Technical Report CoSBi 06/2007

Simulating a faulty mechanism of protein folding in the pathogenesis of familial Parkinson's disease

Paola Lecca

The Microsoft research - University of trento
Centre for Computational and Systems Biology

lecca@cosbi.eu

This is the preliminary version of a paper that will appear in
Online Journal of Bioinformatics 9 (1):30-43, 2008
available at http://www.cpb.ouhsc.edu/ojvr/bioinfo.htm#20072

Simulating a faulty mechanism of protein folding in the pathogenesis of familial Parkinson's disease

Paola Lecca

Abstract

A growing body of evidence suggests that the accumulation of misfolded proteins in brain tissues is a crucial event in the Parkinson's disease neurodegeneration. Pathogenic mutations may directly induce abnormal protein conformations or compromise the ability of the cellular machinery to detect and degrade misfolded proteins. Although the recent explosion in the rate of discovery of genetic defects linked to Parkinson's disease (PD) have provided tangible clues to the neurobiology of the disorder, they have provided neither direct explanation for the disease process or its key biochemical mechanism. The aim of the work is to provide *quantitative* models for *in silico* experiments, that can help the researchers either to elucidate important and still elusive aspects of the Parkinson's disease or to design new wet-experiments.

Here we present three stochastic models of a faulty mechanism of protein re-folding and degradation of misfolded proteins. Our models are specified in biochemical stochastic π-calculus and are based on what is currently known about the genetic mutations causing PD. The expressive capabilities of this formalism in the description of parallel and competive nature of biochemical interactions make it particularly suitable for modeling the intricate mechanism of proteins folding, refolding and eventually degradation. Furthermore, the simulation results point out those kinetic quantitative parameters, whose variations lead to significant changes in the capability of the system to react to the accumulation of dangerous proteins.

1 Introduction

Parkinson's disease (PD) is a chronic, progressive movement disorder, that results from the degeneration of dopamine-producing nerve cells in the substantia nigra. Dopamine is a neurotransmitter that stimulates motor neurons, those nerve cells that control the muscles. When dopamine production is depleted, the motor system nerves are unable to control movement and coordination. Parkinson's disease patients have lost 80% or more of their dopamine-producing cells by the time symptoms appear. The inherited forms of Parkinson are relatively rare but may provide clues to the biological origins of the more common forms of the disease. More specifically, the rationale for studying the rare genetic forms of PD is the expectation that the phenotypic similarity between the genetic and sporadic forms indicates that they share important pathogenic mechanisms, and as consequence, that genetic information will help focus research on a key biochemical pathway responsible for the disease.

Although the scientists have not yet found the exact cause of PD, recent years have seen an explosion in the rate of discovery of genetic defects linked to the pathogenesis of PD. In the mid-1990s three missense mutations in the gene encoding α-synuclein were identified as a possible cause of a dominantly inherited PD [18]. None of these mutations has been found in sporadic PD or individuals without the disease. At the end of 1990s another protein, whose mutation has been associated to PD, has been identified: the parkin. Loss-of-function mutations in the gene encoding parkin cause a recessively inherited form of PD, that usually occurs under age 40 [7].

The causes of these mutations are believed to result from the exposure to environmental toxins, for example pesticides, that inhibits dopamine production and produce free radicals and oxidation damages. The effect of the pathogenic α-synuclein is the inhibition of the degradation of misfolded proteins [11, 18], whose accumulation in the dopaminergic neurons is the determining factor of their death. However, how mutant α-synuclein and parkin variants exactly produces neurotoxicity remains unknown, in part because the theoretical and experimental efforts made to understand the proteins function and the reasons of the block of the proteolitic machinery are just at the beginning. At the present there are no available mathematical or computational models or computer simulated experiment that can aid the researchers to unravel the molecular basis of the neurodegenerative processes of PD. This work proposes two gene-based models of PD and a hypothetical model that relates the onset of the disease to an insufficient quantity of chaperones. The gene-based models concern PD caused by dominant mutations in the α-synuclein gene and PD caused by recessive mutations in the parkin gene. The hypothetical model intends to suggest a third possible mechanism in which the presence of an insufficient quantity of chaperones promotes the accumulation of misfolded proteins. All the three models has

been specified in biochemical stochastic π-calculus [15] and simulated with SpiM (Stochastic PI-Machine) [13]. The choice of such a formalism is motivated by the capability of expressing the parallel and concurrent nature of biological interaction, especialliy at molecular level and the modularity and the adaptability of a biological system as well. Common approaches as those based on differential equations revealed to be not so expressive, especially for handling concurrency and adaptability, i. e. the ability of the system to reconfigure itself in rensponse to environmental stimuli (see [9, 10] for a detailed comparison of the π-calculus with ordinary differential equation formalism).

2 Mechanisms of neurodegeneration

In patients with inherited PD, pathogenic mutations are though to cause disease directly by inducing abnormal and toxic protein conformations [2] or indirectly by interfering with the processes that normally target for degradation the misfolded proteins. The triggers for disfunctional protein metabolism may be oxidative stress. The tissue content of oxidized proteins, which may misfold, increases with age [1], and neurons may be susceptible because they are postmitotic. At the present a definitive diagnosis of PD can only be made by autopsy, and it is based on the presence of intraneuronal proteic inclusions called Lewy bodies (LBs) and on the loss of nigrostriatal dopaminergic neurons In PD caused by dominant mutations of α-synuclein, LBs contain a significant amount of the oxidatively modified variant of this protein. The mutations in the gene encoding parkin protein cause a recessively inherited form of PD. Pathologically, this form of familial PD is associated with a loss of nigrostriatal neurons, but LBs are not typically observed [3]. The ability of the cell to handle misfolded proteins is expressed by some complexes of macromolecules, called *chaperones*.

