北京大学定量生物学中心
CENTER FOR QUANTITATIVE BIOLOGY

The 3rd Worldwide Chinese Computational Biology Conference

第三届世界华人计算生物学大会

August 3-6, 2020

Posters

No.	Name	Affiliation	Abstract
1	Zhe Huai	East China Normal University	The difference of the FMO complex from different green sulfur bacteria: A QM/MM description with PPC charge
2	Jie Liu	Wuhan Institute of Technology	Molecular simulations of polymer membranes for organic solvent nanofiltration
3	Shikhar Saxena	Nanyang Technological University	OnionMHC: peptide - HLA-A*02:01 binding prediction using both structure and sequence feature sets
4	Yuanfei Xue	East China Normal University	Accelerated Computation of Free Energy Profile at Ab Initio QM/MM/PIMD Accuracy via Semi-Empirical Reference-Potential
5	Kun Zhang	Qingdao Agricultural University	Investigation on the fungicide resistance mechanism against FgMyoI inhibitor phenamacril by computational study

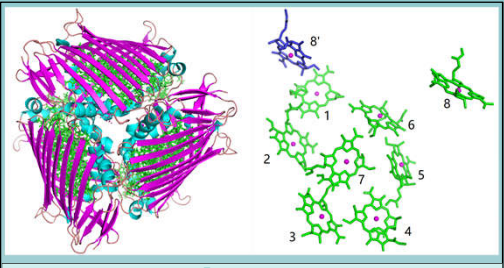
The difference of the FMO complex from different green sulfur bacteria: A QM/MM description with polarized protein-specific charge

Zhe Huai[1], Zhengqing Tong[1], Ye Mei[1,2], Yan Mo[1,2,*]

[1]State Key Laboratory of Precision Spectroscopy and Department of Physics and Institute of Theoretical and Computational Science, East China Normal University , Shanghai, 200062

[2]NYU-ECNU Center for Computational Chemistry at NYU Shanghai,Shanghai,200062

***Email: ymo@phy.ecnu.edu.cn**



Abstract

The relation between structural differences and the difference of absorption spectrums for the two–type FMO complex has not yet understood. To contribute the opinion, we use QM/MM calculations to study the coupling between the protein environment and the pigments. The Polarized Protein–specific Charge (PPC) scheme is used to describe a more realistic protein environment. The spectral densities are calculated to obtain the effect of the protein environment.

Method

Molecular dynamic simulation

- AMBER03 and GAFF force field
- the polarized protein-specific charge (PPC) and Amber charge

The Frenkel exciton model

$$H = \sum_k \varepsilon_k |k\rangle\langle k| + \sum_{k\neq m} V_{km} |k\rangle\langle m|$$

$$V_{km} = f \frac{\mu_k \mu_m}{R_{km}^3} [\vec{\mu}_k \cdot \vec{\mu}_m - 3(\vec{\mu}_k \cdot \vec{n}_{km})(\vec{\mu}_m \cdot \vec{n}_{km})]$$

The Drude bath correlation functions

$$J(\omega) = \frac{2\lambda\gamma\omega}{\omega^2 + \gamma^2}$$

$$C(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \frac{J(\omega) e^{-i\omega t}}{1 - e^{\beta\omega}}$$

$$C(t_i) = \frac{1}{N-i} \sum_{j=1}^{N-i} \Delta E(t_i + t_j) \Delta E(t_j)$$

Results and Discussion

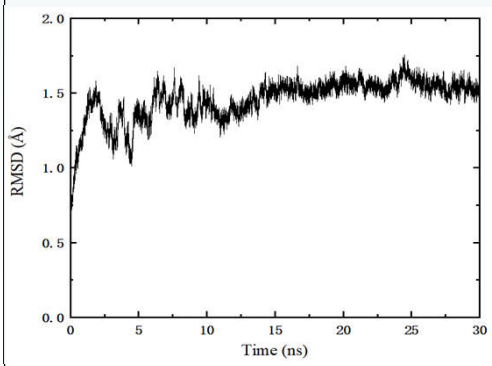


Figure 1. RMSD of the protein backbone for P-type along MD simulations.

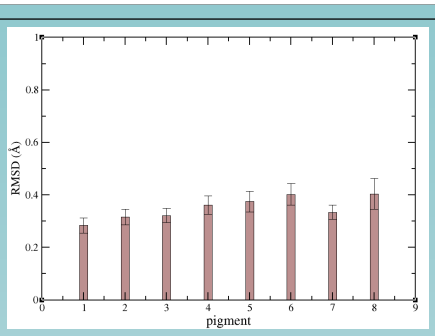


Figure 2. Average RMSDs of the individual pigments for P-type. The standard deviations are indicated with error bars.

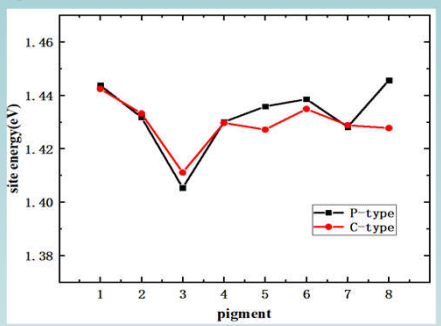


Figure 3. The comparison of site energies between C-type and P-type.

