Leveraging *Neurospora crassa* Fungus and Carboxypeptidase A1 Enzyme to Illuminate Microscale Biodiversity Changes in Response to Global Shifts

Supajit Sraphet¹, Chaisri Tharasawatdipipat², Sivapan Choo-in², Pantip Kayee² and Bagher Javadi³*

¹Institute of Molecular Biosciences, Mahidol University, Nakhon Pathom, 73170, Thailand.
²Department of Applied Sciences, Faculty of Science and Technology, Suan Sunandha Rajabhat University, Bangkok, Thailand.
³Department of Sciences, Faculty of Science and Technology, Suan Sunandha Rajabhat University, Bangkok, Thailand.

Abstract

The profound impacts of global changes on biodiversity necessitate a more comprehensive documentation, particularly at the microscale level. To achieve precise and rapid insights into this unique diversity, the choice of an ideal species candidate is crucial. *Neurospora crassa*, a well-established organism in the field of biology, emerges as a promising candidate for this purpose. In our study, we explore the potential of the Carboxypeptidase A1 (CPA1) enzyme as a valuable tool for profiling global diversity. Our investigation has revealed that CPA1 possesses distinctive characteristics, notably its conserved solvent accessibility. This unique feature makes CPA1 an invaluable asset for microscale studies of global changes. The insights presented in our study serve as a practical blueprint, showcasing the application of structural biology in understanding diversity and global changes within microscale environments.

Keywords: *Neurospora crassa*, Protease (Carboxypeptidase A1), Protein Modeling, Computational Analysis, Biodiversity
INTRODUCTION

Genetic diversity regulated with comprehensive complex mechanisms in different living organisms. The diversity of genes and proteins deduced and regulated from environmental conditions. These regulations constructed the different structural biomacromolecules to fulfill living organisms' essentials. Thus, evolution of specific traits should be traceable in the genetic codes of biomacromolecules as well as their structures. The cell and cell growth should adapt with different environmental stresses. Proteases as the important enzymes in the cells, can disclose this diversity.

Most investigation on proteases was based on their needs in the market. Particularly microbial proteases that are used in many industries such as food processing and detergent. Many proteases investigated by researchers, although protease (carboxypeptidase A1, CPA1) in Neurospora crassa gained inevitable and unavoidable position as it is from one easy culturing model fungus (N. crassa) besides its variety of approved functions in the cell. CPA1 involves in many cells function from protein maturation to immune response and reproduction.

Diversity and climate change in global scale documented in many research articles. The changes are in many aspects of the living systems such as species interactions, marines' communities, and populations as well as pest and disease shift besides evolutionary genetic and plasticity. These prospectives gained more attention in global scale. One important question of the many studies should be about the fingerprint changes of rising temperature and the ability of the earth to maintain the current biodiversity of plants and animals especially related to the shifts in the food supplies. However, the results and conclusion are mostly about the organism from different ecologically spread around the planet, thus the specific research on the organisms from smaller scale or even microscale can provide better picture. The research goals should be the effect of solar radiation and temperature on DNA repair, thermotolerance, drought resistance and different stresses on living organisms and their macromolecules. All the effects would have the direct and indirect changes on biomacromolecules specially proteins and enzymes expression and structures. Therefore, finding the microscale environment as well as ideal species and macromolecule/s to study these effects can really help our understanding on climate change and biodiversity.

To address this issue the N. crassa as the model microscopic fungus belong to the Sordariomycetes class was found around the tropical and subtropical regions and isolated from many environments. Thus, it can be good example to study the evolution based on the climate changes. Short life cycle and easiness in culture made it suitable organism for genetic study. Investigation of different biomacromolecules of N. crassa in genes and proteins structures can reveal different environmental stresses such as temperature and light more easily than another organism. Therefore, structural study of N. crassa macromolecules that can represent different environmental diversity can be very revealing, enlightening, and educational. Thus, the objective of this research paper is to scrutinize the protease (Carboxypeptidase A1) structure in N. crassa with the help of computational approaches. This information can help in designing the lab experiment with more meaningful approach. Meanwhile, it would be a good help to find the better link with environmental conditions and diversity of biomacromolecules.

MATERIALS AND METHODS

The protease CPA1 was retrieved and annotated from the full genome of N. crassa in National Center for Biotechnology Information (NCBI). The specific physicochemical features and sequence analysis were done with the help of Swiss institute of bioinformatics-server. The structural model was determined with homology modeling. The final models were evaluated with Ramachandran Map. Structural characterizations such as Secondary structure prediction and solvent accessible surface area were analyzed with Chou & Fasman secondary prediction and Fraczkiewicz and Braun's method, respectively.