Molecular chaperones interact with unfolded or partially folded protein subunits, e.g. nascent chains emerging from the ribosome, or extended chains being translocated across sub-cellular membranes. They prevent inappropriate association or aggregation of exposed hydrophobic surfaces and direct their substrates into productive folding, transport or degradation pathways. Essential for viability and cell survival, the expression of the molecular chaperones is often increased by cellular stress.

In the healthy cells, if a protein does not assume the correct 3D shape, or a cellular stress induces a right-folded protein to assume a wrong folding, the chaperones re-shape it correctly. In the case in which the protein is not correctly refolded, the cellular proteasome - a system designated to the digestion of cell wastes - degrades it before the faulty protein can cause damages. First the protein *parkin* attaches molecules of *ubiquitin* to the misfolded protein; once the ubiquitin is bound to the faulty protein, it sig-

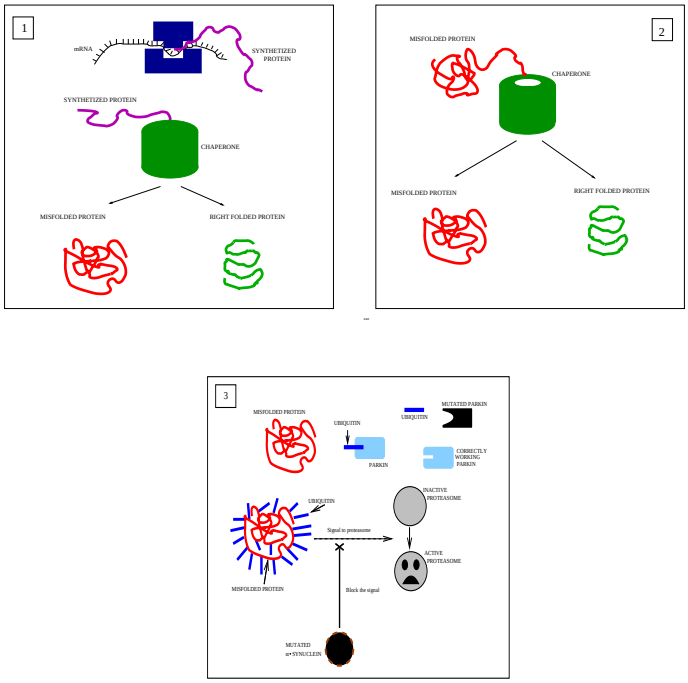

Figure 1: Pathogenesis of PD induced by mutant α-synuclein: 1. the intereation of a nascent protein with a chaperone can results in a right-folded protein or in a misfolded protein; 2. the chaperone attempts to re-fold the faulty protein and the result can be again a right-folded protein or a misfolded protein; 3. therefore,the misfolded protein is drapped by the ubiquitin transported by the parkin protein. A mutant variant of the parkin is not able to transport the ubiquitin on the misfolded protein. The mutant α-synuclein inhibits the activation of the proteasome by the ubiquitin. The mutant α-synuclein seems to be proteasome-proof, but the model presented in this paper takes into account an eventual attempt of the proteasome to attack the faulty α-synuclein. The outcomes of the interactions between the nascent linearprotein and the chaperone, as well as of the interaction between the mutant α-synuclein and the proteasome are stochastically determined by the reaction probabilities derived from the kinetic reaction rates accordingly to the Direct Gillespie algorithm [6]

.

nals to the proteasome of decomposing the protein into its amino-acids, that will be employed somewhere else in the cell. A mutation of the gene of the α-synuclein gives rise to a proteasome-proof variant of α-synuclein, that perturbatively interferes with the communication between ubiquitin and proteasome [11]. If the ubiquitin can not enable the activity of the proteasome, the misfolded protein can not be degraded. Fig. 1 depicts the mechanisms

of neurodegeneration triggered by the mutant variant of α-synuclein.

Some chaperones are non-specific, and interact with a wide variety of polypeptide chains, but others are restricted to specific targets. With aging, the ability of cells to induce a variety of chaperones is impaired as is the activity of the proteasome. Proteasomal disfunctions and the accumulations of misfolded proteins are involved in a vicious cycle, with excess of misfolded proteins inhibiting an already compromised proteasome.

Finally, the mutations of the encoding for parkin, produce a variant of parkin unable to transport the molecules of ubiquitin on the misfolded proteins, that being untargeted, can not be recognized by the proteasome for the degradation.

3 Modeling and simulating in SpiM

The network of interactions that transform a linear protein, turned out by the translation of its messenger RNA in the ribosome, into a functional 3D spatial structure, has been specified in the formalism of biochemical π-calculus.

This abstract formal language, initially developed for the specification of concurrent computational processes, revealed particularly suitable to describe biological molecular systems [9, 8, 10, 15]. A mathematical theory of concurrent processes has been built in the π-calculus [12]. In this calculus, each process, belonging to a set of interacting processes, is defined by its potential communication activities and may be composed either sequentially or concurrently with other processes. Communications occur via channels, indicated by their names. The π-calculus process algebra is an expressive and efficient formal language for modeling biochemical processes. In such systems, multiple processes interact with each other on complementary shared communication channels by sending and receiving messages in a synchronized way. The π-calculus can be used to model a system of interacting bio-molecules, treating molecules and their individual domains as computational processes, where their complementary structural and chemical determinants correspond to communication channels. Moreover chemical interaction and subsequent modification coincide with communication and channel transmission. Finally, the simulation of the dynamic behavior of the system is defined by the operational semantics of the language.

