P(I_{P_0}))$$

Sémantika dynamiky je definována pravidlem  $f_{\alpha} = \frac{v \cdot E \cdot S}{K + S^2}$ .

# *BioPEPA – syntax*

## *Algebra modelových komponent*

$$S ::= (\alpha, k) \text{ op } S \mid S + S \mid C$$

$$\text{op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot$$

$$P ::= P \bowtie_{\mathcal{L}} P \mid S(I)$$

- $\mathcal{L}$  je množina reakcí
- $I$  je iniciální podmínka

# *BioPEPA – syntax*

## *Definice elementárních komponent – substráty*

Každá komponenta  $C$  je charakterizována 5-ticí  $\langle H, N, M_0, M, V \rangle$ , kde:

- $H \in \mathbb{R}^+$  je rozlišení kvantity (velikost jedné “úrovně” koncentrace),
- $N \in \mathbb{N}$  je maximální hodnota kvantity,
- $M_0 \in \mathbb{R}^+ \cup \{-\}$  je iniciální koncentrace,
- $M \in \mathbb{R}^+ \cup \{-\}$  je maximální koncentrace,
- $V$  označuje kompartment do něhož je komponenta přiřazena.

Pozn. (1):  $-$  značí “nedefinovanou hodnotu”

Pozn. (2):  $M, M_0$  jsou uvedeny pro kalibraci se spojitým modelem.



# BioPEPA – syntax

## Definice elementárních komponent – substráty

- diskretizace koncentrace pomocí úrovní stejné délky (v počtu molekul)
- diskrétní sémantika tedy umožňuje kalibraci se spojitou (aproximace)
- počet úrovní pro komponentu  $C_i$  je  $N_i + 1$ , kde  $N_i$  je max. úroveň
- platí:  $N_j = \lceil \frac{M_j}{H} \rceil$   
kde  $M_j \leq M$ ,  $0 \leq j < N$ , reprezentuje hraniční koncentraci mezi levely  $N_j$ ,  $N_{j+1}$
- pro aktuální úroveň  $0 \leq l_j \leq N_j$  je koncentrace definována vztahem  $x_j = l_j \cdot H$

## *BioPEPA – definice biologického procesu*

BioPEPA systém je šestice  $\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P \rangle$  kde:

- $\mathcal{V}$  ... množina kompartmentů,
- $\mathcal{N}$  ... množina kvantit popisujících substráty,
- $\mathcal{K}$  ... množina definic parametrů,
- $\mathcal{F}_R$  ... množina definic dynamiky,
- $Comp$  ... definice komponent,
- $P$  ... komponenta popisující systém.

*BioPEPA – sémantika*

- každá akce  $\alpha$  má přiřazenu kinetickou funkci  $f_\alpha$  nad kvantitami substrátů
- kvantitativní substrátů jsou určeny z kontextu akce  $\alpha$
- kinetické funkce mohou obsahovat kinetické parametry, které musí být asociovány v  $\mathcal{K}$
- příklady typů kinetických funkcí:
  - $fMA(k) = k \cdot \prod_{i=1}^{n_j} (C_i)^{\kappa_i}$  kde  $n_j$  je počet reaktantů reakce  $\alpha_j$  a  $\kappa_i$  je příslušný stechiometrický koeficient reaktantu  $C_i$
  - $fH(v, K, n) = v \cdot C^n / (K + C^n)$  je Hillova kinetika

## *BioPEPA – operační sémantika*

Předpokládáme  $\mathcal{C}$  spočetná množina zahrnující všechny možné modelové komponenty.

BioPEPA systém má asociovanou stochastickou sémantiku, která je definována prostřednictvím *odvozující relace*  $\rightarrow_c \subseteq \mathcal{C} \times \Omega \times \mathcal{C}$ , kde  $\theta \in \Omega$  vyjadřuje kvantitativní informaci potřebnou k vyhodnocení kinetické funkce,

$$\theta := (\alpha, w)$$

kde  $w := [S : op(l, \kappa)] \mid w@w$ ,  $S \in \mathcal{C}$ ,  $l$  hodnota kvantity komponenty  $S$ ,  $\kappa$  stechiometrický koeficient komponenty  $S$  v  $\alpha$ .

