# COMBINING DEVS AND SEMANTIC TECHNOLOGIES FOR MODELING THE SARS-COV-2 REPLICATION MACHINERY

Ali Ayadi
Claudia Frydman

Laboratoire d'Informatique et des Systèmes
52 Avenue Escadrille Normandie Niemen
13397 Marseille Cedex 20, FRANCE
{ali.ayadi,claudia.frydman}@lis-lab.fr

Wissame Laddada
Lina F. Soualmia
Cecilia Zanni-Merk

LITIS, INSA Rouen, Normandie Université,
76801 Saint-Etienne-du-Rouvray, FRANCE
wissame.laddada@univ-amu.fr,
{lina.soualmia, cecilia.zanni-merk}@litislab.fr

India L'Hote
Emeline Grellet
Isabelle Imbert

Laboratoire Architecture et Fonction des Macromolécules Biologiques,
Case 932. 163 Avenue de Luminy, 13009 Marseille, FRANCE
{india.lhote,emeline.grellet,isabelle.imbert}@univ-amu.fr

## ABSTRACT

The search for inhibitors of SARS-CoV-2 viral replication depends on an in-depth knowledge of the different stages of the viral cycle. The macro-molecular level focuses on the interactions between the virus and the infected cell, while the micro-molecular level focuses on the different biochemical reactions leading to the production of new viruses. Here, a hybrid approach for modeling the SARS-CoV-2 viral replication in the micro- and macro-molecular levels is presented. The proposed approach combines ontology engineering and DEVS modeling. Biological knowledge at the micro-level of the viral system is capitalized by ontological models, while the dynamic behavior of SARS-CoV-2 molecular mechanisms are modeled by DEVS models. The proposed DEVS approach uses ontological concepts and SWRL rules to compute the main functions and behaviour of the molecular components involved in the SARS-CoV-2 replication cycle. We illustrate the proposed approach through the simulation of the SARS-CoV-2 proteins production by cellular ribosomes.

**Keywords:** SARS-CoV-2 replication machinery, DEVS formalism, ontology-based modeling and simulation.

## 1 INTRODUCTION

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the virus responsible of the current coronavirus disease-2019 (COVID-19) pandemic. As of 11 April, 2021, there have been 134,957,021 confirmed cases of COVID-19, including 2,918,752 deaths (WHO data). This virus mainly affects the respiratory tract and particularly the lungs. It generally causes flu-like symptoms such as fever, coughing, muscle

aches and fatigue, with an incubation period of 2 to 14 days (Centers for Disease Control Prevention 2020). In more serious cases, the infection can cause severe pneumonia, kidney failure, and eventually death.

Unlike most bacteria, viruses such as SARS-CoV-2 are not autonomous organisms. They are considered as obligate parasites that depend on the host cell they infect, for the replication of their genome and the production of their proteins (structural, non-structural proteins,...). Viruses have developed many strategies to hijack the biological processes of the infected-cell to their benefit. Indeed, they use the host cellular components for their benefit (Davey et al. 2011). It is therefore crucial to understand how this virus interacts with host cell components at different levels. At the micro-level, by understanding the different biochemical reactions leading to the production of viral proteins, modifications that occur to one or more amino acids on proteins, and at the macro-level, by analyzing the interactions among viral and host proteins. Understanding the detailed interactions between viral and host cell components, and how codons and genomic sequences are assembled, can provide important information for the development of new antiviral therapies and vaccine technologies, by blocking these vital processes for the virus.

In this paper, a hybrid approach for modeling and simulating the SARS-CoV-2 viral replication, in the micro- and macro-molecular levels, will be presented. This proposed approach aims to mix and combine knowledge representation, and discrete event simulation (especially the DEVS formalism) in a common modeling framework to face the complexity of understanding the SARS-CoV-2 viral replication at different levels. By leveraging the benefits of both domains, the proposed approach allows to understand, model, and simulate, qualitatively and quantitatively the mechanisms behind the replication machinery of the SARS-CoV-2. On one hand, biological knowledge at the micro-level of the viral system is capitalized and inferred by ontological models. They build an exhaustive semantic vocabulary related to the virus replication, in particular by providing properties and rules addressing the assembly of the genome sequence, nucleotides, codons, and amino acids. On the other hand, DEVS modeling has the interest to be timed, highly modular and hierarchical for the description of the dynamic behavior of SARS-CoV-2 molecular mechanisms and their different state changes.

## 2 BACKGROUND

### 2.1 SARS-COV-2 Life Cycle

The SARS-CoV-2 exhibits a large positive-sense single-stranded RNA genome of approximately 30,000 nucleotides (Shang et al. 2020). The following items describe the three main steps of the SARS-CoV-2 viral cycle:

1. *The viral entry*: To gain access to the host cell cytoplasm, the virus via its spicules (spike S) attach to the host cell receptor called angiotensin-converting enzymes 2 (ACE2). The interaction between S and ACE2 will allow the fusion between the virus and the cell membranes and thus the release of the viral genome into the cytoplasm (Shang et al. 2020).
2. *The viral replication*: Once inside the host cell, the virus hijacks the host cell's machinery, i.e. all production mechanisms comprising translation, replication and transcription, allowing the production of all the elements for new virus particles biogenesis.
3. *The viral particles export*: The genomic RNA will assemble with its structural proteins (M, S, E, and N) to form a new viral particle. Then, the resulting virion is transported to the surface of the infected cell through the secretion pathway and released out of the cell by exocytosis, ready to infect other cells (Shang et al. 2020).