A significant extension of the π-calculus was realized by Priami [14], by developing a stochastic variant of the original operational semantics of the calculus. This stochastic variant introduces the possibility to assign different rates to each involved biochemical reaction, from which it is possible to derive the probability of occurrence, and from there the reaction waiting time, of the different reactions. On the basis of that information it is then possible to implement a race condition, establishing which is the next

reaction and when it occurs.

To summarize and to introduce the essential notation, the most basic process form is a choice $\sum = P_1 + \cdots + P_N$, among zero or more actions exhibited by the processes Ps composing the sum $\sum$. The simplest process is the deadlock, that is a process that can do nothing (usually denoted by $\mathbf{0}$. An action π can be an *output* $\overline{x}(n)$, or an *input* $x(m)$, or a *delays* τ that the process can perform. x is the channel through which the output message n is sent. A process R defined by $R = \pi.R'$, is a process prefixed by an *unguarded* action, i. e. a process offering to perform the action π. Once this action is performed the process R changes to the state defined by the process R'. Two processes P and Q can be combined using parallel composition $P|Q$. Moreover, another basic operator of stochastic π-calculus is the *new* operator (indicated by ν followed by name or a set of channel names). It allows the creation of *fresh* channels. A fresh channels is a channel different from any other channel defined in the system. If a fresh name is defined in the body of a process, that channels a private channel for that process, i. e. it lives in the scope of that process. The delays represent a single communication on a fresh channels, that correspond to an internal evolution of the process. While sending and receiving messages on shared channels represent bi-molecular reactions, delay actions correspond to mono-molecular reactions of to changes of state or conformation of a given chemical. Table 1 lists the axioms of the semantics of biochemical stochastic π-calculus.

In this paper the simulations were obtained using the Stochastic Pi Machine (SpiM), which is the most recent simulator for the biochemical stochastic π-calculus. The simulator in extensively described in [13]. Here we briefly recall that SpiM simulates a given process P by firstly converting the process to a corresponding simulator data structure, consisting of a list of components $A = \sum_1, \ldots, \sum_N$. These list is processed by a procedure based on the Gillespie algorithm [6], that stochastic determines the next interaction channel x and the corresponding reaction time τ. Once an interaction channel x has been chosen, the simulator randomly selects from the list A a component of the form $\sum + x(m).P$, containing an input on channel x, and a different component of the form $\sum' + \overline{x}(n).Q$ containing an output on x. The selected components can then interact in such a way that P is replaced by $P\{n/m\}$ (i. e. in the body of P the variable m is replaced by n) and Q remains unchanged. Finally, the summations $\sum$ and $\sum'$ are discarded and the $P\{n/m\}$ and Q are added to the remainder of the list A.

Accordingly to the stochastic formulation of the chemical kinetics developed by Gillespie, for a reaction μ, the *propensity* a_μ is a function of its kinetic rate r_μ and the number of individual potential copies of reactions μ involving the same reactant species. As stated in the following equation, the reaction propensity is the product of the reaction "rate" and the number of unique reactant combinations.

$$a_\mu = \begin{cases} r_\mu \times (\#A) \times (\#B) & \text{for bi-molecular reactions}^1 \\ r_\mu \times \frac{(\#A)\times(\#A-1)}{2} & \text{for homodimerization} \\ r_\mu \times \#A & \text{for monomolecular reaction} \end{cases}$$

where $\#A$ and $\#B$ are the numbers of reactants of species A and B in the elementary reaction μ. The Gillespie algorithm calculates explicitly *which* reaction occurs next and *when* it occurs. Both question are answered probabilistically by specifying the probability $P(\mu, \tau) = a_\mu \exp(-\tau \sum_j a_j) d\tau$ that the reaction is μ and it occurs a time τ. $P(\mu, \tau)$ can be expressed as the product of two distributions: the distribution for reactions $P(\mu) = a_\mu / \sum_j a_j$, and th distribution for times $P(\tau)d\tau = (\sum_j a_j) \exp(-\tau \sum_j a_j) d\tau$. The algorithm chooses a reaction according to $P(\mu)$, and the time step τ according to an exponential with parameter $\sum_j a_j$.

SpiM simulator uses the notion of channel *activity* to compute the reaction/communication propensity. The activity of channel x in a list L of processes is defined by

$$Act_x(L) = \left[\mathrm{In}_x(L) \times \mathrm{Out}_x(L)\right] - \mathrm{Mix}_x(L)$$

where $\mathrm{In}_x(L)$ and $\mathrm{Out}_x(L)$ are the number of unguarded inputs and outputs on channel x in L, respectively, and $\mathrm{Mix}_x(L)$ is the sum of $\mathrm{In}_x(\sum_i) \times \mathrm{Out}_x(\sum_i)$ for each summation $\sum_i$ in L. By subtracting $\mathrm{Mix}_x(L)$ from the product of the number of inputs and outputs on x, an eventual interaction between an input and an output belonging to the same summation is avoided.

Finally, the dynamic behavior of a process is driven by a race condition. All activities enabled attempt to proceed, but only the fastest one succeeds. The fastest activity is different on successive attempts because durations are random variables. The continuity of the probabilistic distribution ensures that the probability that two activities end simultaneously is zero. Furthermore, exponential distributions enjoy the *memoryless* property: the time at which a certain transition occurs is independent on the time at which it ever occurred before. Therefore there is no need to record the time elapsed to reach the current state.