Sequence alignment and phylogenetic tree were presented with Clustal Omega Multiple
RESULTS

Molecular weight of the CPA1 is around 45 KDa with neutral pH of isoelectric point (pI) (Table 1). The total of negative and positive charges residues in the structure of the CPA1 is equal. Aliphatic index (AI) and GRAVY showed this molecule is thermostable and slightly hydrophilic. The instability index indicated the stability of CPA1. The average estimated half time before degrading in the cell was around 20 to 30 hours. The extinction coefficient at 280 nm in water estimated to be 78520. The molecule consisted of high Ala, Gly, Ser and Thr residues (Table 2). The least residue is Cys that form two disulfide bonds (Cys 250-Cys 274, Cys 326-Cys 361). There is signal peptide on the n-terminal sequence of CPA1.

The 3d model structure provided with homology modeling with the procarboxypeptidase A (1PCa.1.A) template with more than 36 percent sequence identity and 92% coverage (from res24 to res 422) showed high percentage of helices followed by sheets secondary structure (Figure 1). The CPA1 is monomer with zinc ion. The evaluation of predicated model had QMEAN of -3.06 and Cβ=-2.14 with solvation around the -1.85 (Table 3). The model is highly accepted.

Ramachandran map (Figure 2) indicated that more than 93 percent of residues were in highly acceptable position and only the 1.27% was considered as outsider (Table 3). The model was applied for the refinement with Galaxy Refine tool (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE), however the evaluation of the model didn’t change significantly, showed that homology modeling had significantly predicated the correct position of the CPA1 residues.

Protein-protein interaction showed that the CPA1 had interaction with serine peptidase, alpha-1,3-glucosyltransferase, endopeptidase K and glycoside hydrolase (1,4-galacrutonidase) (Figure 3). The whole genome of the N. crassa Table 1. Physiochemical features of N. crassa. Number of amino acids (NAA), molecular weight (MW), isoelectric point (pI), total number of negatively charged residues (Asp + Glu), total number of positively charged residues (Arg + Lys), aliphatic index (AI), grand average of hydropathicity (GRAVY)

<table>
<thead>
<tr>
<th>Entry</th>
<th>MW (kDa)</th>
<th>pI</th>
<th>negatively charged residues (Asp + Glu)</th>
<th>positively charged residues (Arg + Lys)</th>
<th>AI</th>
<th>GRAVY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q7S312</td>
<td>423</td>
<td>45899.58</td>
<td>7.11</td>
<td>38</td>
<td>38</td>
<td>2066</td>
</tr>
</tbody>
</table>
Table 2. Amino acid compositions of the N. crassa CPA1

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Ala</th>
<th>Arg</th>
<th>Asn</th>
<th>Asp</th>
<th>Cys</th>
<th>Gln</th>
<th>Glu</th>
<th>Gly</th>
<th>His</th>
<th>Ile</th>
<th>Leu</th>
<th>Lys</th>
<th>Met</th>
<th>Phe</th>
<th>Pro</th>
<th>Ser</th>
<th>Thr</th>
<th>Trp</th>
<th>Tyr</th>
<th>Val</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>54</td>
<td>15</td>
<td>21</td>
<td>20</td>
<td>4</td>
<td>12</td>
<td>18</td>
<td>36</td>
<td>9</td>
<td>19</td>
<td>30</td>
<td>23</td>
<td>5</td>
<td>19</td>
<td>32</td>
<td>8</td>
<td>23</td>
<td>32</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>Percentage</td>
<td>12.80%</td>
<td>3.50%</td>
<td>5.00%</td>
<td>4.70%</td>
<td>0.90%</td>
<td>2.80%</td>
<td>4.30%</td>
<td>8.50%</td>
<td>2.10%</td>
<td>4.50%</td>
<td>7.10%</td>
<td>5.40%</td>
<td>1.20%</td>
<td>4.50%</td>
<td>3.80%</td>
<td>7.60%</td>
<td>7.60%</td>
<td>1.90%</td>
<td>5.40%</td>
<td>6.40%</td>
</tr>
</tbody>
</table>

Table 3. N. crassa CPA1 model properties were predicted with homology modeling approach. N. crassa CPA1 information for the 3-D structures of as well as the secondary structure of CPA1 were prompted in the Figure 4 and the CPA1 was in the proximity with NCU07516 gene that is RING finger membrane protein.

Surface and nucleus accessibility showed residues VAL72, ILE73, LEU97, GLY178, ALA182, VAL190, ALA191, ASN224, GLY252, ASP254, ALA286, PRO321, GLY323, THR368, GLY369, ASP373, GLY396, ILE404, GLY408 had zero accessibility to the solvent (Table 4). These residues should be involved in the stability and conformational rigidity of the structure. On the other hand, the residues ARG27, LYS168, ARG341, ASP115, TYR110, LYS49, GLU86, LYS300 and ARG133 had the maximum accessibility to the solvent and considered as the functional residues.

Finding the conserved residues in comparison with other CPAs showed that residues 181H, 183R, 184E, 224N, 226D, 227G, 240K, 257R, 258N, 280G, 373D and 410E were totally conserved beside that the residues 67G, 106G, 168K, 210S, 333N, 394A and 67G had high frequency compared to others. In the other positions all kind of the residues were seen and all of them exposed partially on the surface of the enzyme. It is interesting that conserved residues were all functional residues. Except for conserved residues, the tolerance of the residue exchange was acceptable in other positions of the structure. The maximum frequency was related to the positions 333N and 394A that can help to understand the most alternative residue in the structure. The residue exchange in other parts of the enzyme was observed however the frequency of them are not the same (Figure 1).