### 2.2 Ontology Engineering

The modeling process, through ontologies, consolidates the philosophical idea to describe all existing entities and their connections. More broadly, an ontology is used for knowledge and semantic representation. It is a set of taxonomies gathering concepts and roles. The concepts describe all the entities related to a specific field and the roles define relationships that link the concepts. This representation is based on a logic formalism (Description Logic and First Order Logic) that allows performing a reasoning process to build a system based on symbolic artificial intelligence.

### 2.3 Related Work

Modeling and simulations play a key role in virology, allowing us to study viruses and their components. To date, the main modeling methods in this field consist of mathematical and statistical models (Hattaf et al. 2018). This is the case for the models of the infection dynamics of Ebola (Madelain et al. 2018), influenza A (Handel et al. 2018), HIV-AIDS (Li et al. 2017) and Zika (Best and Perelson 2018) viruses, to cite a few that have been studied. These models were formulated as a system of ordinary differential equations which require optimization (Fabreti et al. 2019). However, important computation power is sometimes needed to have quantitative solutions to these mathematical models and often, analytical solutions do not exist for them. Moreover, in such mathematical models, the number of parameters is huge. Consequently, obtaining a value for each parameter through laboratory experiments or theoretical calculations is impossible (de Oliveira et al. 2016). Furthermore, as discussed above, the interplay between viral and human proteins happens at different levels (at a macro-level: interactions between cellular components, and at a micro-level: mutations in the viral RNA genome that will have repercussions on amino-acids and consequently, on the protein sequence) and time scales. Indeed, the viral cycle stages timing is key to the virus survival. To take into account the dynamics of multi-component and multi-scale phenomenon, we propose a hybrid approach that relies on the use of semantic technologies associated with simulation to model and simulate the SARS-CoV-2 amplification cycle.

## 3 PROPOSED APPROACH: COMBINED ONTOLOGICAL AND DEVS MODELS

### 3.1 Approach Overview

The proposed hybrid approach relies on the use of semantic technologies associated with discrete event simulation to characterize the SARS-CoV-2 RNA genome replication. The first part of the proposed approach is based on a formal model, particularly an ontology associated with reasoning rules expressed with Semantic Web Rule Language (SWRL) expressions, which integrates the available knowledge about the SARS-CoV-2 replication process. The main purpose of this unified model is to allow a semantically rich representation of the dynamic properties of the SARS-CoV-2 RNA synthesis machinery which is more complex than representing simple biological concepts. The second part of the proposed approach aims to generate discrete event simulation models, especially DEVS models, from the SARS-CoV-2 domain ontology combined with reasoning rules developed in the first part. These DEVS models will be implemented to simulate the replication cycle of the SARS-CoV-2 polymerase. The source code of both ontological and DEVS simulation models are made publicly available on BioPortal (BioPortal 2021) and on Github repository (GitHub 2021), respectively.

### 3.2 Ontological Modeling Phase

Modeling the replication machinery of SARS-CoV-2 leads us to define several taxonomies. First, we describe all objects from the replication as concepts (classes). Hence, two main classes are defined, namely, *Replication_element* and *Genome*. The latter gathers all the concepts related to the genome organization such as *Nucleobase*, *Open_reading_frame*, *Polyprotein*, *Accessory_protein*, *Structural_protein*, etc. The *Accessory_protein* includes sixteen Non-Structural Proteins (NSPs) (*NSP_1*, ..., *NSP_16*). The *Structural_protein* is enriched with the subclasses: *E*, *M*, *N*, and *Spike*. The class *Replication_element* encompasses all elements needed for the replication process such as *Ribosome* that includes *tRNA*, *Amino_acid*, *Protease*, *Replication_Transcription_Complex*, *RNA_genome*, *Messenger_RNA*, ...

In addition to this terminology, our ontology takes into account relationships (roles through Object and Data properties) between the concepts. Each role $R$ that links two concepts $x$ and $y$ is described by the binary predicate $R(x,y)$. For example, *has_first_base (Codon, Nucleobase)*, *has_second_base (Codon, Nucleobase)*, and *has_third_base (Codon, Nucleobase)* are roles defining the fact that a codon is a sequence of three nucleobases. The property *has_next (Nucleobase/Codon, Nucleobase/Codon)* associates each nucleobase to its next nucleobase and each codon to its next codon describing hence, the genome sequence by mean of nucleobases or the polyprotein sequence via codons (amino acids). We associate to this two classes a rank with the data property *has_rank (Nucleobase/Codon, Nucleobase/Codon)*. On the whole, our ontology is developed with eighty-four classes and fifteen properties.

To enhance the semantic of our ontology, a reasoning process is considered through axioms that define the 22 amino acids which enter into the composition of viral proteins translated by ribosomes. Reasoning rules were also defined to described the continuous genomic RNA replication as well discontinouous transcription of subgenomic mRNAs.

### 3.3 Mapping Ontological Model to DEVS Model Phase

This second step consists of translating the ontological model, associated with reasoning rules, to DEVS models. As depicted in Table 1, the concepts (or classes) of the ontology, formalizing objects in the replication machinery, are translated into inputs or outputs of atomic or coupled DEVS models. For example, the concept *mRNA* should become an input port *mRNA_in* in the DEVS model, etc. All the axioms and algorithms in the formal model are translated into atomic or coupled DEVS models. For example, the axiom describing the complex RTC (also called Replication transcription complex) in the replication machinery is defined by the following axiom: $Com\_RTC \equiv Element\_replication$ and *has_genome_element* only ($nsp7$ and $nsp8$ and $nsp10$ and $nsp12$ and $nsp14$),
and will be transformed into an atomic model named "Complex_RTC". Similarly, the SWRL reasoning rules associated with the ontological model, representing transitions in the replication machinery, are either translated or implemented inside the internal or external transition functions. A detailed example of this case is provided in the proof of concept section (Section 4).