3.1 First model: misfolded protein accumulation induced by mutant α-synuclein

The system of bio-molecules and cellular structures driving the proteins folding has been implemented as process SYSTEM, given by the parallel composition of 8 processes, that represent their homonymous molecules: LINEAR_PROTEIN, RIBOSOME, RNA_α_SYNUCLEIN, CHAPERONE, PROTEASOME, UBIQUI-
TIN, PARKIN, and STRESS. The specification showed in Table 2 describes

$$(\ldots + (\overline{x}\langle z\rangle, r).Q) \mid ((x(y), r).P + \ldots) \xrightarrow{x, r_b \cdot 1 \cdot 1} Q \mid P\{z/y\}$$

$$\frac{P \xrightarrow{x, r_b \cdot r_0 \cdot r_1} P'}{P \mid Q \xrightarrow{x, r_b \cdot r_0' \cdot r_1'} P' \mid Q} \quad \begin{cases} r_0' = r_0 + In_x(Q) \\ r_1' = r_1 + Out_x(Q) \end{cases}$$

$$\frac{P \xrightarrow{x, r_b \cdot r_0 \cdot r_1} P'}{(\nu\, x)P \xrightarrow{x, r_b \cdot r_0 \cdot r_1} (\nu\, x)P'}$$

$$\frac{Q \equiv P \quad P \xrightarrow{x, r_b \cdot r_0 \cdot r_1} P' \quad P' \equiv Q'}{Q \xrightarrow{x, r_b \cdot r_0 \cdot r_1} Q'}$$

Table 1: Rules of the semantics of biochemical stochastic π-calculus. A reaction is implemented by the three parameters r_b, r_0 and r_1, where r_b represents the basal rate, and r_0 and r_1 denote the quantities of interacting molecules, and are compositionally computed using the two functions In_x and Out_x defined below. These two functions inductively count the number of receive and send operations on channel x. The first axiom of the BioSpi reduction semantics corresponds to usual reactions involving two different molecules, the second rule corresponds to homo-dimerization reactions, involving the same molecular species. The third rules states that if a process P evolves into a process P' through a communication on channel x, the restriction of channel x to the scope of process P does not affect the transition of P into P'. Finally, the fourth axiom states that if P is structurally congruent to Q and P evolves into P', that is structurally congruent to Q', then Q evolves into Q'.

a faulty non deterministic operation of the chaperones system, whose twofold interaction with a protein can result in a right-folded protein or in a misfolded one according to the value of the associate reaction rate. In this model we assume that the rate of interaction between a chaperone and a linear protein is equal to the rate of a a typical fast protein folding: 1 μs^{-1}. CHAPERONE is the process abstracting the cellular chaperone and LINEAR_PROTEIN is the one representing the linear chain of amino-acids derived by the translation of the messenger RNA. The linear protein process "'physically binds" to the chaperone by sending a private name *to_chaperon* to the process CHAPERONE via the shared channel *bind_to_chaperone*. The result of this interaction is the complex PROTEIN_CHAPERONE, that can undergo two internal modifications: one resulting in a right-folded protein, that in this model is a deadlock process, and the other resulting in a misfolded protein. We have called MISFOLDED' the misfolded protein generated by the interaction between chaperone and linear protein, and MISFOLDED" the misfolded protein generated by the interaction between

chaperone and MISFOLDED'. MISFOLDED" is thus the process representing the protein that has been not correctly re-folded even during a second interaction with a chaperone. Therefore, this misfolded protein is targeted by parkin according to the following sequence of reactions: parkin performs first a physical binding with ubiquitin and then with the misfolded protein. The the sub-component ubiquitin of the trimer formed by parkin, ubiquitin, and misfolded protein, sends a signal to proteaseome to activate it. The physical binding of the parkin with the ubiquitin is modeled as the sending from the process PARKIN to the process UBIQUITIN of the private name *parkin_ubiquitin* on the public channel *bind_to_ubiquitin*. After binding, PARKIN changes to PARKIN_BOUND and, analogously, UBIQUITIN changes to UBIQUITIN_BOUND. The formation of the trimer including parkin, ubiquitin and misfolded protein is modeled as two sequential physical bindings on the public channels *to_misfold1* and *to_misfold2*. In particular, PARKIN_BOUND interacts with MISFOLDED" on *to_misfold1* and UBIQUITIN_BOUND interacts with MISFOLDED" on *to_misfold2*.

In our model we hypothesize that the process STRESS, representing the oxidative cellular stress, interferes with the translation of the transcript of the α-synuclein gene. The messenger RNA of the α-synuclein (RNA_A_SYNUCELIN) after the interaction with RIBOSOME on the channel *bind_to_ribosome* can either be correctly translated (i. e. it performs a communication on channel *translate* and it does not have any other evolution) or it receive a signal from STRESS on channel *stress* and transforms into MUTATED_A_SYNUCLEIN.

The sub-part MISFOLDED"_UBIQUITIN of the trimer midolfed protein-ubiquitin-parkin is defined as a choice: it can either communicate with the process PROTEASOME via the channel *to_proteasome* and activate it for the degradation or it can receive a signal from MUTATED_A_SYNUCLEIN on the channel *signal_from_synuclein* and, consequently, remains intact.

The specification takes into account the attempts to the proteolitic machinery of the cell to degrade the mutant variant of the α-synuclein, by defining the behavior of the process MUTATED_A_SYNUCLEIN as a choice between the communication with MISFOLDED"_UBIQUITIN and the communication on the channel *degrade*, representing the eventuality of a degradation. The kinetic rates and the initial amoutn of reacting molecules are listed in Table 3.