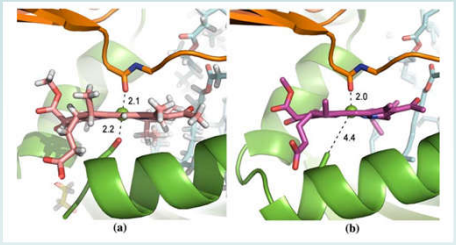


Figure 4. The local environment of the eighth pigment for P-type (a) and C-type (b).

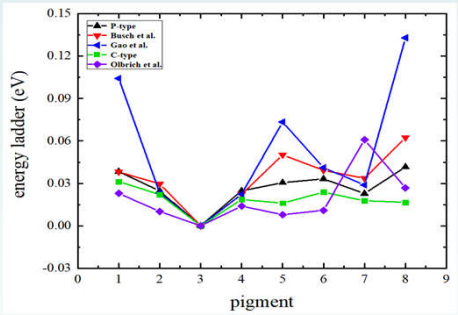


Figure 5. Comparison of site energies obtained in this work with previous studies by Busch et al. (red triangles), Gao et al. (blue circles), and Olbrich et al. (purple squares).

exciton level	exciton energies	BChl 1	BChl 2	BChl 3	BChl 4	BChl 5	BChl 6	BChl 7	BChl 8
1	1.453	0.44	0.26	0.00	0.01	0.11	0.00	0.00	0.17
2	1.451	0.05	0.01	0.02	0.18	0.36	0.36	0.02	0.00
3	1.445	0.26	0.02	0.00	0.00	0.04	0.01	0.00	0.67
4	1.435	0.02	0.00	0.01	0.36	0.03	0.28	0.30	0.00
5	1.429	0.11	0.00	0.00	0.09	0.29	0.30	0.20	0.00
6	1.424	0.11	0.67	0.02	0.02	0.02	0.00	0.00	0.15
7	1.419	0.02	0.00	0.09	0.28	0.14	0.04	0.44	0.00
8	1.402	0.00	0.03	0.86	0.07	0.00	0.00	0.04	0.00

Table 1. Exciton energies (in units of ev) for P-type and the contribution of each pigment to exciton levels.

exciton level	exciton energies	BChl 1	BChl 2	BChl 3	BChl 4	BChl 5	BChl 6	BChl 7	BChl 8
1	1.451	0.67	0.14	0.00	0.01	0.00	0.00	0.15	0.03
2	1.447	0.00	0.02	0.16	0.28	0.13	0.30	0.10	0.02
3	1.433	0.02	0.02	0.00	0.34	0.14	0.36	0.11	0.01
4	1.430	0.02	0.44	0.00	0.00	0.00	0.00	0.01	0.53
5	1.426	0.01	0.10	0.02	0.10	0.13	0.26	0.29	0.10
6	1.424	0.29	0.22	0.02	0.00	0.02	0.00	0.20	0.24
7	1.417	0.00	0.02	0.10	0.27	0.36	0.04	0.17	0.05
8	1.407	0.00	0.03	0.69	0.00	0.21	0.05	0.00	0.02

Table 2. Exciton energies (in units of ev) for C-type and the contribution of each pigment to exciton levels.

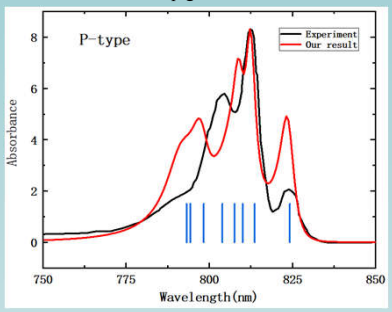


Figure 6. Absorption spectrum of P-type (red line) compared with the experimental spectrum (black line) and exciton energy of each pigment (blue line).

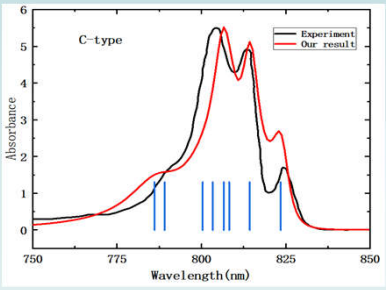


Figure 7. Absorption spectrum of C-type (red line) compared with the experimental spectrum (black line) and exciton energy of each pigment (blue line).

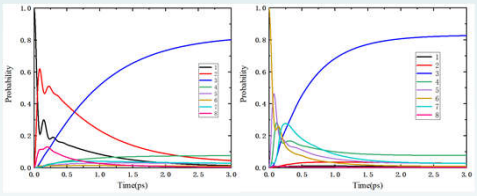


Figure 8. Population dynamics of P-type at 77K when BChla-1 and BChla-6 were initially excited respectively.

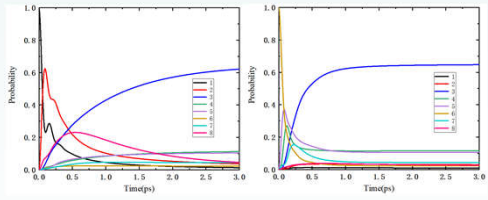


Figure 9. Population dynamics of C-type at 77K when BChla-1 and BChla-6 were initially excited respectively.

References

[1] Jia, X. et al. *Sci. Rep.* **2015**, 5: 17096.