After translating our ontological model and its associated reasoning rules, we obtain a DEVS model describing the different steps of the SARS-CoV-2 replication cycle, starting from its attachment to a host cell, traversing its replication process, to the release of new virions. The obtained model consists of a coupled DEVS model describing the host cell, having as input a spike S and for output the new virions. This coupled model is constituted of three coupled models describing the cell membrane, cytoplasm, and secretory pathway. The cell membrane consists of a single atomic model representing the ACE2 receptor. The coupled model of the cytoplasm contains seven atomic models, both ribosomes HR1 and HR2, four complex RTCs (Complex_RTC, RTC1, RTC2, and RTC3), and the viral protease. Lastly, the secretory pathway comprises two atomic models, the ERGIC component and the Golgi apparatus.

Table 1: Mapping ontological model with DEVS model.

| **Ontological model** | **Replication machinery** | **DEVS model** |
|---|---|---|
| *Concepts*<br>E.g.: mRNA, pp1a,... (sequence of nucleobases) | *Object* | *Inputs/Outputs*<br>E.g.: mRNA_in, pp1a_out,... |
| *Axioms* and *Algorithms*<br>E.g.: Host Ribosomes,... | *Transition* | *Atomic and coupled models*<br>E.g.: HR atomic model,... |
| *SWRL rules*<br>E.g.: Translation, Transcription, ... | *Process* | *External transition & output functions*<br>E.g.: $\delta_{ext_{HR}}$, $\lambda(HR)$,... |

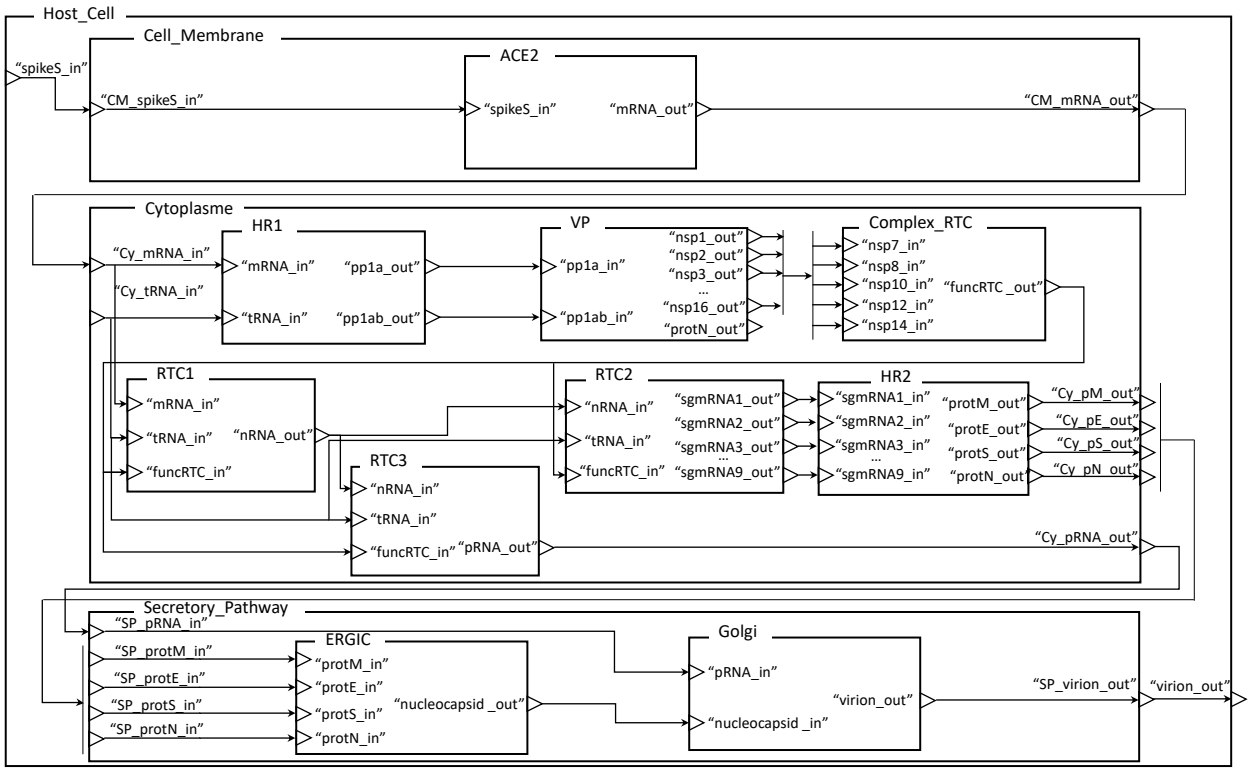


Figure 1: DEVS hierarchical representation of the SARS-CoV-2 replication machinery, including the angiotensin-converting enzymes 2 (ACE2), first host ribosome (HR1), viral protease (VP), four replication transcription complex (Complex_RTC, RTC1, RTC2, and RTC3), second host ribosome (HR2), ER-Golgi intermediate compartment (ERGIC), and Golgi apparatus (GOLGI).