**DISCUSSION**

The scientific approach to find the interaction of biological diversity and climate condition is exploring based on the ideal model species and focusing on the diversity of major biological macromolecules in that species. Thus, structural biology of the biomacromolecule can be very suitable and informative specially with the help of computational analysis. Here N. crassa from the Sordariaceae family can be good example for the structural analysis of CPA1. This fungus was long time an ideal model for research in different aspects of molecular biology and...
genetics. Research on circadian rhythm-based physiological regulation\textsuperscript{53}, RNA interference (RNAi) post-transcriptional gene silencing\textsuperscript{54,55}, and DNA methylation-mediated epigenetic control\textsuperscript{56} have been done previously with this model organism. The good source of genetic sequences in public databases such as NCBI and many biological molecular tools and tractability as well as rapid culturing and single gene knock out collection made this fungus more attractive in filamentous ascomycete. The entire genome of \textit{N. crassa} includes the seven chromosomes is available in the public data bases.\textsuperscript{57} This information can help to find the effect of the light and temperature\textsuperscript{58} particularly on CPA1 gene very easily during the 22 hours Lab work.\textsuperscript{59} Even complete growth cycle of \textit{N. crassa} can easily observed and documented in less than one week. Additionally, the effect of the different environment on different life cycle time\textsuperscript{60} as well as gene annotation for specific genes would be easily documented.\textsuperscript{61}

\textbf{Structurally characterization of biomacromolecules provides great insight in defining the evolutionary and biodiversity changes during the time for specific species.\textsuperscript{52-64}} Carboxypeptidases are the hydrolytic enzyme\textsuperscript{65} with the ability to cleave the c-terminal peptide bond of proteins and releasing free amino acids. They have many functional roles in cells such as degradation and modulation of intracellular proteins. They categorizing in three distinct groups including serine, metal and cysteine. Here in the structure of CPA1 the 3dmodel had zinc as the metal in the structure. CPA1 had two disulfide bonds in the structure that were also observed in animal cells polypeptides carboxypeptidases.\textsuperscript{66-71} Generally, peptidase or protease categorized on seven group based on the catalytic residues: serine,
cysteine, threonine, aspartic, glutamic, asparagine and metalloproteases. The mechanism of action to cleave the peptide bond is by making the amino acid residue nucleophilic with catalytic triad. The histidine residue involved in the activation the serine, cysteine, or threonine as a nucleophile.

One important aspect for the investigation on CPA1 of the N. crassa is finding the changes in the active optimal pH. Furthermore, more information of specific zero accessibility to solvent residues identified in our research would help to understand the effect of different environmental stresses on evolution of this structure, additionally post translational modification of them would be informative in the specific environment. On the other hand, the residues with maximum accessibility to solvent could be very good target to find the environmental variability of these amino acids in microscale global changes. Combination of the solvent accessibility and the conserved residues could give the more information about the structural insight. This information (conserved and accessible solvent) in combination would help to find less and more tolerable part of the structure.

Figure 2. Ramachandran plot (A) and ERRAT plot (B) (https://saves v6.0.mbi.ucla.edu/)
Figure 3. Protein-protein interaction of CPA1

Figure 4. Syntenic genes and full genome structure of *N. crassa*. CPA1 indicated in red
to examine for assessment of warming and global changes in macro and microscale. Meanwhile changing the amino acids would change many features of protein that provide great connection with environmental condition.

Reports on computational post translational Carboxypeptidase showed the slight modification of the molecules to performing its function.\textsuperscript{72-76} This post translational included the phosphorylation, O-glycosylation, acetylation in different sites of Serine, Threonine and Tyrosine.

**CONCLUSION**

Computational analysis of CPA1 from \textit{N. crassa}, the model fungus presented. For future studies the Lab isolating the new carboxypeptidase from this species in different climates (such as tropical or subtropical) and finding the SASA and comparing the results would provide better insight into the interaction of environmental stresses on this enzyme and generally biodiversity. Furthermore, finding the effect of light and temperature on CPA1 gene mutation and even in the post-translational modification of the structure would give better insight into global warming and biodiversity.

**ACKNOWLEDGMENTS**

The authors would like to thank the Faculty of Science and Technology, Suan Sunandha Rajabhat University and Mahidol University, Thailand, for their support.

**CONFLICT OF INTEREST**

The authors declare that there is no conflict of interest.

**AUTHORS’ CONTRIBUTION**

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

**FUNDING**

This work was supported by Suan Sunandha Rajabhat University (Grant no. 11773/2566).

**DATA AVAILABILITY**

All datasets generated or analyzed during this study are included in the manuscript.

**ETHICS STATEMENT**

Not applicable.

**REFERENCES**


72. Saleem M, Lamb BC, Nevo E. Inherited differences in crossing over and gene conversion frequencies between wild strains of Sordaria fimicola from