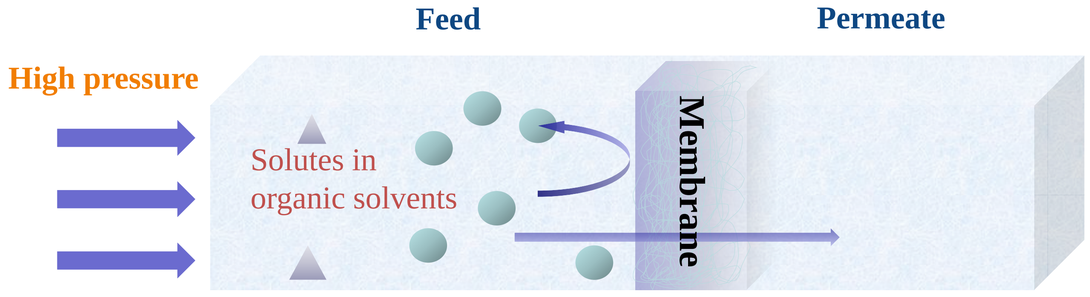
[2]Ji, C., Mei, Y. & Zhang, J. Z. H. *Biophys. J.* **2008**, 95: 108

北京大学定量生物学中心 CENTER FOR QUANTITATIVE BIOLOGY
The 3rd Worldwide Chinese Computational Biology Conference
武汉工程大学 Wuhan Institute of Technology

Molecular simulation of polymer membranes for organic solvent nanofiltration

Liu Jie*
Wuhan Institute of Technology, Email: ljie@wit.edu.com

Background



Organic Solvent Nanofiltration (OSN)

■**Organic solvent nanofiltration**

✓Recover and reuse organic solvents

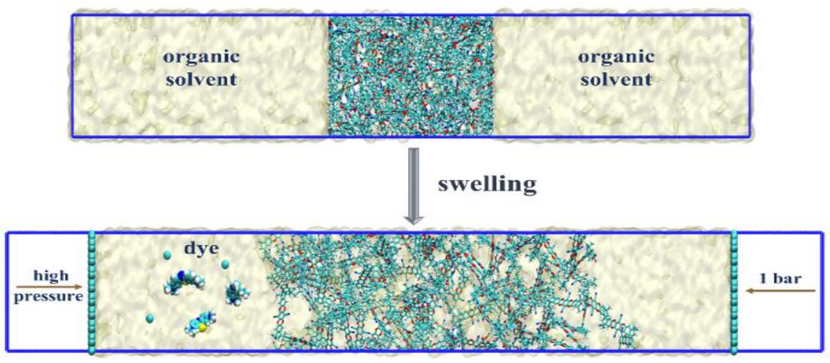
✓Purify and re-concentrate solutes, such as pharmaceutical industry

■**Polymer membrane**

✓Structural diversity

✓Good solvent resistance

✓Low fabrication cost



Swelling is ubiquitous for polymer membranes in a solvent; the membrane structures and properties usually vary substantially after swelling.

New simulation protocol for swelling

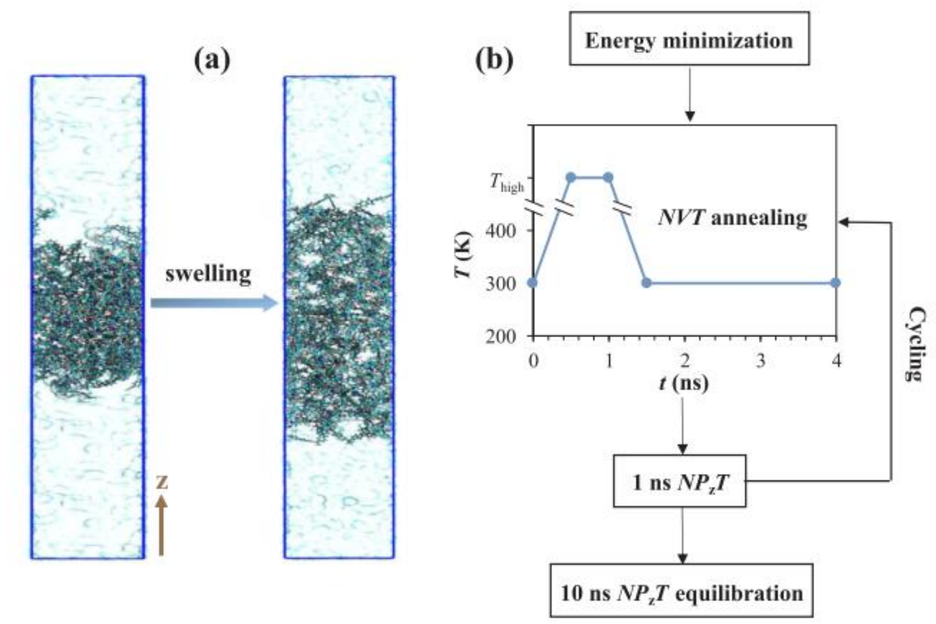
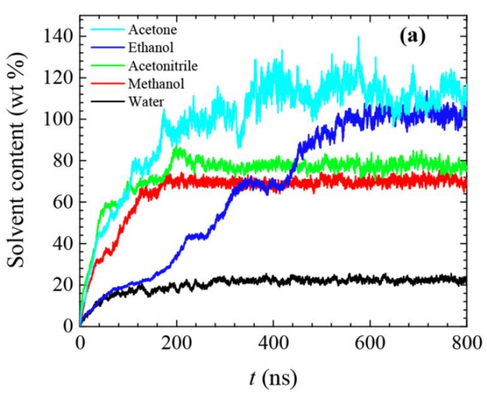
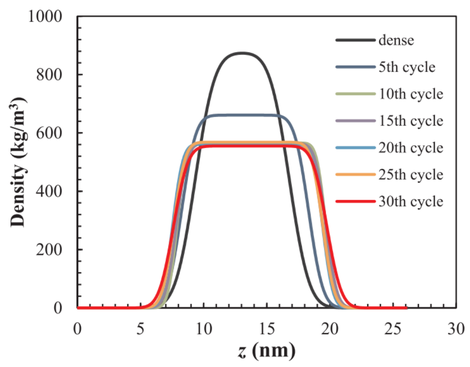


Fig. 1. Membrane swelling. (a) simulation system before and after swelling. (b) new efficient simulation protocol.