Relace  $\rightarrow_c$  je definována jako minimální relace splňující pravidla na násl. slidu.

# BioPEPA – odvozující relace

$$\text{prefixReac} \quad ((\alpha, \kappa)\downarrow S)(l) \xrightarrow{(\alpha, [S: \downarrow(l, \kappa)])}_c S(l - \kappa) \quad \kappa \leq l \leq N$$

$$\text{prefixProd} \quad ((\alpha, \kappa)\uparrow S)(l) \xrightarrow{(\alpha, [S: \uparrow(l, \kappa)])}_c S(l + \kappa) \quad 0 \leq l \leq (N - \kappa)$$

$$\text{prefixMod} \quad ((\alpha, \kappa)op S)(l) \xrightarrow{(\alpha, [S: op(l, \kappa)])}_c S(l) \quad \text{with } op = \ominus, \oplus, \Theta \text{ and } 0 \leq l \leq N$$

$$\text{choice1} \quad \frac{S_1(l) \xrightarrow{(\alpha, w)}_c S'_1(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha, w)}_c S'_1(l')} \quad \text{choice2} \quad \frac{S_2(l) \xrightarrow{(\alpha, w)}_c S'_2(l')}{(S_1 + S_2)(l) \xrightarrow{(\alpha, w)}_c S'_2(l')}$$

$$\text{constant} \quad \frac{S(l) \xrightarrow{(\alpha, S': [op(l, \kappa)])}_c S'(l')}{C(l) \xrightarrow{(\alpha, C: [op(l, \kappa)])}_c S'(l')} \quad \text{with } C \stackrel{def}{=} S$$

$$\text{coop1} \quad \frac{P_1 \xrightarrow{(\alpha, w)}_c P'_1}{P_1 \boxtimes_{\mathcal{L}} P_2 \xrightarrow{(\alpha, w)}_c P'_1 \boxtimes_{\mathcal{L}} P_2} \quad \text{with } \alpha \notin \mathcal{L}$$

$$\text{coop2} \quad \frac{P_2 \xrightarrow{(\alpha, w)}_c P'_2}{P_1 \boxtimes_{\mathcal{L}} P_2 \xrightarrow{(\alpha, w)}_c P_1 \boxtimes_{\mathcal{L}} P'_2} \quad \text{with } \alpha \notin \mathcal{L}$$

$$\text{coop3} \quad \frac{P_1 \xrightarrow{(\alpha, w_1)}_c P'_1 \quad P_2 \xrightarrow{(\alpha, w_2)}_c P'_2}{P_1 \boxtimes_{\mathcal{L}} P_2 \xrightarrow{(\alpha, w_1 @ w_2)}_c P'_1 \boxtimes_{\mathcal{L}} P'_2} \quad \text{with } \alpha \in \mathcal{L}$$

*BioPEPA – stochastická sémantika*

Stochastická sémantika BioPEPA systému je určena relací:

$$\rightarrow_s \subseteq \mathcal{P} \times \Gamma \times \mathcal{P}$$

kde  $\gamma \in \Gamma$ ,  $\gamma := (\alpha, r_\alpha)$ ,  $r_\alpha \in \mathbb{R}^+$  určuje parametr exponenciální distribuce události v čase (rate).

