Poster I-74

Molecular dynamic and structural analysis from the ab initio model of HC-Pro, a plant virus multifunctional protein

Authors:
Antonio Sergio Kimus Braz (*Depto Genetica - IBB - UNESP Botucatu*)
Sérgio Oyama Jr. (*Centro de Biologia Molecular Estrutural, Laboratório Nacional de Luz Síncrotron*)
Ivan de Godoy Maia (*Depto Genetica - IBB - UNESP Botucatu*)

Short Abstract: The HC-Pro encoded by Potyvirus is a multifunctional protein involved in several steps of the virus life cycle. In this study, we build a three-dimensional model of HC-Pro using an ab initio method, followed by molecular dynamics of two domains derived from this structure.

Long Abstract:
Helper-component proteinase (HC-Pro) is engaged in several steps of the potyviral life cycle, playing several critical functions : autocatalytic protease ; virus transmission by aphid vectors; genome amplification; virus movement; suppressor of RNA silencing. HC-Pro contains three intrinsic functional domains: an N-terminal domain essential for aphid transmission, a central domain that binds nucleic acid in a nonspecific manner and is implicated in a variety of processes, including suppression of RNA silencing and a C-terminal domain that is reminiscent cysteine-peptidases. A number of well-defined essential motifs were identified within the corresponding functional domains of HC-Pro. A conserved KITC motif present in the protein N-terminus, has been shown to promote virus retention in aphid stylets during transmission. In all members of the Potyvirus genus, this motif lies in a highly conserved Cys-rich region predicted to fold in a putative zinc-finger. Notably, mutations in this region affect aphid transmissibility and symptom development. A conserved PTK motif present in the C-terminal half of the protein was also shown to be necessary for aphid transmission, since it is probably involved in the interaction with the viral coat protein. In the central domain that accounts for all additional functions of HC-Pro, a conserved IGN motif plays a critical role in virus genome amplification while a specific CC/SC motif has been implicated in movement. However, these functions are assumed to be an indirect result of the RNA silencing suppression activity of HC-Pro.
It is widely accepted that HC-Pro is functional in solution as a dimer. In biologically active fractions the presence of oligomeric forms of the protein with prevalence of dimers is observed. The N-terminal region was found to be involved in HC-Pro self-association, mutations in conserved His and in two Cys residues within the N-terminal putative zinc-finger domain interfered with dimerization. In the C-terminal region of HC-Pro was also identified as a key factor for dimer formation, a motif located within the C-terminal 128-residues was found to be involved in self-association. The combined results indicate that two distinct contact sites on each HC-Pro monomer might be involved in dimerization.
Despite a wealth of functional data, little is known about the three-dimensional structure of this protein. When searching for a suitable template structures in the Protein Data Bank, we were unable to find any structure sharing sequence similarity to potyvirus HC-Pro. For this

reason it was impossible to use homology and threading methods to generate a three-dimensional structure model for this protein. Instead, an ab initio method for protein structure prediction (HMMSTR/Rosetta) was employed. Since the entire HC-Pro protein is relatively large and has no homologues as a whole, we decided to analyze structural regions smaller than 150 amino acids independently. However, two regions of the protein were highly disordered in our model: the RNA binding domain A, a loop-rich region, and the C-terminal region. Both regions were also shown to be very variable when comparing different models. Therefore, these two portions were not included in subsequent MD simulations. The other two sections, an N-terminal alpha helix rich region (first 90 amino acids) and a central beta sheet rich region (between amino acids 271 and 394), were well structured and relatively constant in all sets of generated models. Curiously, these regions seemed to be capable of keeping their structure in a way that was independent of the neighboring regions. Hence, we focused on the structure of these protein sections and only the corresponding models were submitted to MD simulations (using GROMACS) to test the stability of the modeled conformation. A remarkable attribute of the modeled N-terminal region was the amphipathic properties of the alpha helixes. In this model, all hydrophobic faces of the alpha helixes were turned to the same side, indicating that the predicted structure might not be stable as a monomer. Taking into account that HC-Pro dimer formation is essential for its biological activity, we opted to set up the MD of this section as a dimer in which the hydrophobic areas of each monomer were placed facing each other. The MD results of the N-terminus dimeric model were particularly notable, since a good accommodation of the polypeptide chains was observed. In particular, the formation of small beta sheets in the dimeric monomer-monomer contact area could be detected . The dimeric structure remained stable in a simulation time of 5 ns (starting from 2 ns), thus reflecting the overall stability of the proposed theoretical conformation. Apart from its function in dimer formation, the N-terminus of HC-Pro has been also postulated to bind metal ions and two conserved residues (His-23 and Cys-34) have been implicated in this binding activity. In our model, both these residues are facing the interior of the monomer structure. When the modeled structure was submitted to a search for possible binding atoms and molecules with the SUMO software, a great probability for zinc binding coordinated by residues His-23, Cys-34 and Cys-53 was obtained. Moreover, a Lys residue (Lys-50) present in this putative zinc pocket and implicated in virus transmission by aphids is placed in an exposed and accessible position in the final model.

The sequence stretch corresponding to the second modeled region of HC-Pro comprises amino acid residues 271 to 394. This portion was well structured in the complete model and independently folded. The overall topology of this central domain consists of a series of beta sheets followed by two alpha helixes. After a simulation time of 5 ns the conformation suffered small structural accommodations but remained very close to the initial structure, thus reflecting the overall stability of the proposed model. In the final structure, the functional CSC and PTK motifs were observed with their very exposed residues inside the loops of the beta sheet rich region.