The simulation protocol can highly **efficiently** simulate membrane swelling, which is approximately **one order of magnitude faster than normal simulation method**.

Swelling progress

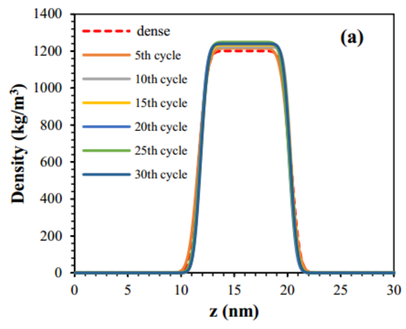


Fig. 4. Fitted density profiles of PI swelling in water started with two different dense membranes

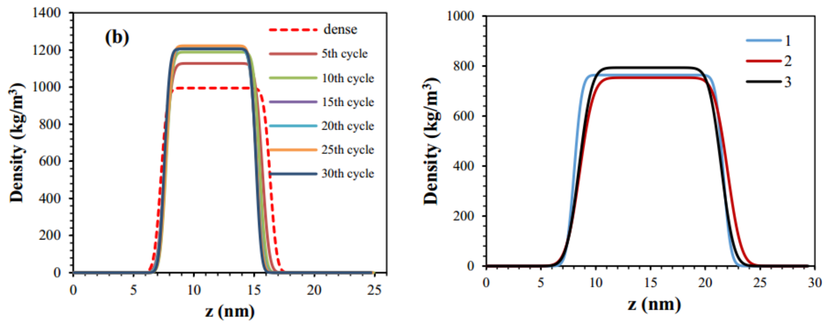
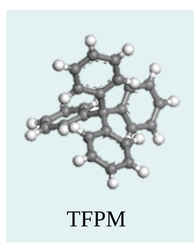


Fig. 5. Fitted density profiles of PI swelling in methanol from three independent runs.

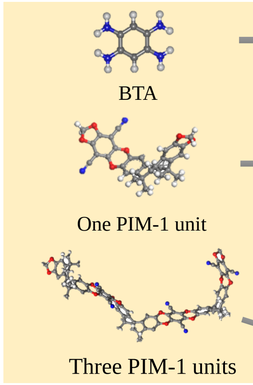
Polymer membranes for OSN

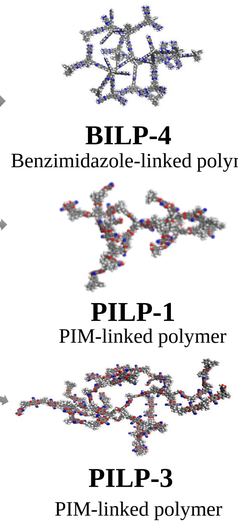
8 Tetrahedral Centers



TFPM (Tetrakis(4-formylphenyl)methane)
BTA (1,2,4,5-Benzenetetramine tetrahydrochloride)
PIM (Polymers of intrinsic microporosity)

