VCell installation and set-up

Go to http://www.vcell.org/ , select "RUN VCELL" from the top menu, choose the proper OS and click the link to VCell **6.1**.

Internet-dependence: After installation, you can run the local VCell application without an Internet connection and without logging in to the model database; in this case, however, you will only be able to work with local files being saved by using File > Export. Whenever you are connected to the Internet, VCell will automatically check for and download any updated version.

Navigation through VCell Modeling Framework

The Virtual Cell opens with four panels.

- The **upper left panel** of VCell provides navigation links to different steps of the VCell modeling process that become active in the **top right panel**.
- Select any element in the top right panel and you can view/edit its **Object Properties** in the **bottom right panel**.
- The **bottom left panel** shows models stored in VCell databases and other resources.
- This guide is using a tutorial model **Rule-based_Ran_transport** located in VCell software Tutorial folder in bottom left panel under **VCellDB -> BioModels**. These models are accessible when users log in to VCell from the Internet.

# Quick tour of compartmental rule-based modeling

Fig. 1. Molecule
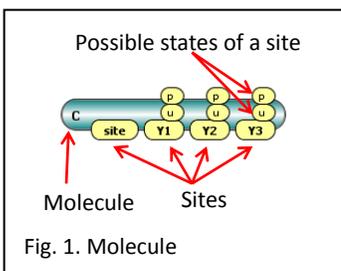
Rule-based modeling involves the representation of **species** as structured objects consisting of **molecules** and molecular interactions as **reaction rules** for transforming the attributes of these objects. It allows one to systematically incorporate site-specific details about molecular interactions into a model.


Fig 2. Anchoring molecules

Molecules (Fig. 1) are the principal component of rule-based model specification. Molecules are comprised of sites that can bind to each other, both within a molecule and between molecules. Sites typically represent physical parts of proteins, such as extracellular and trans-membrane domains or phosphotyrosines of a receptor. Sites may also be associated with a list of states, intended to represent states or properties of the site, e.g. phosphorylation status.
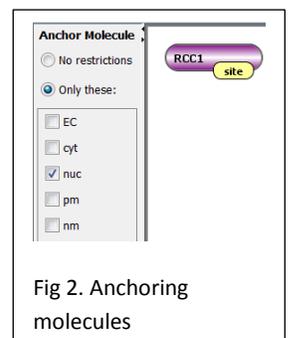
Molecule can be anchored to specific locations (Fig. 2), so all species that contain this molecule will be located in these locations only.

Fig 3. Species comprised of multiple molecules
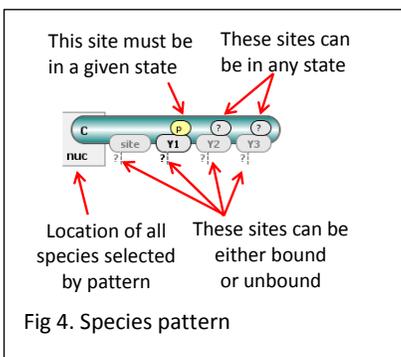
Species (Fig. 3) are composed of molecules with bonds connecting binding sites. If a Molecule has multiple states, each state must be specified - otherwise the species will present a pool of different chemical entities and will not be valid.

Fig 4. Species pattern
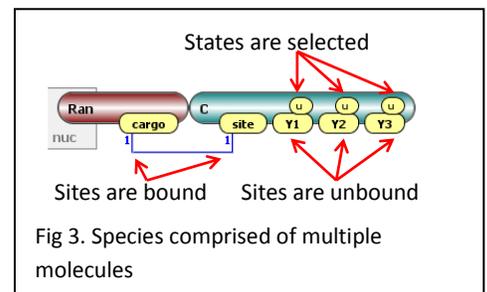
Species Patterns (Fig. 4) specify a set of possible species to be selected as