The Figs. 2 (A)-(G) show the time evolution of the amount of MISFOLDED' (solid line) and MISFOLDED" (dotted line) for different values of the initial number of stress processes (N_s) and stress signaling rate (r_s). The plots (A)-(C)-(E)-(G) were generated by fixing at 10 μs^{-1} the value of the stress signaling rate and varying the number of processes STRESS by 10, 100, 200, and 1000 respectively. On the contrary, the plots (B)-(D)-(F)-(H) were generated by fixing at 100 the amount of processes STRESS and

$SYSTEM := LINEAR\_PROTEIN \mid PARKIN \mid UBIQUITIN$
$RNA\_A\_SYNUCLEIN \mid RIBOSOME \mid STRESS$
$PROTEASOME \mid CHAPERONE$

Interaction between parkin and ubiquitin

$PARKIN := (\nu\ parkin\_ubiquitin)\overline{bind\_to\_ubiquitin}\langle parkin\_ubiquitin\rangle.PARKIN\_BOUND$
$PARKIN\_BOUND :=$
$(\nu\ tm1)\overline{to\_misfold1}\langle tm1\rangle.PARKIN\_BOUND\_MISFOLDED''(tm1)$
$PARKIN\_BOUND\_MISFOLDED''(tm1) := \overline{tm1}$
$UBIQUITIN := bind\_to\_ubiquitin(pu).UBIQUITIN\_BOUND(pu)$
$UBIQUITIN\_BOUND(pu') :=$
$(\nu\ tm2)\overline{to\_misfold2}\langle tm2\rangle.UBIQUITIN\_BOUND\_MISFOLDED''(pu', tm2)$
$UBIQUITIN\_BOUND\_MISFOLDED''(pu'', tm2') :=$
$\overline{pu''}.MISFOLDED''\_UBIQUITIN(tm2')$
$MISFOLDED''\_UBIQUITIN(tm2'') :=$
$\overline{to\_proteasome}.degrade + signal\_from\_synuclein.MISFOLDED''\_UBIQUITIN(tm2'')$

Interaction between chaperone and protein

$LINEAR\_PROTEIN :=$
$(\nu to\_chaperon)\overline{bind\_to\_chaperone}\langle to\_chaperon\rangle.PROTEIN\_CHAPERONE(to\_chaperone)$
$PROTEIN\_CHAPERONE(to\_chaperone') := \overline{to\_chaperon'}.MISFOLDED' + \overline{to\_chaperon'}$
$MISFOLDED' := (\nu\ to\_chaperone\overline{bind\_to\_chaperon}MISFOLDED\_CHAPERONE$
$MISFOLDED\_CHAPERONE := \overline{to\_chaperon}.MISFOLDED'' + \overline{to\_chaperon}$
$MISFOLDED'' :=$
$to\_misfold1(tm1').to\_misfold2(tm2').MISFOLDED''\_PARKIN\_UBIQUITIN$
$MISFOLDED\_PARKIN\_UBIQUITIN := tm1.parkin\_ubiquitin.PARKIN$
$CHAPERONE :=$
$bind\_to\_chaperon(to\_chaperone'').CHAPERONE\_BOUND(to\_chaperone'')$
$CHAPERONE\_BOUND(to\_chaperone''') := \overline{to\_chaperone'''}.CHAPERONE$

Translation of protein, perturbation from the process STRESS and communication with proteasome

$PROTEASOME := to\_proteasome.ACTIVE\_PROTEASOME$
$ACTIVE\_PROTEASOME :=!degrade.PROEASOME$
$RNA\_A\_SYNUCLEIN := \overline{bind\_to\_ribosome}.(translate + stress.MUTATED\_A\_SYNUCLEIN)$
$MUTATED\_A\_SYNUCLEIN :=$
$\overline{signal\_from\_synuclein}.MUTATTED\_A\_SYNUCLEIN + degrade$
$RIBOSOME := bind\_to\_ribosome.\overline{translate}.RIBOSOME$
$STRESS := \overline{stress}.STRESS$

Table 2: Stochastic π-calculus specification of a model of a faulty mechanism of protein folding and degradation. If the concent of a send or receive action is not specified, it means it is non relevant.

considering four different values of the stress signaling rate: 0.01, 1, 100, 1000 μs^{-1}, respectively. The proteins MISFOLDED' that are not degraded become proteins MISFOLDED". In Figs. 2 (C)-(E)-(G), after about 5 μs^{-1}, the amount of MISFOLDED" assumes a linear behavior slowly increasing