16 Linear linkers





BILP-4
Benzimidazole-linked polymer

From experiment
Chem. Mater. 2012, 24, 1511-1517

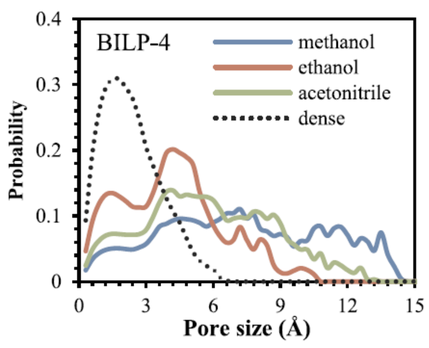
PILP-1
PIM-linked polymer

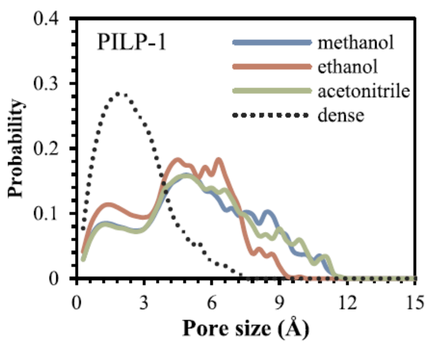
Designed

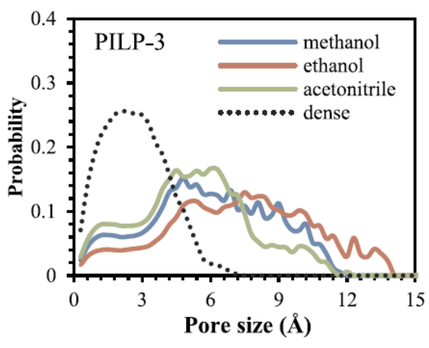
PILP-3
PIM-linked polymer

Designed

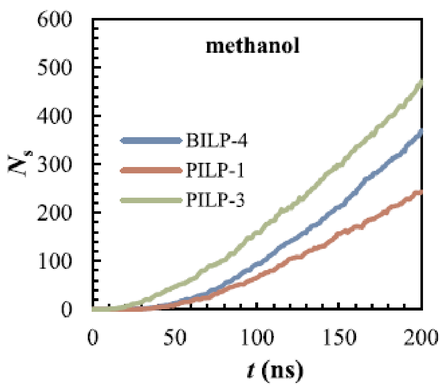
Pore size distribution

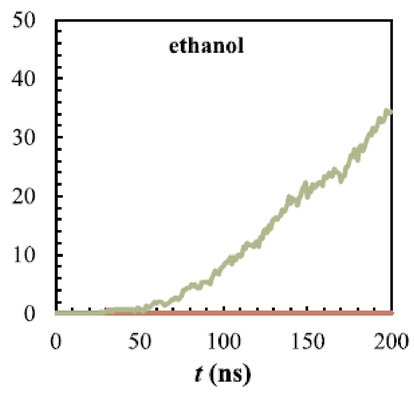


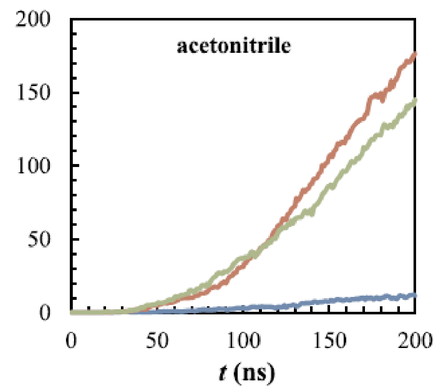




Solvent flows through membranes







Conclusion

- ❑ **A molecular simulation protocol is developed** to examine the swelling of polymer membranes in solvents.
- ❑**Designed two microporous polymer membranes (MPMs)** and predicted their OSN performance.
- ❑The solvent permeation through polymer membrane is governed by the pore size and/or membrane-solvent interaction in a complex manner.
- ❑ For all the three membranes, the rejection of methylene blue is 100%.

Liu J., Jiang J. W., Microporous benzimidazole-linked polymer and its derivatives for organic solvent nanofiltration, ***Polymer*** 2019, 185, 121932.

Liu, J., Xu, Q. S., Jiang, J. W., A molecular simulation protocol for swelling and organic solvent nanofiltration of polymer membranes, ***J. Mem. Sci*** 2019, 573, 639-646.

OnionMHC: peptide - HLA-A*02:01 binding prediction using both structure and sequence feature sets

Shikhar Saxena, Yuguang Mu

School of Biological Sciences, NTU, 60 Nanyang Drive, Singapore - 637551

Abstract:

The peptide binding to Major Histocompatibility Complex (MHC) proteins is an important step in the antigen-presentation pathway. Thus, understanding the binding potential of peptides with MHC is essential for the design of peptide-based therapeutics. Most of the available machine learning based models predict the peptide-MHC binding based on the sequence of amino acids alone, not characterizing the structural features of the peptide-MHC complex.

Given the importance of structural information in determining the stability of the complex, we have considered both the peptide sequence as well as the complex structural features to predict peptide binding to HLA-A*02:01. We have applied machine learning techniques through the natural language processing (NLP) and convolutional neural network to design a model that outperforms the existing state-of-art models. Our model shows that, information from both sequence and structure domains would result in an enhanced performance in binding prediction compared to information from one domain alone. Our model has achieved the state-of-the-art result in most of the weekly benchmark datasets provided by Immune Epitope Database (IEDB).

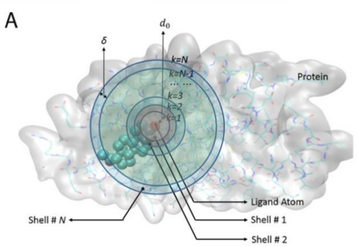
Introductions

Cancer cells are known to produce unique peptide signatures (neoepitopes), some of which are presented on their outer surface and recognized by T cells. For successful presentation of these antigenic peptides, the peptide should be able to bind to the Major Histocompatibility Complex (MHC) receptor with enough strength thus, making it the most crucial and selective step. Thus, there is a need to develop algorithms that predict accurately the binding affinity of peptides to MHC.

In this study, we have used both complex structural feature and sequence feature sets to predict the binding affinity of the peptides to widely studied human HLA-A*02:01 allele. The structural features we have used are the contact pairs between the atoms in peptide and MHC. The contact pairs based structural features are known to perform well for the small molecule and protein complex in the PDBbind dataset. Here, we tested the similar contact-pairs based structural features for peptide-MHC complex. We compared our model's performance with the current state-of-the art baseline model NetMHCpan4.0 and other different models on the independent IEDB benchmark dataset.