participants in reaction rules and in observables. Patterns are comprised of molecules. The states of sites may be left unspecified; thus a pattern may select multiple species. Moreover, binding sites may have implicit binding status (has external bond or may be bound) where its binding partner is not explicitly defined. Such patterns may be inclusive of species that contain molecules not explicitly specified in a pattern but being possibly bound to molecules within it.
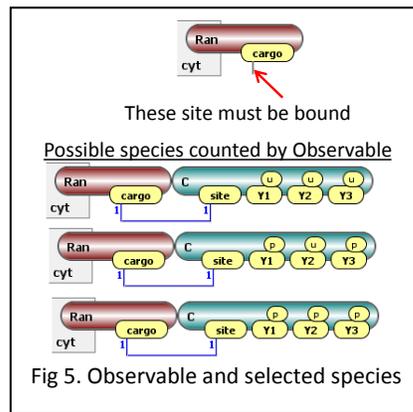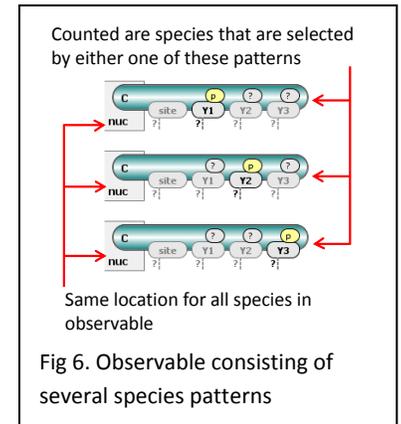

Fig 5. Observable and selected species

Fig 6. Observable consisting of several species patterns

Observables (Fig. 5) are used to specify model outputs, which are functions of the population levels of multiple chemical species that share a set of properties. For example, to simulate the tyrosine phosphorylation level of a particular protein, one needs to specify an Observable that is a function for the total amount of all chemical species containing the phosphorylated form of this protein. Observables are computed over a set of chemical species (Fig. 6) that share similar properties.

Fig 7. Reaction rule for binding molecules.

Reaction rules define transformation of multiple species at once (Fig 7, 8). Species to be transformed are selected by reactant pattern(s). Product pattern(s) define the end result of transformation. Product pattern may differ from reactant by adding, deleting or re-assigning of molecules, adding or deleting bonds and changing site states.
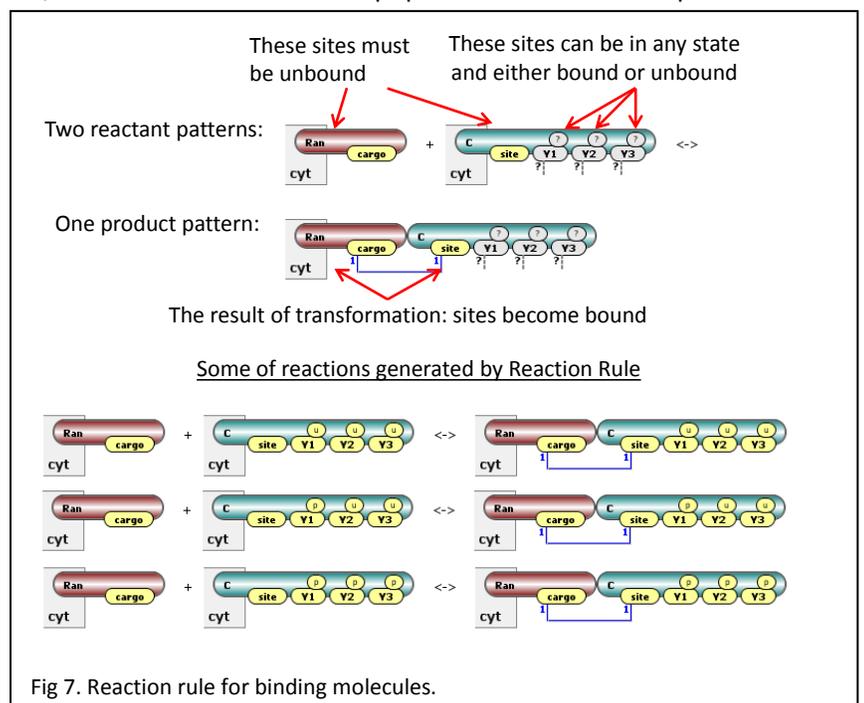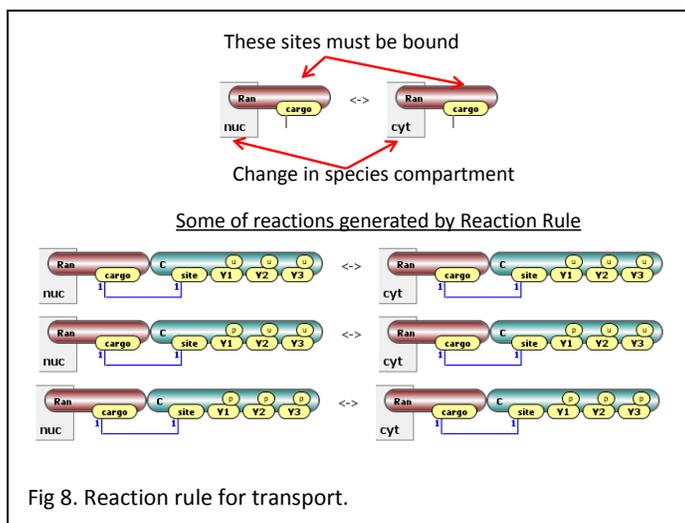