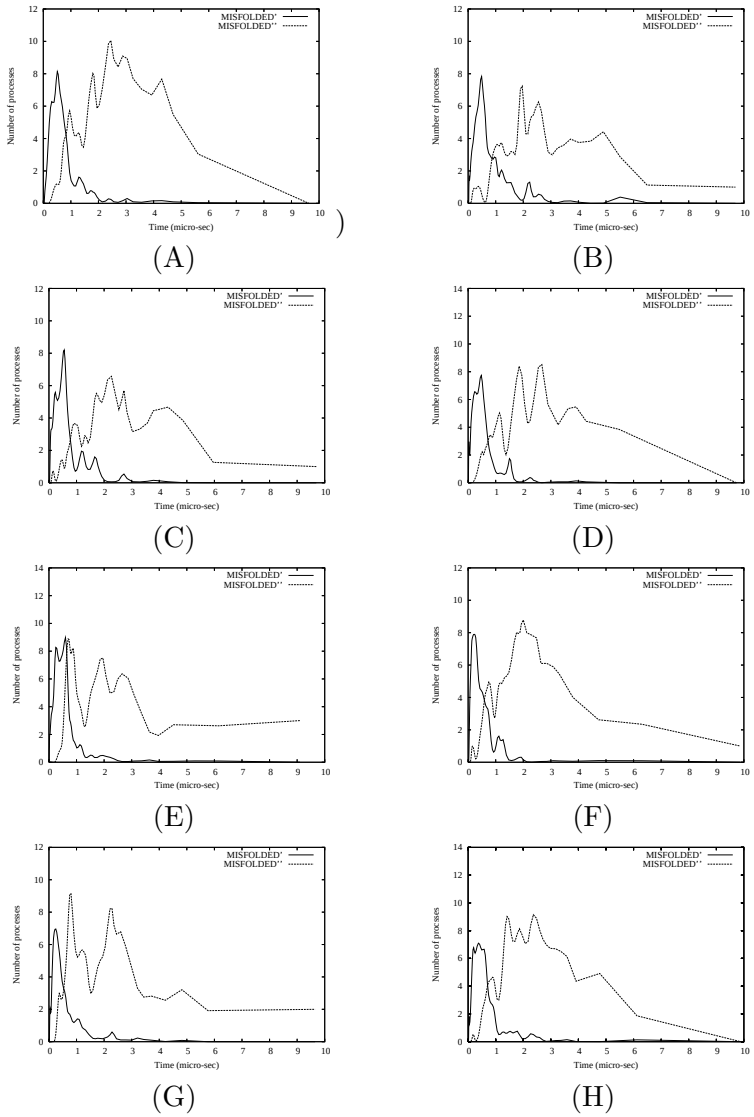

Figure 2: (A) $r_s = 10\ \mu s^{-1}$, $N_s = 10$; (B) $r_s = 0.01\ \mu s^{-1}$, $N_s = 100$; (C) $r_s = 10\ \mu s^{-1}$, $N_s = 100$; (D) $r_s = 1.0\ \mu s^{-1}$, $N_s = 100$; (E) $r_s = 10\ \mu s^{-1}$, $N_s = 200$; (F) $r_s = 100.0\ \mu s^{-1}$, $N_s = 100$; (G) $r_s = 10\ \mu s^{-1}$, $N_s = 1000$; (H) $r_s = 1000.0\ \mu s^{-1}$, $N_s = 100$. The rates used in these simulations has been taken from [4].

with the increase of the number of processes STRESS. In Figs. 2 (D)-(F)-(H) the number of MISFOLDED" decreases independently of the rate of perturbation. Fig. 2 (A) indicates that if $N_s \leq 10$ and $r_s = 10.0\ \mu s^{-1}$, the MISFOLDED" proteins are rapidly degraded after about 4.4 μs^{-1}, as for the cases in Fig. 2 (D)-(F)-(H). This result asserts that in presence of

a scarce STRESS, even if its rate of perturbation is significant, the cell is still able to respond to the accumulation of faulty proteins. These results suggest that the intensity of cellular stress causing the α-synuclein mutations is more significant than the rate at which the stress interferes with the translation of this protein. It means that the probability of producing mutated α-synuclein proteins essentially depends on the amount of stress that interacts with the cell system. A particular attention should be payed to the plot of Fig. 2 (B), showing the simulation obtained with 100 processes STRESS and a stress rate of 0.01 μs^{-1}. With such a rate, after 6 μs, we do not observe the rapid decrease in the number of MISFOLDED", as we can observe in the plots (D)-(F)-(H). Most probably, this fact may reveal the existence of "threshold" phenomena in the onset of the degradation of the misfolded proteins. On the basis of these simulations, the proteolitic machinery seems be activated only for stress signaling rate greater at least than 0.01 μs^{-1}.

Channel	Reaction	Rates (in μs^{-1})
bind_to_chaperon	The native protein interacts with the chaperone	1.0
bind_to_ubiquitin	The parkin binds to the ubiquitin	1.0
to_misfold1	The bound state of the parkin to ubiquitin binds to the misfolded protein	0.1
to_misfold2	The bound state of the ubiquitin to parkin binds to the misfolded protein	0.1
to_proteasome	Signaling from the ubiquitin to the proteasome for the degradation of misfolded protein	10.0
degrade	Degradation of the misfolded protein by the proteasome	10.0
bind_to_ribosome	Binding between native protein and ribosome	10.0
translate	Translation of the native protein	10.0
stress	Perturbative signal from the stress factor	*
signal_from_synuclein	Communication between misfolded protein and mutant α-synuclein	100.0

Process	Number
LINEAR_PROTEIN	100
CHAPERONE	*
PARKIN	100
UBIQUITIN	100
PROTEASOME	100
RIBOSOME	100
RNA_A_SYNUCLEIN	100
STRESS	*

Table 3: Channels rates and number of processes (molecules) used in the models. The fields marked with the symbol "'*" is different in the three considered models (see Fig. 2).

3.2 Second model: misfolded protein accumulation induced by mutant parkin

The mutant parkin does not transfer the ubiquitin to the misfolded protein, that remains untargeted for the degradation. In order to model this situation the specification showed in Table 2 has been modified in the following way. The process PARKIN still communicates with the process UBIQUITIN to realize the binding, but the dimer UBIQUITIN_BOUND formed by this interaction, is now a deadlock process, unable to do any other reactions. Thence, in this model the process PROTEASOME is not necessary, because the reactions of polyubiquitination do not occur any more.

Fig. 3 shows that the number of misfolded proteins derived by a wrong refolding process, presents a rapid increase within the first 5 μs. After that time the number stabilizes about 100.