Relace  $\rightarrow_s$  je definována pomocí pravidla:

$$\text{Final} \frac{P \xrightarrow{(\alpha_j, w)}_c P'}{\langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, \text{Comp}, P \rangle \xrightarrow{(\alpha_j, r_\alpha[w, \mathcal{N}, \mathcal{K}])} \rightarrow_s \langle \mathcal{V}, \mathcal{N}, \mathcal{K}, \mathcal{F}, \text{Comp}, P' \rangle}$$

kde

$$r_\alpha[w, \mathcal{N}, \mathcal{K}] = \frac{f_\alpha[w, \mathcal{N}, \mathcal{K}]}{H}$$

- $H$  je velikost kroku
- $f_\alpha[w, \mathcal{N}, \mathcal{K}]$  značí kin. funkci  $f_\alpha$  vypočítanou v přísl. kontextu

## *BioPEPA – stochastická sémantika*

Evaluace  $f[w, \mathcal{N}, \mathcal{K}]$ :

- pro každou komponentu  $C_i$  je odvozena koncentrace  $l_i \cdot H$
- volný výskyt  $C_i$  je nahrazen  $(l_i \times H)^{\kappa_{ij}}$ , kde  $\kappa_{ij}$  je stechiometrický koeficient substrátu  $C_i$  v reakci  $R_j$
- pozn.: pro  $H = 1$  a vhodné omezení kinetických funkcí dostáváme přímou stochastickou sémantiku (relace  $\rightarrow_s$  definuje CTMC, který bychom získali interpretací stochastické Petriho sítě odpovídající BioPEPA systému)

*BioPEPA – příklad*

Uvažujme reakci  $\alpha : 2X + Y \xrightarrow{k} 3Z$ . Kinetika je definována funkcí  $v = k \cdot X^2 \cdot Y$ . Definujeme komponenty BioPEPA:

$$X \stackrel{\text{def}}{=} (\alpha, 2)\downarrow X \quad Y \stackrel{\text{def}}{=} (\alpha, 1)\downarrow Y \quad Z \stackrel{\text{def}}{=} (\alpha, 3)\uparrow Z$$

Reakční systém je definován:

$$X(l_{X_0}) \bowtie_{\{\alpha\}} Y(l_{Y_0}) \bowtie_{\{\alpha\}} Z(l_{Z_0})$$

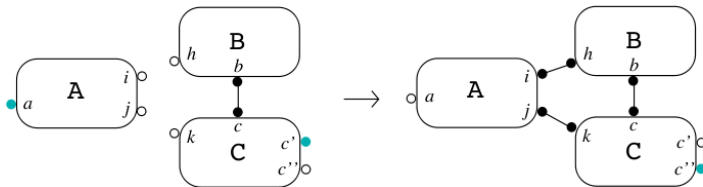
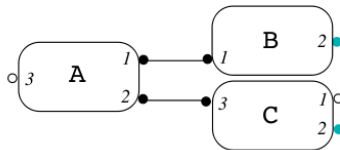
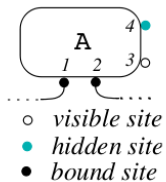
kde  $l_{X_0}, l_{Y_0}, l_{Z_0}$  značí iniciální úrovně koncentrace, kin. funkce je definována  $f_\alpha = fMA(k)$ .

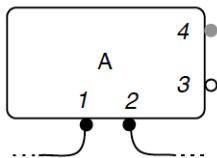
Pro reakci bude odvozením stochastické sémantiky vypočítán rate

$$r_\alpha = \frac{k \cdot (l_X \cdot H)^2 (l_Y \cdot H)}{H}$$

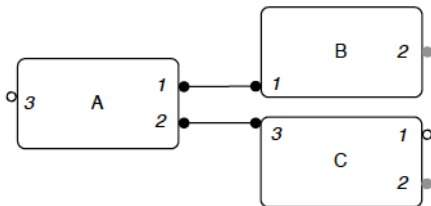
Reakce bude provedena pouze v situaci:  $N_X$  je nejméně 3,  $N_Z$  je nejméně 4.