## 4 PROOF OF CONCEPT: SARS-COV-2 PROTEINS TRANSLATION BY CELLULAR RIBOSOMES MODELING AND SIMULATION

### 4.1 Description

The *biosynthesis of proteins*, also called *translation*, is a key phase for all organisms and in the case of the SARS-CoV-2, this stage will produce viral proteins involved in the viral replication as well in the formation of new virions. This translation process is carried out in each cell by very sophisticated molecular machines,

the *ribosomes* (Sarnow et al. 2005). Ribosomes comprise two distinct subunits. The small subunit allows the decoding of genetic information carried by mRNAs, while the large subunit is the site of protein synthesis through the catalysis of a bond between amino acids. The translation process includes all the mechanisms required to convert mRNA into a polypeptide sequence. During this process, mRNAs are translated by triplets of nucleotides (or codons) into one of the twenty-two amino acids constituting the proteins. The translation occurs in three major steps (Figure 2), the initiation, elongation, and termination steps (Schmeing and Ramakrishnan 2009).

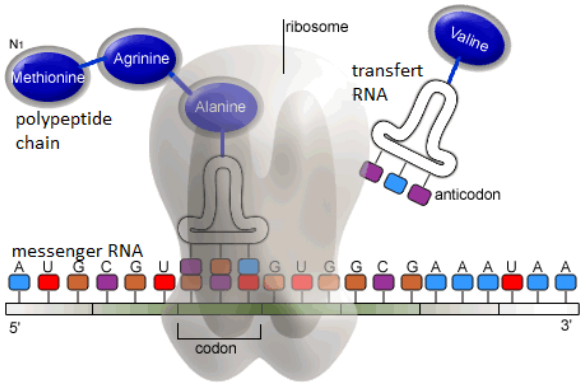


Figure 2: The translation of mRNA to polyprotein chain initiated by a ribosome (Avissar et al. 2018).

1. The *translation initiation step* involves the recognition of the unique initiator codon (AUG) of the mRNA and the binding of the two ribosome subunits, forming a competent ribosome for translation.
2. The *elongation step* corresponds to the movement of the ribosome along the mRNA and the sequential attachment of the amino acids of the newly biosynthesized protein.
3. During the *termination step*, when the ribosome reaches a stop codon (UAG, UGA or UAA), the synthesized protein is released, and the ribosome's subunits are separated.

### 4.2 Simulation Results of the SARS-COV-2 Proteins Translation

#### 4.2.1 Ontological Model of the SARS-COV-2 Proteins Translation

As described above, the translation process involves decoding the messager RNA by means of codons (triplet of nucleobases). Since each *Codon* will be associated to an *Amino Acid*, the sequence of codons will form a chain that produces a *Protein*. Besides this codon definition, our ontology takes into account the biological phenomenon of ribosomal frameshifting. Considering the coronaviruses, sometimes the ribosome slips back one nucleobase and continues the translation process enabling viruses to pack more information into their genomes (Dinman 2006). Therefore, the Data property (role) *frameshifting (Nucleobase, xsd:boolean)* is added to the ontology and describes where the frameshifting occurs. *sequence_polyprotein_pp1a (Codon, Codon)* and *sequence_polyprotein_pp1ab (Codon, CodonFS)* are also defined to delimit respectively the polyproteins sequences of pp1a and pp1ab. The class *CodonFS* describes codons generated after the frameshifting.

The axioms formalizing the twenty-two amino acids formed by the combination of three nucleobases (codon) are based on description logic. Table 2 shows an example of axiom formalization for amino acids: Methionine and Asparagine.

Table 2: Axioms to infer amino acids from codons.

| Example of amino acids definition |
|---|
| Methionine ≡ Global_codon ⊓ (∃ has_first_base.Adenine) ⊓ (∃ has_second_base. Uracil) ⊓ (∃ has_third_base.Guanine) |
| Asparagine ≡ Global_codon ⊓ (∃ has_first_base.Adenine) ⊓ (∃ has_second_base.Adenine) ⊓ (∃ has_third_base. (Cytosine ⊔ Uracil)) |

Rules are developed to handle the transitions that occur during the translation process, such as the codons definition or the ribosomal frameshifting. We exploit SWRL (Horrocks et al. 2004) a Semantic Web Rule Language to formalize each rule. A rule is a conjunction of predicates forming a head (consequence: the results of inference) and a body (conditions to fulfil and get inferences). In our ontology, nine rules are defined to describe only the translation process. Figure 3 illustrates an example of some rules that infer knowledge related to the ribosomal frameshifting phenomenon (where the frameshifting occurs and codons generated after the frameshifting).

**SWRL Rule 3**

```
Stem_Loop(?s) ^ has_beginning_base(?s, ?b) ^ has_Rank(?b, ?r1) ^ Codon(?x) ^ Codon(?y) ^ Codon(?z) ^
has_next(?x, ?y) ^ has_next(?y, ?z) ^ has_third_base(?x, ?u1) ^ Uracil(?u1) ^
has_first_base(?y, ?u2) ^ Uracil(?u2) ^ has_second_base(?y, ?u3) ^ Uracil(?u3) ^
has_third_base(?y, ?a1) ^ Adenine(?a1) ^ has_first_base(?z, ?a2) ^ Adenine(?a2) ^
has_second_base(?z, ?a3) ^ Adenine(?a3) ^ has_third_base(?z, ?c) ^ Cytosine(?c) ^
has_Rank(?u1, ?r2) ^ swrlb:subtract(?fs, ?r1, ?r2) ^ swrlb:lessThanOrEqual(?fs, 15 )
 => frameshifting(?u1, true)
```