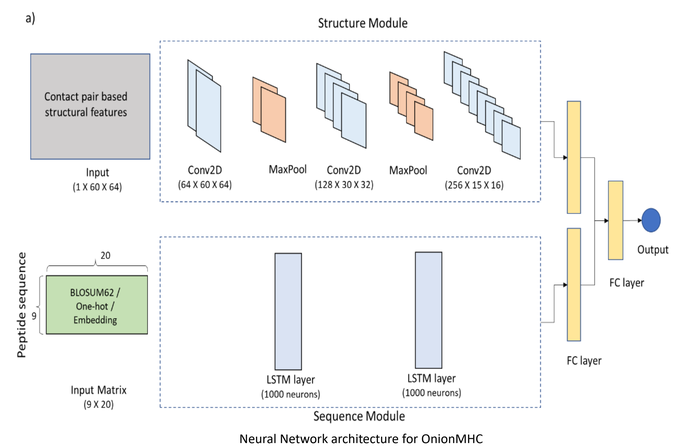
Method:

A



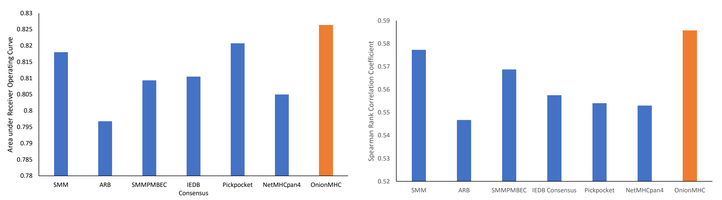
Structure-based features adapted from OnionNet

Adapted from Zheng et. al 2019

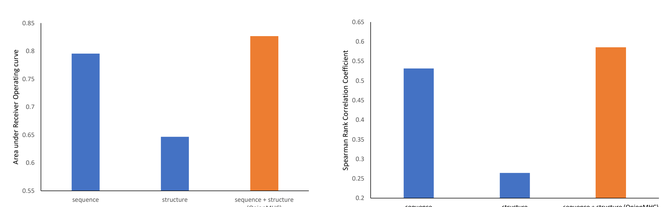


Neural Network architecture for OnionMHC

Results:



Independent Evaluation of OnionMHC and other models on the IEDB benchmark dataset. Performance measured in terms of area under receiver operating curve and spearman rank correlation coefficient.



Model ablation Analysis to understand the contribution of each module towards prediction of binding affinity.

Sequence module is the main contributor but structure module also contains sufficient information required for binding affinity prediction.

Conclusion

Here, a deep learning-based model has been developed to predict the binding affinity of peptide with the HLA-A*02:01 receptor. The model employs both structure as well as sequence feature sets to make binding prediction which is quite different from the previous structure only or sequence only based approaches. Since, in these structure-based features, the whole peptide is treated as a ligand, the residue-wise contribution towards peptide binding is lost. Thus, adding the sequence-based features with the structure-based features would allow the network to learn from the atomic interactions between the peptide-mhc as well as the different peptide residues contributing towards the binding. The combination of both structure and sequence features increases the model's performance compared to either of the features used also as shown in model ablation analysis.

Future Directions:

- The model can be further extended to predict the binding affinities of peptides to different alleles, a pan-alleles model.
- The model can also be extended to accommodate the peptides of different length.

References:

1. Zheng, L.; Fan, J.; Mu, Y., OnionNet: a Multiple-Layer Intermolecular-Contact-Based Convolutional Neural Network for Protein-Ligand Binding Affinity Prediction. *ACS Omega* **2019**, 4, 15956-15965.

2. Trolle, T.; Metushi, I. G.; Greenbaum, J. A.; Kim, Y.; Sidney, J.; Lund, O.; Sette, A.; Peters, B.; Nielsen, M., Automated benchmarking of peptide-MHC class I binding predictions. *Bioinformatics* **2015**, 31, 2174-81.

Acknowledgements:

SG NSCC SINGAPORE National Supercomputing Centre

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Accelerated Computation of Free Energy Profile at Ab Initio QM/MM/PIMD Accuracy via Semi-Empirical Reference-Potential

Yuanfei Xue[1], Jianing Wang[1], and Ye Mei[1,2]

[1]*State Key Laboratory of Precision Spectroscopy, School of Physics and Electronic Science, East China Normal University, Shanghai, 200062, China*

[2]*NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai, 200062*

Introduction

There are increasingly interest of exploring the nuclear quantum effects (NQEs). One of the most popular approaches to achieve this requirement is path integral molecular dynamics (PIMD). However, the computational cost owing to PIMD simulations is tens to hundreds of times more expensive than that of classical molecular dynamics (MD).

The semi-empirical reference-potential (RP) method is proposed to combine with PIMD to accelerate computation. It is proved that the computational cost could be saved two orders at least.

Objectives

To study the NQEs of one system at ab initio QM/MM accuracy with Semi Empirical level cost.

- **Verifying RP at quantum PIMD.**
- **Comparing the computational efficiency.**

Methods

- **Molecular dynamics simulations**
  - **Amber19 package**
  - **PIMD simulations mainly**
  - **Classical MD for comparison**
- **PIMD simulations[1]**
  - $Z(\beta) = \mathrm{Tr}(e^{-\beta\hat{H}/P})^{/P}$
  - $$H_P = \sum_{j=1}^{P} \frac{\mathbf{p}_j^2}{2m} + V(\mathbf{q}_j) + \frac{1}{2} m\omega_P^2 (\mathbf{q}_j - \mathbf{q}_{j+1})^2$$
- **RP method[2]**
  - **Weighted average**

$$\langle \hat{\mathbf{A}} \rangle = -\ln \sum_{n=1}^{N} \omega(\mathbf{r_n}) \hat{\mathbf{A}}(\mathbf{r_n})$$