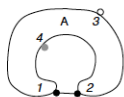
$\kappa$ -calculus

$\kappa$ -calculus – vazebná místa

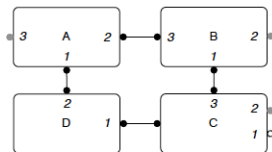
- 1,2 – navázaná vazebná místa
- 3 – skrytá vazebná místa
- 4 – viditelná vazebná místa

*$\kappa$ -calculus – interakce proteinů*

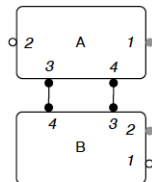
- proteiny mohou vytvářet **komplexy** spojováním na vazebných místech
- kolekce proteinů a jejich komplexů tvoří **roztok (solution)**

*$\kappa$ -calculus – příklady spojení*

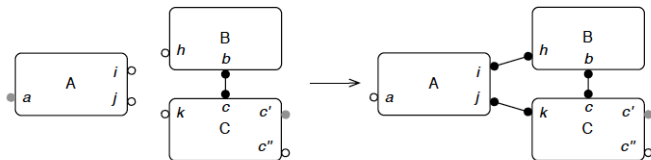
*a self-complexation*

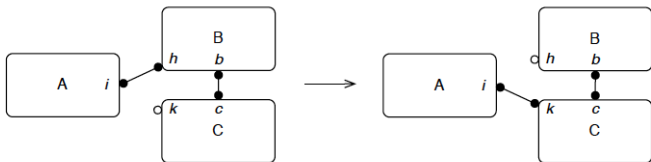


*a ring-complex*



*a double-contact*

$\kappa$ -calculus – komplexace

*$\kappa$ -calculus – požadavek monotonicity*

NEPOVOLENÁ OPERACE, nesplňuje monotonicitu updatu vazeb

$\kappa$ -calculus – syntax

- spočetná množina **jmen proteinů**  $\mathcal{P}$ , zn.  $A, B, \dots \in \mathcal{P}$
- spočetná množina **jmen hran**  $\mathcal{E}$ , zn.  $x, y, \dots \in \mathcal{E}$
- **signatura**,  $f : \mathcal{P} \rightarrow \mathbb{N}$ 
  - $f(A)$  zn. počet vazeb proteinu  $A$
  - $(A, i)$  značí  **$i$ -té vazebné místo** na  $A$

$\kappa$ -calculus – rozhraní

- lib. paricální zobrazení  $\mathbb{N} \rightarrow \mathcal{E} \cup \{h, v\}$  nazýváme **rozhraní**, zn.  $\rho, \sigma, \dots$
- vazebné místo  $(A, i)$  nazýváme:
  - **viditelné**, pokud  $\rho(i) = v$
  - **skryté**, pokud  $\rho(i) = h$
  - **vázané**, pokud  $\rho(i) \in \mathcal{E}$
- rozhraní reprezentuje **stav** roztoku



*$\kappa$ -calculus – příklad rozhraní*

Předpokládejme  $A$  t.ž.  $f(A) = 3$ , pak  $\rho(1) = v$ ,  $\rho(2) = h$ ,  $\rho(3) = x$  je rozhraní proteinu  $A$ , v němž místo 1 je viditelné, místo 2 skryté, a místo 3 vázané na jméno hrany  $x \in \mathcal{E}$ . Symbolicky píšeme:

$$\rho = 1 + \bar{2} + 3^x$$

*$\kappa$ -calculus – syntax roztoku* $\mathcal{S} :=$ 

0      prázdný roztok

 $A(\rho)$     protein $\mathcal{S}, \mathcal{S}$     skupina $(\nu x)(\mathcal{S})$  vazba

Pro množinu jmen  $x_1, x_2, \dots, x_n$  používáme notaci  $\tilde{x}$  a vazbu rozšíříme na  $(\nu \tilde{x})(\mathcal{S})$ .

