

Proposal for hierarchical extension of SBGN Entity Relationships Diagrams

Michael Blinov

Center for Cell Analysis and Modeling, University of Connecticut Health Center

E-mail: blinov@uchc.edu

Many thanks to Anatoly Sorokin, Mirit Aladjem and the rest of SBGN forum for fruitful discussions

Draft September 6th, 2009

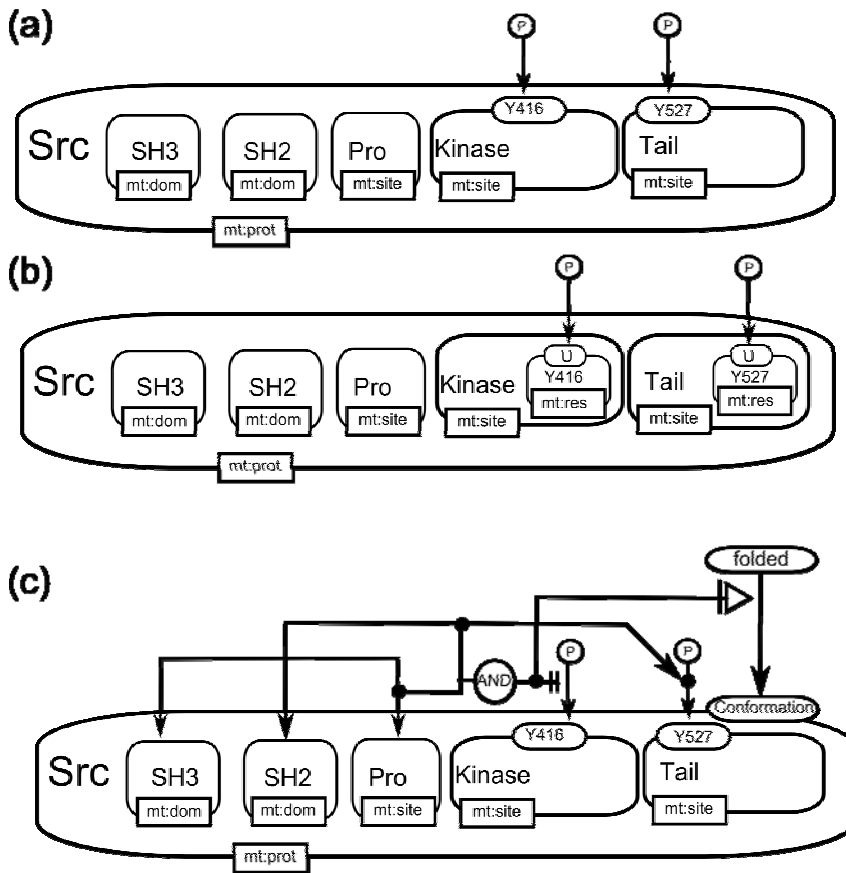
1. Introduction

This proposal extends the current description of Entity Relationships (ER) diagrams by allowing a hierarchical nesting of entities. Thus, it provides a description of domains, sites and individual atoms of biomolecules. The proposal is designed to cover all the requirements and capabilities of non-hierarchical fine structure proposal of MIM group. The proposal is consistent with rule-based description used in BioNetGen software and it is expected to be consistent with SBML level 3 extension describing multi-state multi-component species. It still does not cover all requirements for specification of rule-based models in SBGN, such as specifying topology of complexes or convenient representation of competitive binding. Please see the proposal on “Entity relationships representation of rule-based models” for extra features required for representation of rule-based models as ER diagram.

2. Hierarchical structure of entities

Entity glyphs can be nested one into another, in a manner similar to specification of complexes in Process Diagrams part of SBGN. Examples of nesting entities include proteins that contain catalytic domain that contain tyrosines. Each entity can carry any state variable that can be modified in any allowable way. Figure 1 illustrates two diagrams of the Src protein containing several levels of hierarchy but the same details: (a) protein and domains having state variables; (b) protein, domains, and individual tyrosines having state variables. These descriptions can be mixed and matched.

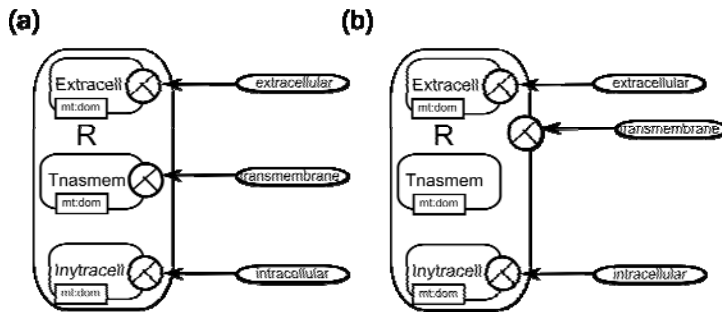
Each entity (without regards to what level hierarchy it belongs to) can interact with any other entity. Figure 1c illustrates an interaction between domains of the same protein, such as binding of SH3 domain of Src to Pro domain. Note that interaction between entities within protein may lead to change in state variable assigned to the whole protein.



3. Inheritance of locations

State variables assigned to hierarchically nested entities are independent and don't know about each other. However, certain conventions are required in two cases: to specify locations and existence.

Each entity can have an optional location. Consistency of locations is responsibility of a user. The convention is that entities of the lower level of hierarchy inherit location of encompassing entity if their own location is not specified. However, location of entity overrides the location of encompassing entity. Figure 2a illustrates an example where every entity representing domains have a specific location, while a protein does not have a location. Individual tyrosines inherit locations of containing domains. Figure 2b illustrates an example where the receptor is assigned to a membrane; however, individual domains are located in different compartments. Location of transmembrane domain is inherited from the location of receptor.

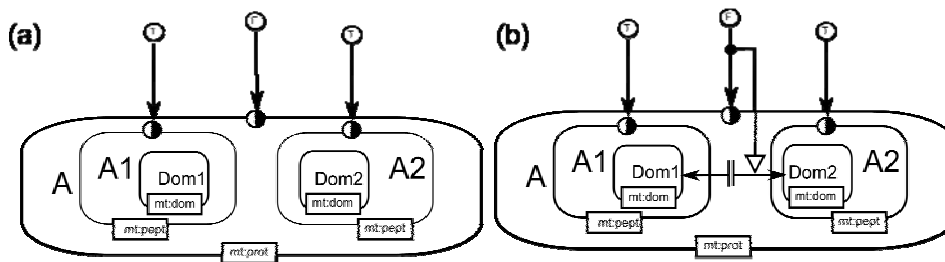


4. Existence and cleavage

Existence glyph follows similar conventions to location. If the encompassing entity is destroyed, all entities inside are destroyed by default (no need to specify glyphs) unless their existence flag is set to positive.

Generally, cleavage can be represented by destroying the top level entity. The user is advised to introduce “intermediate entities” that represent parts of cleaved protein. This way, after deleting the top level entity cleaved parts become “the top level”, as in Figure 3a. The apparent problem is the type of new entities: they can not become proteins.

Normally, entities representing domains of a protein don’t need to be connected via interactions. Their relative position can be specified using unit of information and location. However, to specify cleavage of an entity, domains can be connected via “interaction” arc, as shown in Figure 3b.



5. Conclusions

The approach has a significant advantage that in multiple cases it allows flattening of the diagram. Generally, if no cleavage is included and locations of all entities are specified, all entities can be moved outside of encompassing entities with the resulting flattened ER diagram carrying the same information as the hierarchical one.

Also, by placing entity labels in the predefined order, the proposed cartoon will be almost identical to the one suggested in Fine-structure proposal of MIM group.