- **Weight at target level of theory**

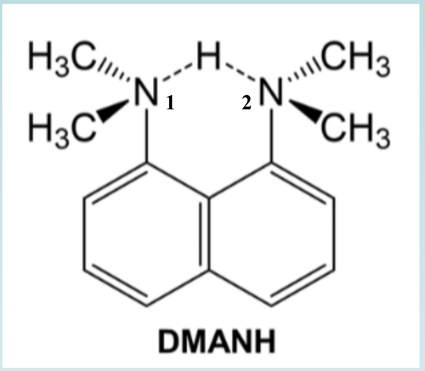
$$\omega_e(\mathbf{r}_n) = \frac{e^{-\beta[\Delta U(\mathbf{r}_n) - f_e]}}{\sum_{k=1}^{K} N_k e^{-\beta[U_k(\mathbf{r}_n) - f_k]}}$$

$$\Delta U = U_H - U_L$$

- **The potential of mean force**

$$F(\mathbf{x}) = -\beta^{-1} \ln \sum_{n=1}^{N} \omega_e(\mathbf{x}) \delta(\mathbf{x}_\eta - \mathbf{x})$$

System Studied



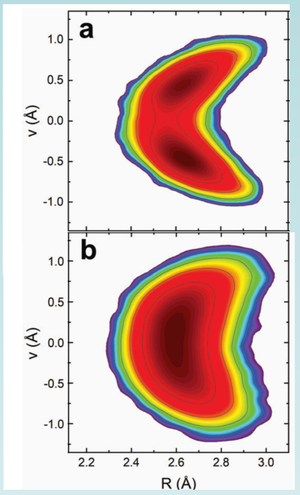


Figure 1. DMANH molecule and free energy profile of (a) ab initio MD and (b) ab initio PIMD simulations in organic solvents[3]. v = $|H\text{-}N_1|$ - $|H\text{-}N_2|$, R= $|N_1\text{-}N_2|$

Results

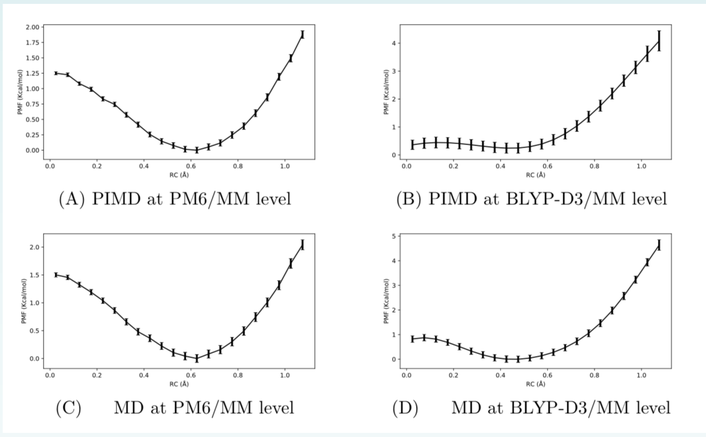


Figure 2.Free energy profile of different simulations of DMANH in aqueous solution at variant Hamiltonian level. RC = $|H\text{-}N_1|$ - $|H\text{-}N_2|$.

- **The efficiency has been enhanced by about 580 folds with Semi-Empirical Reference-Potential.**

References

(1) Feynman, R. P.; Hibbs, A. R. Quantum Mechanics and Path Integrals; McGraw-Hill: New York, 1965

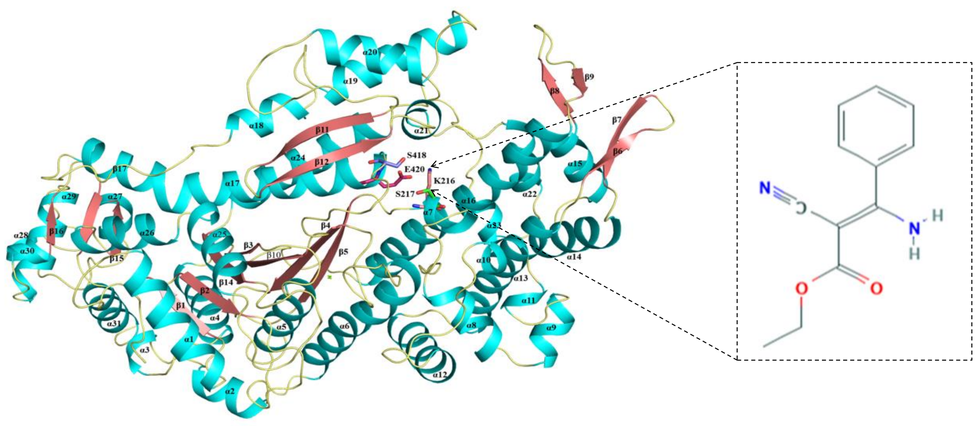
(2) Li, P.; Jia, X.; Pan, X.; Shao, Y.; Mei, Y. Accelerated Computation of Free Energy Profile at ab Initio Quantum Mechanical/Molecular Mechanics Accuracy via a Semi- Empirical Reference Potential. I. Weighted Thermodynamics Perturbation. *J. Chem. Theory Comput*. 2018, 14, 5583–5596.