$\kappa$ -calculus – syntax roztoku

- induktivně definujeme množinu **volných jmen** v roztoku  $\mathcal{S}$ , zn.  $\text{fn}(\mathcal{S})$ :

$$\begin{aligned} \text{fn}(0) &= \emptyset \\ \text{fn}(A(\rho)) &= \text{fn}(\rho) \\ \text{fn}(\mathcal{S}, \mathcal{S}') &= \text{fn}(\mathcal{S}) \cup \text{fn}(\mathcal{S}') \\ \text{fn}((\nu x)(\mathcal{S})) &= \text{fn}(\mathcal{S}) \setminus \{x\} \end{aligned}$$

- výskyt  $x$  v  $\mathcal{S}$  je *vázaný* pokud je zanořen pod operátorem  $\nu x$
- roztok  $\mathcal{S}$  je *uzavřený* pokud všechny výskyty jmen v  $\mathcal{S}$  jsou vázané (zejména  $\text{fn}(\mathcal{S}) = \emptyset$ )

*$\kappa$ -calculus – strukturní kongruence*

Definujeme relaci *strukturní kongruence*, zn.  $\equiv$ , jako nejmenší ekvivalenci uzavřenou syntakticky asociativní a komutativní operací seskupení “,” ,  $\alpha$ -ekvivalencí (injektivní přejmenování vázaných jmen) a 0 chápanou jako neutrální prvek, dále splňující pravidla:

1.  $(\nu x)(\nu y)(\mathcal{S}) \equiv (\nu y)(\nu x)(\mathcal{S})$ ,
2.  $(\nu x)(\mathcal{S}) \equiv \mathcal{S}$ , pro  $x \notin \text{fn}(\mathcal{S})$ ,
3.  $(\nu x)(\mathcal{S}), \mathcal{S}' \equiv (\nu x)(\mathcal{S}, \mathcal{S}')$ , pro  $x \notin \text{fn}(\mathcal{S}')$ .

*Přejmenování*

**Přejmenování**  $r$  je konečná parciální injekce na množině  $\mathcal{E} \cup \{h, v\}$ , která je identitou na  $\{h, v\}$  a zobrazuje  $\mathcal{E} \rightarrow \mathcal{E}$ .

$\kappa$ -calculus – propojenost

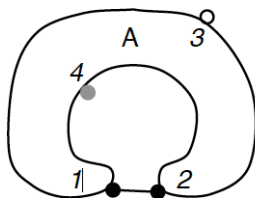
Definujeme **propojený** roztok strukturní indukci:

- $A(\rho)$  je propojený,
- je-li  $\mathcal{S}$  propojený, pak  $i(\nu x)(\mathcal{S})$  je propojený,
- jsou-li  $\mathcal{S}, \mathcal{S}'$  propojené a  $\text{fn}(\mathcal{S} \cap \mathcal{S}') \neq \emptyset$ , pak  $\mathcal{S}, \mathcal{S}'$  je propojený,
- je-li  $\mathcal{S}$  propojený a  $\mathcal{S} \equiv \mathcal{T}$ , pak  $\mathcal{T}$  je propojený.

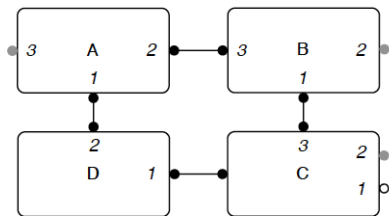
*$\kappa$ -calculus – vizuální omezenost*

Roztok je **vizuálně omezený**, pokud:

- výskyt lib. volného jména v  $\mathcal{S}$  je max. dvojnásobný
- vazba  $(\nu x)(\mathcal{S})$  váže buď 0 nebo 2 výskyty  $x$

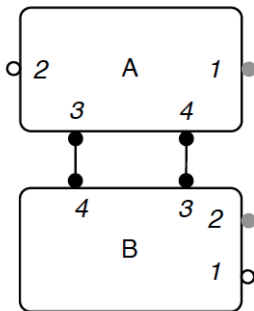
*$\kappa$ -calculus – příklad*

$$(\nu x)(A(1^x + 2^x + 3 + \bar{4}))$$

*$\kappa$ -calculus – příklad*

$$(\nu wxyz)(A(1^x+2^x+3), B(1^z+\bar{2}+3^y), C(1+\bar{2}+3^z+4^w), D(1^w+2^x))$$