Fig 8. Reaction rule for transport.

Network generation: given an initial set of species with non-zero amounts (**seed species**) and a set of reaction rules, every rule in the set is applied to those species among the seed species that can be selected as reactants, generating proper well-defined reactions and an extended set of species (seed species and species that are products of these reactions). The set of rules is again applied to this extended set of species to generate even more reactions and more species. This process is iteratively repeated until no new species and reactions are generated.

BioNetGen (**Bio**logical **Ne**tworks **Gen**erator,

http://pubmed.org/19399430) and **NFSim** (**N**etwork-**F**ree **Sim**ulator, http://pubmed.org/21186362) are two back-end engines that are used in VCell 6.1 to work with rule-based models.

Load an existing rule-based model

Tutorial rule-based models are located in **Tutorial** in bottom left panel under **VCellDB -> BioModels**. These models are accessible when users log in to VCell from the Internet.

If you have a BNGL code, you can choose **File -> Import**. The code will be imported and a rule-based model will be created with two applications: Network-Free called NFSim, and deterministic network application called BioNetGen. If some features of BNGL file are not supported by VCell, a pop-up window is launched with a suggestion to correct unsupported features (Fig 9). When some features in BNGL file (such as fixed value of concentration) are supported by BioNetGen but not NFSim, only a single deterministic rule-based application will be created. You will be asked about units and simulation volume during import.
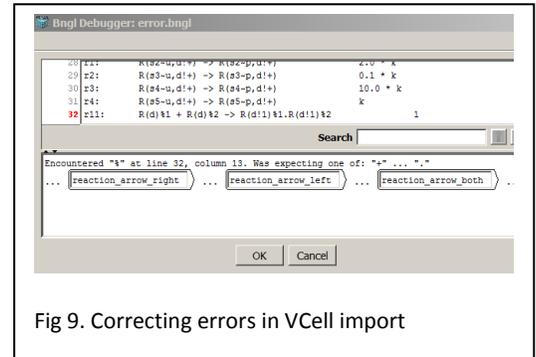

Fig 9. Correcting errors in VCell import

After you're done model creation/editing, you can save your private copy of it: **File > Save As**.

Create a new rule-based model

Please note that **Help->Help** menu provides a very detailed searchable description of all steps in modeling, definitions and specifications. This guide is intended to give just a short overview of modeling process.