**SWRL Rule 4**

```
Nucleobase(?u) ^ frameshifting(?u, true) ^ has_Rank(?u, ?r0) ^ Nucleobase(?x) ^ Nucleobase(?y) ^
Nucleobase(?z) ^ has_next(?x, ?y) ^ has_next(?y, ?z) ^ has_Rank(?x, ?r1) ^ has_Rank(?z, ?r3) ^
swrlb:mod( 0 , ?r1, 3 ) ^ swrlb:greaterThanOrEqual(?r1, ?r0) ^ CodonFS(?c) ^ has_Rank(?c, ?r4) ^
swrlb:add(?s, ?r3, 1 ) ^ swrlb:divide(?d, ?s, 3 ) ^ swrlb:equal(?d, ?r4)
=> has_first_base(?c, ?x) ^ has_second_base(?c, ?y) ^ has_third_base(?c, ?z)
```

**SWRL Rule 5**

```
CodonFS(?c1) ^ has_third_base(?c1, ?x) ^ CodonFS(?c2) ^ has_first_base(?c2, ?y) ^ has_next(?x, ?y)
=> has_next(?c1, ?c2)
```

Figure 3: Ribosomal frameshifting rules.

After enriching the data with the genome sequence (sequence of nucleobases representing the mRNA as an input), the reasoning process infers several information (as an output) including the main knowledge in this context of protein synthesis which is the polyproteins delimitation of pp1a and pp1ab (chain of amino acids).

### 4.2.2 DEVS Model of the SARS-COV-2 Proteins Translation

By applying the steps described in Section 3, we converted the ontological model of the SARS-CoV-2 proteins translation into a DEVS model. The obtained ribosome DEVS model (Eq. (1)) is described as follows:

$$HR = (X_{HR}, Y_{HR}, S_{HR}, \delta_{ext_{HR}}, \delta_{int_{HR}}, \lambda_{HR}, ta_{HR}) \tag{1}$$

**The set of inputs ($X_{HR}$)** The set of input values represents information about the presence of both the messenger RNA (mRNA) and transfer RNA (tRNA) via HRIn1 and HRIn2 inputs, respectively. These input ports are modeled by concepts in the ontological model. Thus, the set of input values of the HR model ($X_{HR}$)

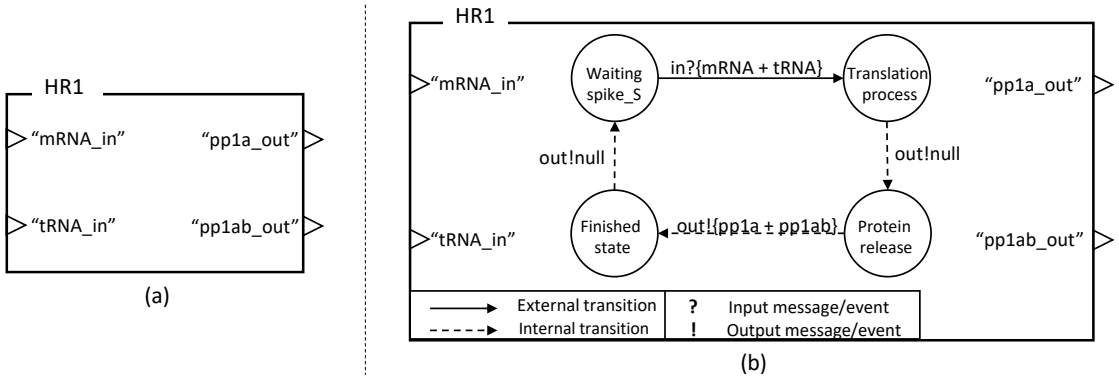


Figure 4: The first host ribosome (HR1) atomic model (a) and its different states (b).

can be defined as follows (Eq. (2)):

$$X_{HR} = \{(HRIn1, mRNA), (HRIn2, tRNA)\} \tag{2}$$

Where: $\{HRIn1, HRIn2\} \in HRIP$ the HR inputs, $mRNA \in X_{HR,HRIn1}$ and $tRNA \in X_{HR,HRIn2}$. The *set of HR input ports* denoted by $HRIP$ consists of two input ports ($HRIP = \{HRIn1, HRIn2\}$), where $HRIn1$ and $HRIn2$ the input ports by which the the mRNA and tRNA are presented to the ribosome, respectively.

**The set of outputs ($Y_{HR}$)** After the translation process, the host ribosome will provide two polyproteins $pp1a$ and $pp1ab$. These two polyproteins will be released through the output ports $pp1a\_out$ and $pp1ab\_out$. In the ontological model, these output ports are represented by concepts. Thus, the set of HR outputs ($Y_{HR}$) can be defined as follows (Eq. (3)):

$$Y_{HR} = \{(HROut1, pp1a), (HROut2, pp1ab)\} \tag{3}$$

Where $pp1a \in Y_{HR,HROut1}$, $pp1ab \in Y_{HR,HROut2}$, and $\{HROut1, HROut2\} \in HROP$ the HR outputs. The *set of output ports* denoted by $HROP$ consists of two output ports ($HROP = \{HROut1, HROut2\}$) where $HROut1$ and $HROut2$ are the two output ports of the HR by which the two polyproteins $pp1a$ and $pp1ab$ are released.