*$\kappa$ -calculus – příklad*

$$(\nu xy)(A(\bar{1} + 2 + 3^x + 4^y), B(1 + \bar{2} + 3^y + 4^x))$$

*$\kappa$ -calculus – relace růstu*

- cílem je omezit prostor možných interakcí (reakcí)
- nástrojem bude monotonost updatů vazeb
- formálně zachycena **relací růstu**  $\leq$  definovanou na parciálních rozhraních
- idea:  $\rho \leq \rho'$  odpovídá situaci  $\rho'$  obsahuje více vazeb než  $\rho$
- relace je parametrizována danou množinou jmen  $\tilde{x}$  reprezentující právě nové (přidané) hrany

*$\kappa$ -calculus – relace růstu*

$$\text{(CREATE): } \frac{x \in \tilde{x}}{\tilde{x} \vdash \iota \leq \iota^x}$$

$$\text{(HV-SWITCH): } \frac{}{\tilde{x} \vdash \bar{\iota} \leq \iota}$$

$$\text{(VH-SWITCH): } \frac{}{\tilde{x} \vdash \iota \leq \bar{\iota}}$$

$$\text{(REFLEX): } \frac{\tilde{x} \cap \text{fn}(\rho) = \emptyset}{\tilde{x} \vdash \rho \leq \rho}$$

$$\text{(SUM): } \frac{\tilde{x} \vdash \rho \leq \rho' \quad \tilde{x} \vdash \sigma \leq \sigma'}{\tilde{x} \vdash \rho + \sigma \leq \rho' + \sigma'}$$

$\kappa$ -calculus – relace růstu

- relace  $\leq$  definována pouze pro rozhraní
- rozšíříme na skupiny proteinů

*Definice (Preprotein)*

Pro libovolný protein  $A$  a *parciální rozhraní*  $\rho$  nazveme  $A(\rho)$  **preproteinem**.

*Definice (Preroztok)*

Libovolnou skupinu preproteinů nazveme **preroztokem**.

*$\kappa$ -calculus – relace růstu (rozšíření na preroztoky)*

$$\text{(NIL): } \frac{}{\tilde{x} \vdash 0 \leq 0} \quad (0 \text{ is the empty solution})$$

$$\text{(GROUP): } \frac{\tilde{x} \vdash \mathcal{S} \leq \mathcal{S}' \quad \tilde{x} \vdash \rho \leq \rho' \quad \text{dom}(\rho') \subseteq \mathfrak{s}(A)}{\tilde{x} \vdash \mathcal{S}, A(\rho) \leq \mathcal{S}', A(\rho')}$$

$$\text{(SYNTH): } \frac{\tilde{x} \vdash \mathcal{S} \leq \mathcal{S}' \quad \text{fn}(\rho) \subseteq \tilde{x} \quad \text{dom}(\rho) = \mathfrak{s}(A)}{\tilde{x} \vdash \mathcal{S} \leq \mathcal{S}', A(\rho)}$$

- interpretace  $\tilde{x} \vdash \mathcal{S} \leq \mathcal{S}'$ : nové vazby přidány ( $\mathcal{S}'$  zahrnuje více spojení)
- (SYNTH) povoluje vytváření nových proteinů s plnými rozhraními

$\kappa$ -calculus – reakce

Nechť  $\mathcal{S}, \mathcal{S}'$  jsou preroztoky.

*monotonní reakce*

$r_1 : \mathcal{S} \rightarrow (\nu\tilde{x})\mathcal{S}'$  je monotonní reakce pokud:

- $\tilde{x} \vdash \mathcal{S} \leq \mathcal{S}'$
- $\mathcal{S}$  a  $(\nu\tilde{x})\mathcal{S}'$  jsou vizuálně omezené
- $\mathcal{S}'$  je propojený

*antimonotonní reakce*

$r_2 : (\nu\tilde{x})\mathcal{S} \rightarrow \mathcal{S}'$  je antimonotonní reakce pokud:

- reakce  $\mathcal{S}' \rightarrow (\nu\tilde{x})\mathcal{S}$  je monotonní.