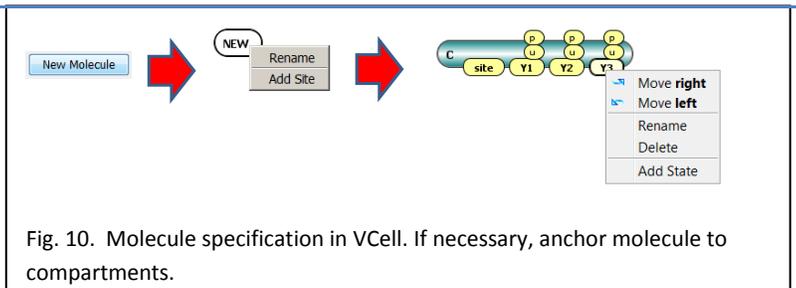1. **Create one or more Molecules** (Fig. 10).

Fig. 10. Molecule specification in VCell. If necessary, anchor molecule to compartments.
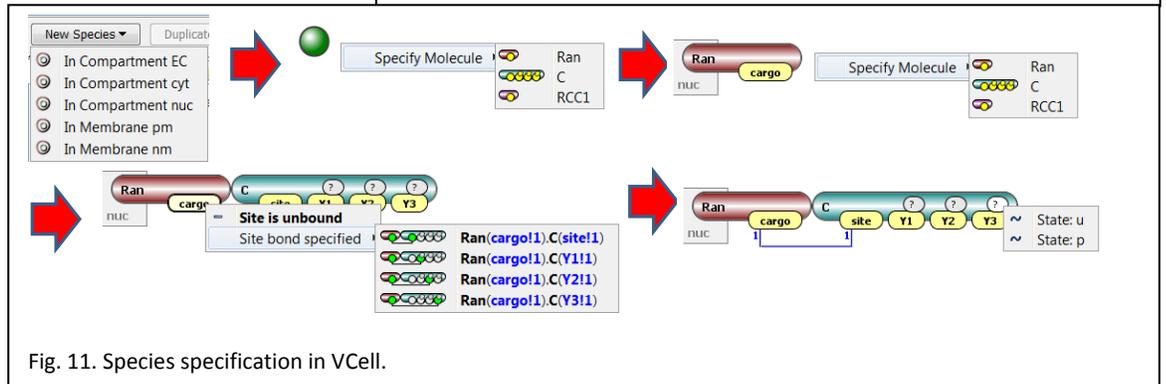
2. **Create one or more Species** (Fig 11).

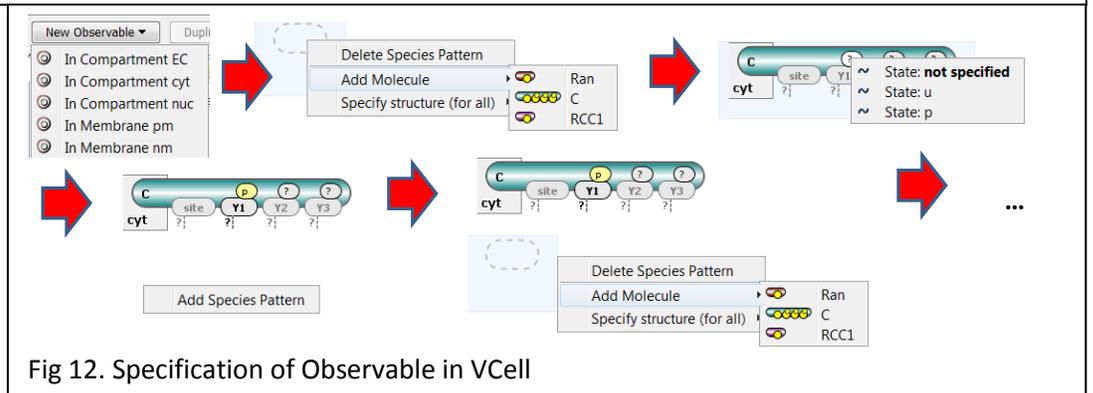Fig. 11. Species specification in VCell.

3. **Create one or more Observables** (Fig 12).

Fig 12. Specification of Observable in VCell

4. Create one or more Reaction Rules (using **Editor** panel in **Reactions**, Fig. 13).

After reaction rule is specified, click on **Kinetics** tab and set whether the rule is reversible, forward and reverse microscopic rate constants for default mass action kinetic law. By default, all reactions generated by a reaction rule have a mass-action kinetic law with forward and reverse rate constants specified here.