**The set of states ($S_{HR}$)** As shown in Figure 4b the HR atomic component has four states, (i) the *WAITING* state: a passive state in which the ribosome is waiting for both *mRNA* and *tRNA*. (ii) The *TRANSLATION_PROCESS* state: a state in which the biosynthesis of proteins is being performed. The duration of this state is defined by the time advance function. (iii) The *PROTEIN_RELEASE* state: a transitory state in which the HR releases the synthesized proteins $pp1a$ and $pp1ab$ in the cytoplasm. And, finally (iv) the *FINISHED* state: a transitory state showing that the HR finished its task and will return to the initial passive state. Thus, the set of states of HR ($S_{HR}$) can be defined as follows (Eq. (4)):

$$S_{HR} = \{"WAITING","TRANSLATION\_PROCESS","PROTEIN\_RELEASE","FINISHED"\} \times \mathbb{R}_0^+ \tag{4}$$

**The external transition function ($\delta_{ext_{HR}} : Q \times X_{HR} \rightarrow S_{HR}$)** This function makes a state transition when an external event happened. For the host ribosome HR, the external events are reflected by the presence and attachment of the *mRNA* and *tRNA* on the HR input ports, $HROut1$ and $HROut2$ (Figure 4b). The HR external transition function is defined as follows (Eq. (5)):

$$\delta_{ext_{HR}}(phase, \sigma, e, x) = \begin{cases} (TRANSLATION\_PROCESS, translationProcessTime), & \text{if } phase = "WAITING" \\ (phase, \sigma - e), & \text{if } phase \in \{"TRANSLATION\_PROCESS","PROTEIN\_RELEASE", \\ & \quad "FINISHED"\} \end{cases} \tag{5}$$

Where $Q = \{(s,e)|s \in S_{HR}, \text{and } 0 < e < ta_{HR}(s)\}$ is the set of total states, $e$ is the elapsed time in the state $s$, and $\sigma$ the resting time in the current state. These states are defined by SWRL rules in the ontological model. For example, the TRANSLATION_PROCESS is computed by the different rules presented in the previous section (Section 4.2.1).

**The internal transition function ($\delta_{int_{HR}} : S_{HR} \rightarrow S_{HR}$)** The internal transition function defines the next state for the host ribosome HR, as a result of the elapsed time without an external event has taken place. Once, the *TRANSLATION_PROCESS* state is successfully done, an internal transition is required to change the state of the HR from the translation state to the *PROTEIN_RELEASE* state. Similarly, once the polyproteins $pp1a$ and $pp1ab$ have been released in the cytoplasm, an internal transition is necessary to change the HR state to *FINISHED*. A last internal transition switches the HR state to *WAITING*. The HR internal transition function is defined as follows (Eq. (6)):

$$\begin{aligned}
&\delta_{int_{HR}}(\text{"}TRANSLATION_PROCESS\text{"}, translationProcessTime) = (\text{"}PROTEIN_RELEASE\text{"}, proteinReleaseTime)\\
&\delta_{int_{HR}}(\text{"}PROTEIN_RELEASE\text{"}, proteinReleaseTime) = (\text{"}FINISHED\text{"}, finishedTime)\\
&\delta_{int_{HR}}(\text{"}FINISHED\text{"}, finishedTime) = (\text{"}WAITING\text{"}, \infty)
\end{aligned} \tag{6}$$

**The output function ($\lambda_{HR} : S_{HR} \rightarrow Y_{HR}$)** The output function generates an external output just before an internal transition takes place. It is defined as follows (Eq. (7)):

$$\begin{aligned}
&\lambda_{HR}(\text{"}TRANSLATION_PROCESS\text{"}, \sigma) = \emptyset\\
&\lambda_{HR}(\text{"}PROTEIN_RELEASE\text{"}, \sigma) = (pp1a, pp1ab)\\
&\lambda_{HR}(\text{"}FINISHED\text{"}, \sigma) = \emptyset\\
&\lambda_{HR}(\text{"}WAITING\text{"}, \sigma) = \emptyset
\end{aligned} \tag{7}$$

**The time advance function ($ta_{HR} : S_{HR} \rightarrow \mathbb{R}^{+}_{0,\infty}$)** This function defines the time that the HR is expected to spend in each state. This function is not supported by the ontological model. The time advance function of HR is defined as follows (Eq. (8)):

$$\begin{aligned}
&ta_{HR}(\text{"}TRANSLATION_PROCESS\text{"}, \sigma) = translationProcessTime\\
&ta_{HR}(\text{"}PROTEIN_RELEASE\text{"}, \sigma) = proteinReleaseTime\\
&ta_{HR}(\text{"}FINISHED\text{"}, \sigma) = finishedTime = 0\\
&ta_{HR}(\text{"}WAITING\text{"}, \sigma = \infty
\end{aligned} \tag{8}$$

### 4.3 Benefits

The implementation of the SARS-CoV-2 proteins translation model was carried out using the CD++ Builder toolkit (Wainer 2002), a platform based on the DEVS formalism. Figure 5 shows a snapshot of the simulation results in CD++. The figure depicts the different steps of the replicative cycle of the virus: the black frame illustrates the creation and initialization of the ribosome in a waiting state. The green frame represents the external transition function that changes the state of the ribosome to the translation process state due to the availability of both mRNA and tRNA. The first blue frame displays the internal transition occurring after the synthesis of proteins. Next the output function, represented by the red frame, sends the synthetized proteins (pp1a and pp1ab) produced by the ribosome to the output ports of the ribosome. Finally, the two blue frames display the internal transition function switching the state of the ribosome from the proteins release state to the finished state and then to the initial waiting state. Through the ontological modeling, the DEVS simulator

considers micro-level phenomena, such as the definition of codons of genomic sequences, the modeling of many types of translational frameshifting, etc. Such micro-level phenomena cannot be supported by the DEVS model. While the DEVS simulation describes in detail the different states of the HR atomic model over time. Additionally, the DEVS model allows computing the production percentage of each polyprotein. According to Snijder et al. (Snijder et al. 2003), a large part of the ribosomal production (about 70%) is dedicated to the production of the pp1a protein and only 30% is dedicated to the production of pp1ab.