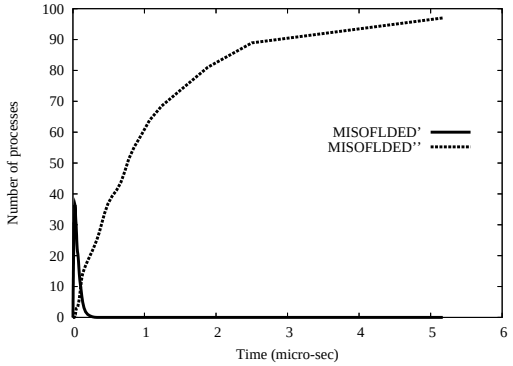

Figure 3: Number of non correctly refolded proteins in PD induced by mutant parkin. The curve of MISFOLDED' zeros before 5 μs, indicating that the production of MISFOLDED" starts since the beginning of the simulation and increases as the square root of the time, without giving to the proteosomal mechanism of the cell any chance to react.

3.3 Third model: misfolded protein accumulation as function of chaperones number

The aim of this experiment is to investigate a possible relationship between the number of chaperones and the ability of the cell of degrading misfolded proteins in the case in which mutant α-synucleins are operating in the system. In particular, the purpose is to validate the thesis for which the greater the number of chaperones is, the lower the rate of decreasing is and consequently the more efficient the ability of the cell to respond to misfolded proteins. By varying the number of instances of process CHAPERONE and setting to 1000 the instances of process STRESS, the plots of Figs. 4 and 5 are obtained.

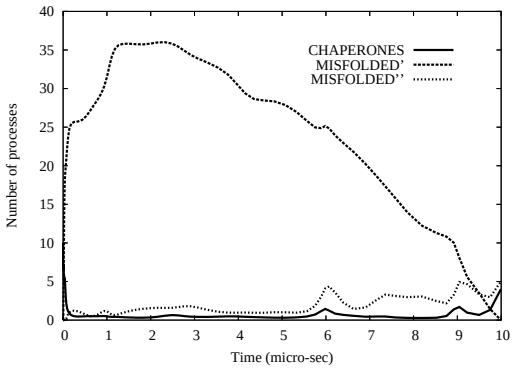

Figure 4: Variation of number of chaperones and wrongly refolded proteins in PD induced by mutant α-synuclein. The initial number of chaperones is 10. This simulation shows that this number is not adequate to defend the cell from the increasing of faulty proteins.

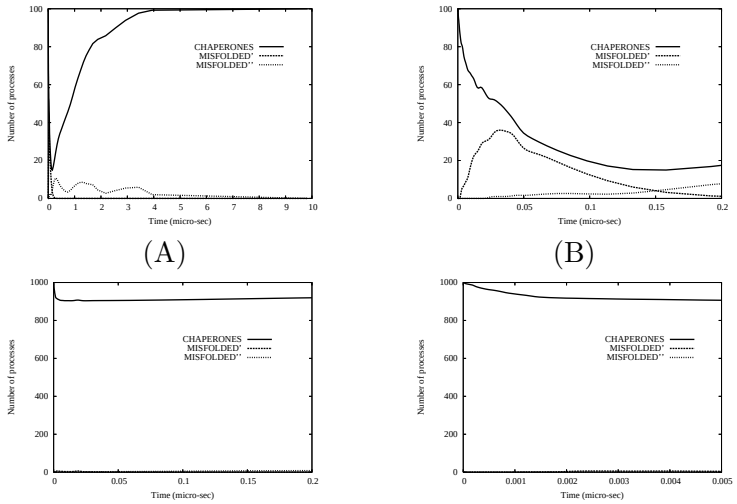

Figure 5: Variation of number of chaperons and wrongly refolded proteins in PD induced by mutant α-synuclein. The initial number of chaperones is 100 in the plot (A) and 1000 in the plot (C). The plots (B) and (D) are a zoom of the plots (A) and (C) to better visualize the time behavior of the processes in the first 0.2 μs^{-1} and 0.005 μs^{-1}, respectively. A sufficiently large number of chaperones seem to ensure the cell the possibility to activate the proteasomes and consequently to decrease the number of faulty proteins.

The plots of Fig. 4 indicates that starting with an initial number of 10 chaperones, the curve of MISFOLDED" has a wide pick decreasing after 4

μs and zeroing about 10 μs, exactly when the number of MISFOLDED" starts to increase. The simulations showed in Fig. 5 (A)-(B) indicate that with an initial number of 100 chaperones, the system is only partially able to control the number of misfolded proteins. Namely, the number of MISFOLDED" zeroes only after 6 μs and, unlike the case of Fig. 4, the number of MISFOLDED" is significantly greater than zero since the beginning of the simulation. On the contrary, the curves of Fig. 5 (C)-(D) shows that with 1000 instances of chaperones, the number of misfolded proteins remains null during all the simulation time. This results confirm what very recent wet experiments are starting to recognize, i. e. that a therapeutically induced increase of chaperones can enhance the cellular environment for protein folding and stability. In particular, in a recent review D. F. Smith, L. Whitesell and E. Katsanis [17] provided an overview of protein misfolding as a basis for disease and have provided a prospective look at pharmacological approaches that may help to prevent or resolve protein-folding problems. In this review they asserted that specific mechanisms for inducing production and accumulation of intracellular chemical chaperones are potentially useful for preventing and correcting protein misfolding; perhaps drugs will be discovered that serve this purpose. Smith et al. recall also that an alternative approach would be to administer, by tissue perfusion, nontoxic chemical chaperones that could be taken up by cells. The feasibility of this approach has been already examined in cell culture systems.