Fig 13. Specification of reaction rule for a transport reaction rule
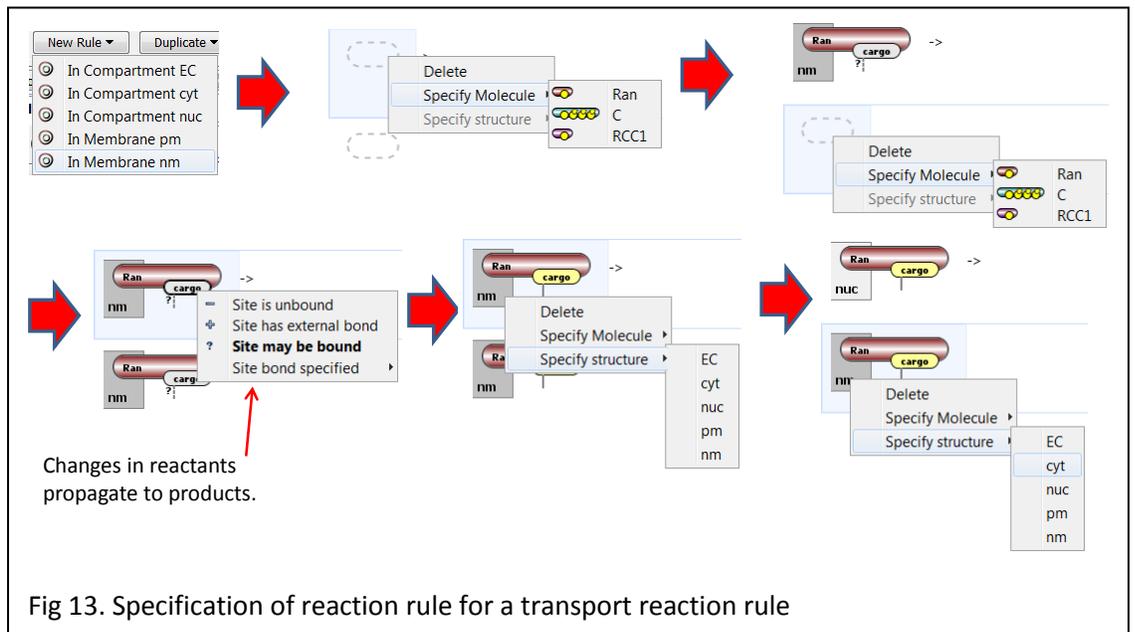
However, these constants can be adjusted by statistical and symmetry factors (see "Expert Options" section below).

Simulate rule-based model

Once rule-based features are specified in **Physiology**, a user has multiple choices of how to simulate. As standard with Virtual Cell, a Rule-Based BioModel may have numerous **Applications**, including **Deterministic**, **Stochastic** and **Network-Free** applications. Each Application, in turn can have multiple **Simulations**, in which different numerical values or simulation conditions are used.

- **Deterministic** and **Stochastic applications** rely on BioNetGen to create the reaction network. Network specification parameters (**Maximal Number of Iterations** to apply rules and **Maximal Number of Molecules per Species** in generated network) can be used to control the size of the network, which otherwise can become infinite or too large. After the network is generated, regular deterministic or stochastic simulators can be run.
- **Network-free Application** simulation avoids network generation and directly simulates the amounts of model observables using NFSim engine. Rather than generating and tracking all possible chemical species, network-free approach follows only the molecular configurations that exist at a given time.

Deterministic Non-spatial Applications: $\frac{d f}{d t}$

- Select Application in the upper left panel, right click **-> New Application -> Deterministic**. By default, new Application is a single compartment.
- Go to **Specifications -> Species** to specify initial values of species and fixed values (clamping).
- Go to **Specifications -> Reactions** to enable/disable some of reaction rules.