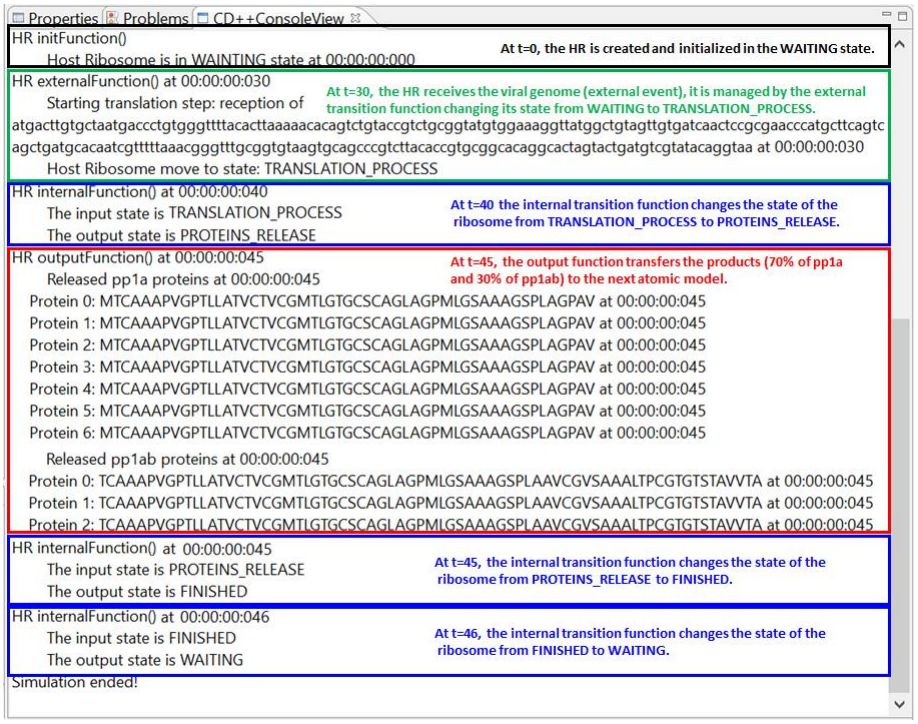


Figure 5: Simulation of the obtained DEVS model in the CD++Builder development environment.

To conclude, this hybrid approach exploits the advantages of the ontological model to enrich the DEVS simulation model. Among these benefits, we can mention the consideration of a very detailed level regarding the assembly of codons, the modeling of frameshifting, the modeling of mutations in genomic sequences, and so on. The proposed DEVS modeling and simulation approach uses the ontological concepts, axioms and SWRL rules to compute and model the main functions and behaviour of the molecular components involved in the replication cycle of SARS-CoV-2. Moreover, the proposed modeling and simulation approach is sufficiently flexible to be applied to other biological and health applications, such as the propagation of infectious diseases among a group of individuals, to establish appropriate prevention measures, and limit the spread of this epidemic, where contamination is a social issue; as well as to model and simulate the multi-scale mechanical behavior of living materials and multi-physical phenomena.

## 5 CONCLUSIONS AND FUTURE WORK

The need to integrate the multiplicity of knowledge and scales of description for modeling complex systems calls for combining knowledge representation and simulation methods. To face the complexity of modeling and simulating the SARS-CoV-2 replication machinery, we proposed an approach combining these two domains, ontology engineering, and modeling and simulation (especially the DEVS formalism). Thus, the challenge addressed in this paper is to mix both domains in a common hybrid approach to gain a micro-macro modeling and simulation approach for understanding the SARS-CoV-2 replication machinery. In

this approach, biological knowledge at the micro-level of the viral system is capitalized and inferred from ontological models, while the complex dynamic behavior of SARS-CoV-2 molecular mechanisms and their different state changes in time are modeled by DEVS models.

Future work on the proposed hybrid approach should focus on the development of advanced Cell-DEVS models, timed cellular model specification based on DEVS with explicit timing delays, to simulate the multiscale replication of the SARS-CoV-2 in cellular tissues. This multilevel simulation model may consider the host immune response, and cellular tissue damage in both time and space. Besides, the proposed approach needs additional work in verification and validation for checking its vitality. An issue that we hope to explore shortly.

## ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support from the project-based funding agency for research in France, Agence Nationale de la Recherche (ANR) under grant agreement No. ANR-20-COVI-0006-01, and the REACTing Inserm.

## REFERENCES

Avissar, Y., J. Choi, J. DeSaix, V. Jurukovski, R. Wise, and C. Rye. 2018. "Biology: OpenStax".

Best, K., and A. S. Perelson. 2018. "Mathematical Modeling of Within-Host Zika Virus Dynamics". *Immunological Reviews* vol. 285 (1), pp. 81–96.

BioPortal 2021. "OntoRepliCov Ontology". https://bioportal.bioontology.org/ontologies/ONTOREPLICOV. Accessed May 28, 2021.

Centers for Disease Control Prevention, . 2020. "Similarities and Differences Between Flu and COVID-19". *Retrieved July* vol. 13, pp. 2020.