4 Conclusions and future directions

Any advance in understanding the genetic basis of neurodegeneraive disorders like parkinsonism can open up new lines of investigation. Over the past decade, converging lines of research revealed that the common pathogenic mechanism underlying many of neurodegenerative diseases is the aggregation and the deposition of misfolded proteins leading to progressive central nervous system degeneration. This process develops insidiously over the lifetime of an individual, even though they do not manifest clinically until middle or late life. The cause of this prolonged preclinical phase is not completely understood, by it certainly points out the requirement for progressive damages to specific brain regions prior to clinical manifestation of the disease, as well as the unfavorable kinetics of protein misfolding [5]. Cells have adapted sophisticated quality-control mechanisms to protect against the accumulation of misfolded and aggregated proteins. Molecular chaperones promote the correct protein folding and the proteins that remain misfolded are degraded by the ubiquitin-proteasome system. Furthermore, genetic mutations of α-synuclein and parkin cause an abnormal processing of misfolded proteins that overwhelms the quality-control system of the cell. In this context, a computational model of the mechanisms regulating protein

processing can facilitate the development of rationally designed therapies to treat and prevent these disorders. Moreover, this work has showed how modeling and simulation can be used also to reveal similarities and differences between the effects of different causes. This kind of knowledge can be the starting point to project new pharmacological strategy to defeat the disease attacking it from different points and with different methods [11]. A long term goal could be the creation of drugs inspired to the activity of the chaperones [17]. Our model showed, in particular, that a scarce number of chaperones has the same effects of the perturbation introduced by the faulty α-synuclein and parkin, and that there may exist "threshold" phenomena on the rates of toxic chemical absorption triggering the beginning of protein aggregation. Such an information can be used to design the guidelines of new wet-experiments.

Our model does not take into account the possible interactions between parkin and α-synuclein, because at the present they are poorly understood. Therefore, a future direction to extend our work will be the addition of such information. Finally, we remark that the usage of the biochemical π-calculus language in the specification of our model revealed noticeably adapt to describe the concurrent and parallel nature of the reaction involved in the protein processing. Unlike the most common language of differential equation, it offers a new point of view of a biological system by switching from the direct modeling the dynamics of the system component to the modeling of the system components them selves. This new paradigm allows the expression of the compositional and modular nature typical of the biochemical networks.

References

[1] K. B. Beckam and B. N. Ames, *The free radicals theory of aging matures*, Physiol. Rev. 78, 547-581, 1998

[2] R. Bussel Jr. and D. Eliezer, *Residual structure and dynamics in Parkinson's disease-associated mutants of alpha-synyclein*, J. Biol. Chem., 276, 45996-46003, 2001

[3] W. Dauer and S. Przedborski, *Parkinson's Disease: Mechanisms and Models*, Neuron, Vol. 39, pp. 889-909, September 11, 2003

[4] W. A. Eaton, V. Munoz, S. J. Hagen, G. S. Jas, L. J. Lapidus, E. R. Henry and J. Hofrichter, *Fast kinetics and mechanisms in protein folding*, Annu. Rev. Biophys. Biomol. Struct, 29: 327-59, 2000.

[5] M. S. Forman, J. Q. Trojanowsky and V. M-Y Lee, *Neurodegenerative diseases: decade of discoveries paves the way to therapeutic brekthroughs*, Nature Medicine, vol. 10, n. 10, October 2004.

[6] D. T. Gillespie, *Exact stochastic simulation of coupled chemical reactions.*, J. Phys. Chem. 81, pp. 2340-2361, 1977

[7] P. J. Kahle and C. Haass, *How does parkin ligate ubiquitin to Parkinson's disease?*, EMBO report 2004.

[8] C. Kuttler, J. Niehren and R. Blossey, *Gene Regulation in the Pi Calculus: Simulating Cooperativity at the Lambda Switch.* Proceedings of BioConcur 2004.

[9] P. Lecca and C. Priami, *Cell cycle control in eukaryotes: a biospi model*, Technical Report DIT-03-045, University of Trento, 2003.

[10] P. Lecca, C. Priami, P. Quaglia, B. Rossi, C. Laudanna and G. Costantin. *A stochastic process algebra approach to simulation of autoreactive lymphocyte recruitment.* SIMULATION: Transactions of the society for modelling and simulation international , vol. 80, 273-288, 2004.

[11] A. M. Lozano and S. K. Kalia, *Nuovi orizzonti per la cura del Parkinson*, Le Scienze (italian edition of Scientific American), October 2005

[12] R. Milner, *Communicating and Mobile Systems: the π-calculus*, Cambridge University Press, 1999.

[13] A. Phillips and L. Cardelli, *A Correct Abstract Machine for the Stochastic Pi-calculus*, In Procs. of BioConcur 2004.

[14] C. Priami, *Stochastic π-calculus*, The Computer Journal, 38, vol. 6, 578-589, 1995.

[15] A. Regev, W. Silverman, and E. Shapiro, *Representation and simulation of biochemical processes using the pi-calculus process algebra*, In Procs of the Pacific Symposium of Biocomputing (PSB2001), 6: 459-470, 2001.

[16] M. Y. Sherman and A. L. Goldberg, *Cellular defenses against unfolded proteins: a cell biologist thinks about neurodegenerative diseases*, Neuron 29, 15-32, 2001.

[17] D. F. Smith, L. Whitesell, and E. Katsanis, *Molecular chaperones: biology and prospects for pharmacological intervention*, Pharmacological Reviews vo. 50 No. 4, 1998.

[18] M. Vila and S. Przedborski, *Genetic clues to the pathogenesis of Parkinson's disease*, Nature Medicine, July 2004.