- Go to **Specifications -> Network** to check how a reaction network looks like for different constraints on network generation. **Apply** constraints after you're happy with the network.
- Go to **Specification->Simulations** to set simulation parameters and run.
    - o Create and manipulate Simulations using the icons on the left to **Add New**, **Copy**, **Edit**, **Delete**.
    - o Click **Edit** to check solver settings. Use the **Parameter**s tab in the edit simulation window to create Simulations with specific parameter values being overridden.
    - o **Green triangle** on the right sends simulation to the server. Simulation results will be available after simulation is completed. Logging in is required. After green button is pressed, any other work with VCell may be continued, or a user can close VCell completely. Simulation results will be saved in the database.
    - o **Blue triangle** on the right performs simulations on a client's computer. No other VCell operations are possible till simulation is completed.
    - o Simulation results can display Observables, Species (that were specified in Physiology), Generated Species (those generated in the network – their structure can be seen in **Specification -> Network**) and Other – which include reaction fluxes and user-defined functions.
    - o Ctrl and Shift can be used to select multiple outputs.
    - o Simulation Results for "green" run can be viewed in a separate window invoked by clicking on the **Results** icon on the left. Results can be viewed while a Simulation is still running; the data displayed will update automatically at the same time with the Simulation status.

Deterministic Spatial Applications: d/dt
- Most settings are the same as for deterministic application, plus add extra specifications:
- Go to **Specifications -> Geometry** to specify 2D or 3D geometry and map it to compartments.
- In **Specifications -> Species** define diffusion constant for seed species. All newly generated species will have the same diffusion constants.
- Go to **Specification->Simulations** to set simulation parameters and run.
    - o All parameters same as in deterministic non-spatial.
    - o Under Edit -> Mesh, define Mesh size for spatial simulations, under Edit -> Solver specify solver settings.

Stochastic Non-spatial Applications:
- Most settings are the same as for deterministic non-spatial.
- **Specifications -> Species** will give a value of either concentrations, or particle numbers. Simulations will be performed in particles. Conversion is done using compartment volume that is defined in **Specifications->Geometry**. Make sure simulation volume is appropriate to avoid very large number of particles.
- Simulation results can be seen in concentrations and number of particles (name_Count).

Network-Free Applications:
- Most settings are the same as for stochastic non-spatial application.
- This is a stochastic application, so compartment volume must be defined in **Specifications->Geometry**
- **Specifications -> Network-Free** can be used to generate a rule-based model from a regular reaction network. For a purely rule-based model pressing this button will generate a new rule-based model identical to the existing. Check "Expert Options" section for more details.
- NFSim simulator has special settings – check them carefully under **Simulation -> Edit->Solver**.

Stochastic spatial Applications:
- Same as for stochastic non-spatial.

- Go to **Specifications -> Geometry** to specify 2D or 3D geometry and map it to compartments.
- In **Specifications -> Species** define diffusion constant for seed species. All newly generated species will have the same diffusion constants.
- **Specifications -> Species** will give a value of either concentrations, or particle numbers. Simulations will be performed in particles. Conversion is done using compartment volume that is defined in **Specifications->Geometry**. Make sure simulation volume is appropriate to avoid very large number of particles.
- Simulation results can be seen in concentrations and number of particles (name_Count).

Working with polymers

- If a reaction rule has multiple identical molecules, they are enumerated to provide one-to-one mapping from reactants to products. The match is established automatically, but can be reassigned with a right click on a molecular shape.
- Observables can be specified either as multimolecular as in Figs 5 & 6. Observables that count species comprising of multiple identical molecules are defined as polymers (Fig. 14).
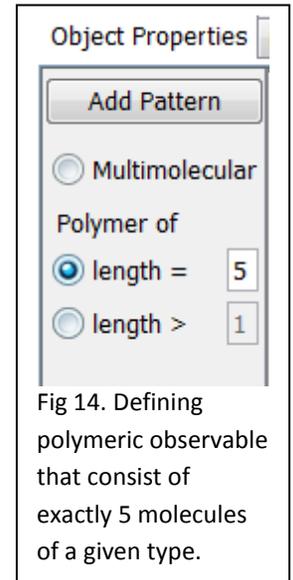
Fig 14. Defining polymeric observable that consist of exactly 5 molecules of a given type.