Davey, N. E., G. Travé, and T. J. Gibson. 2011. "How Viruses Hijack Cell Regulation". *Trends in Biochemical Sciences* vol. 36 (3), pp. 159–169.

de Oliveira, L. P., D. Hudebine, D. Guillaume, and J. J. Verstraete. 2016. "A Review of Kinetic Modeling Methodologies for Complex Processes". *Oil & Gas Science and Technology–Revue d'IFP Energies nouvelles* vol. 71 (3), pp. 45.

Dinman, J. D. 2006. "Programmed Ribosomal Frameshifting Goes Beyond Viruses: Organisms From All Three Kingdoms Use Frameshifting to Regulate Gene Expression, Perhaps Signaling a Aaradigm Shift.". *Microbe (Washington, D.C.)* vol. 1 (11).

Fabreti, L. G., D. Castro, B. Gorzoni, L. M. R. Janini, and F. Antoneli. 2019. "Stochastic Modeling and Simulation of Viral Evolution". *Bulletin of Mathematical Biology* vol. 81 (4), pp. 1031–1069.

GitHub 2021. "DEVS Model". https://github.com/AliAyadi/DEVS_Model_SARS{-}CoV{-}2_replication. Accessed May 28, 2021.

Handel, A., L. E. Liao, and C. A. Beauchemin. 2018. "Progress and Trends in Mathematical Modelling of Influenza A virus Infections". *Current Opinion in Systems Biology* vol. 12, pp. 30–36.

Hattaf, K., A. M. Elaiw, A. A. Lashari, and N. Yousfi. 2018. "Mathematical Modeling in Virology by Differential Equations". *International Journal of Differential Equations* vol. 2721481, pp. 1687–9643.

Horrocks, I., P. F. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, and M. Dean. 2004. "SWRL: A Semantic Web Rule Language Combining OWL and RuleML". *W3C Member Submission* vol. 21 (79), pp. 1–31.

Li, Z., Z. Teng, and H. Miao. 2017. "Modeling and Control for HIV/AIDS Transmission in China Based on Data From 2004 To 2016". *Computational and Mathematical Methods in Medicine* vol. 2017.

Madelain, V., S. Baize, F. Jacquot, S. Reynard, A. Fizet, S. Barron, C. Solas, B. Lacarelle, C. Carbonnelle, and F. Mentré. 2018. "Ebola Viral Dynamics in Nonhuman Primates Provides Insights Into Virus Immuno-Pathogenesis and Antiviral Strategies". *Nature Communications* vol. 9 (1), pp. 1–11.

Sarnow, P., R. Cevallos, and E. Jan. 2005. "Takeover of Host Ribosomes by Divergent IRES Elements". *Biochemical Society Transactions* vol. 33 (6), pp. 1479–1482.

Schmeing, T. M., and V. Ramakrishnan. 2009. "What Recent Ribosome Structures Have Revealed About The Mechanism of Translation". *Nature* vol. 461 (7268), pp. 1234–1242.

Shang, J., Y. Wan, C. Luo, G. Ye, Q. Geng, A. Auerbach, and F. Li. 2020. "Cell Entry Mechanisms of SARS-CoV-2". *Proceedings of the National Academy of Sciences* vol. 117 (21), pp. 11727–11734.

Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. Poon, Y. Guan, M. Rozanov, W. J. Spaan, and A. E. Gorbalenya. 2003. "Unique and Conserved Features of Genome and Proteome of SARS-Coronavirus, an Early Split-Off From the Coronavirus Group 2 Lineage". *Journal of Molecular Biology* vol. 331 (5), pp. 991–1004.

Wainer, G. 2002. "CD++: a Toolkit to Define Discrete-Event Models". *Software, Practice and Experience* vol. 32 (3), pp. 1261–1306.

## AUTHOR BIOGRAPHIES

**ALI AYADI** is a post-doctoral researcher at Aix Marseille Université and a member of the Laboratoire d'Informatique et des Systèmes (LIS), Marseille, France. He is working on conceptual representation and knowledge based simulation. His email address is ali.ayadi@lis-lab.fr.

**CLAUDIA FRYDMAN** is a full Professor at Aix-Marseille Université. She is also a member of the Laboratoire d'Informatique et des Systèmes (LIS), she has been a referee for several scientific journals and a member of the program committee in various international conferences. She is working on knowledge based simulation. Her email address is claudia.frydman@lis-lab.fr.

**WISSAME LADDADA** is a post-doctoral researcher at Aix Marseille Université in collaboration with INSA Rouen Normandie. She is working on knowledge based systems. Her email address is wissame.laddada@univ-amu.fr.

**LINA F. SOUALMIA** is an Associate Professor in Computer Science at Université de Rouen and member of the TIBS team at the LITIS Laboratory. She is working in the domain of Artificial Intelligence applied to Health. Her email address is lina.soualmia@litislab.fr.

**CECILIA ZANNI-MERK** is a Full Professor in Computer Science at INSA Rouen Normandie and the head of the MIND team at the LITIS laboratory. Her research focuses in conceptual representation and inference processes applied to problem-solving. Her email address is cecilia.zanni-merk@insa-rouen.fr.

**INDIA LHOTE** is an engineer in biological experimentation and instrumentation at the AFMB Laboratory. Her email address is india.lhote@univ-amu.fr.

**EMELINE GRELLET** is a Ph.D. student working on host-pathogen interactions and member of the AFMB Laboratory. Her email address is isabelle.imbert@univ-amu.fr.

**ISABELLE IMBERT** is a Full Professor in Biology at Aix-Marseille Université and member of the Architecture and Function of Biological Macromolecules Laboratory (AFMB - CNRS/Aix-Marseille University). She is considered as one of the French experts on coronavirus. Her email address is isabelle.imbert@univ-amu.fr.