Expert Options

- If a reaction rule has multiple identical molecules, they are enumerated to provide one-to-one mapping from reactants to products. The match is established automatically, but can be reassigned with a right click on a molecular shape.
- Every VCell BioModel can be enhanced by creating Molecules and assigning molecular composition to species.
- One can mix reactions and reaction rules (see *Mix_Reactions_Rules* model in Tutorial folder). In particular, species that are used as initial seed species for rule-based models can be used as species in reaction networks.
- Only mass-action kinetics is supported in reaction rules. In deterministic applications, one can put expressions depending on Species as forward and reverse rates (see *RB_Enzyme_Kinetics* model in Tutorial folder). In Stochastic and Network-Free Applications, only numerical expressions are allowed.
- Using Network-Free Application, please be aware:
    o While it can be used for every VCell BioModel obeying the limitations above, it is designed to be used to simulate rule-based models.
    o This is a stochastic simulation, so simulation results will vary by run.
    o It operates with particles, so using "copy as" to convert a deterministic application to a Network-Free application may lead to a very large number of particles. This number can be adjusted by changing simulation volume in **Specifications -> Geometry**. The default maximal number of particles is 200,000 (it may be increased in **Simulations -> Edit -> Solver**)
    o The user is unable to see what species are populated during a simulation. To test if reaction rules produce expected species and reactions, the user is advised to generate a deterministic rule-based application and test network properties.
- Using Deterministic and Stochastic Applications, one can create a new VCell BioModel that consists of all generated species and reactions; species have molecular details, while reactions carry rule name under reaction name. Note that reaction rates in reaction network are adjusted for symmetry factors and statistical factors:
    o The rate law associated with a rule is a microscopic rate for reactions of the form A+B -> products(). If a rule generates a reaction of the form A+A -> products(s), a reaction rate for such symmetric reaction is multiplied by a **symmetry factor** of 1/2.

- In some cases, a reaction generated by a rule can occur in multiple ways that are indistinguishable. For example, a rule A.A -> A.A' can be applied in two different ways: either the first A wil be modified, or the second A will be modified. In these cases, the single-site rule rate law is multiplied by a statistical factor 2 to obtain the rate of the reaction.
- For any VCell BioModel that satisfies above constraints (mass-action only) one can create a rule-based model by creating Network-Free Application and going to **Specifications -> Network-Free -> Create new Rule-Based VCell BioModel.** If no rule-based elements were present, new molecule will be created for every species and trivial rules corresponding reactions will be done. If species have molecular structures, they will be used in rule-based model. Note that the same reaction expression treated as a reaction and as a reaction rule will have different rates:
  - VCell will take care of reactions of the form A + A -> product(s), and the rate of generated reaction rules will be multiplied by 2 to account for symmetry factors.
  - VCell does not take care of reactions where reactants have symmetric sites and reaction rule will have a statistical factor. If this is not properly treated by a modeler, the resultant rule-based model may show different simulation results than its network precursor.
- If a reaction rule has multiple identical molecules, they are enumerated to provide one-to-one mapping from reactants to products. The match is established automatically, but can be reassigned with a right click on a molecular shape.

Limitations compared to stand-alone tools (temporary, will be lifted in future releases of VCell). (for BioNetGen/NFSim experts)

- Only mass-action kinetic laws are supported. No functions are supported (only expressions in forward and reverse rates). Total rate is not supported.
- Molecules cannot have identical sites.
- No exclude/include operations are supported.
- All reaction rules with identical molecules are automatically assigned matches between reactant and product patterns.
- Matches between molecules in reactants and products in BNGL file are not imported, but can be reassigned within VCell interface.
- To speed-up simulations, the number of generated species cannot exceed 800 and the number of reactions should be less than 2,000. If network constraints provide larger numbers, they will not be accepted for deterministic and stochastic simulations.