(3) S. Zhou and L. Wang Symmetry and 1H NMR chemical shifts of short hydrogen bonds: impact of electronic and nuclear quantum effects. *Phys.Chem.Chem.Phys.*, 2020, 22, 4884

Investigation on the fungicide resistance mechanism against FgMyoI inhibitor phenamacril by computational study

Kun Zhang[a], Juan Du[a]*, XiaoJun Yao[b]
[a] Shandong Province Key Laboratory of Applied Mycology, College of Life Science, Qingdao Agricultural University, Qingdao 266109, China
[b] College of Chemistry and Chemical Engineering, Lanzhou University, Lanzhou 730000, China

Abstract



The pathogenic fungus Fusarium graminearum (F.graminearum) causes the fusarium head blight (FHB), which is a global problem for agricultural industry due to its infection strategies and uncontrollable characteristics. The cyanoacrylate fungicide phenamacril is one of the most powerful fungicides for controlling FHB by inhibiting the ATPase activity of the sole class I myosin of only a subset of F.graminearum (FgMyoI). F.graminearum is insensitivity to phenamacril owing to FgMyoI appears single point mutation at K216E, S217P,or E420G. How these mutations affect the interaction mode between FgMyoI and phenamacril is not well understood. In the present study, we investigated the resistance mechanism against phenamacril at atomic level by analyzing the interaction mode between phenamacril and FgMyoI^{WT}, FgMyoI^{K216E}, FgMyoI^{S217P}and FgMyoI^{E420G}, using multiple computational methods, including homology modeling, molecular docking, molecular dynamics simulations, residue interaction network analysis, binding free energy calculation and principle component analysis (PCA). The binding free energy calculation suggests that the binding between FgMyoI and phenamacril is stronger in the wild type (WT) than that in the mutated types (MTs). Further, in comparison with the WT, the interaction mode between phenamacril and residues Lys216, Ser217, Glu420 significantly alters in MTs. Mutations in the protein led to phenamacril resistance in the F.graminearum due to the inefficient binding. In summary, this study provides novel insight to understand the interaction mechanism between FgMyoI and phenamacril and useful information for the rational fungicide design.

Results

Table.1 Summary of the five simulation systems.

System	Starting structure	Water molecules	Total atoms	Time
1	**FgMyoI**	10962	101629	100 ns
2	**FgMyoI^{WT}**	10963	101658	100 ns
3	**FgMyoI^{K216E}**	10964	101651	100 ns
4	**FgMyoI^{S217P}**	10963	101661	100 ns
5	**FgMyoI^{E420G}**	10963	101651	100 ns

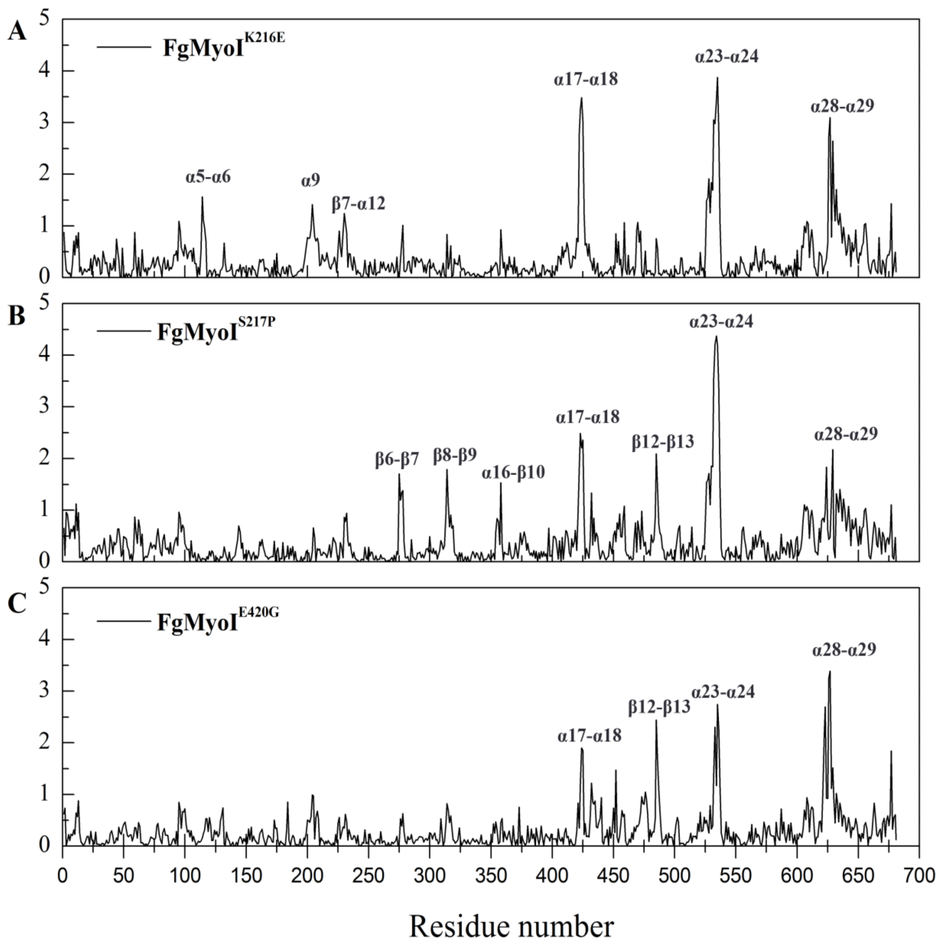


Fig.1 ΔRMSF plots of Ca atoms of three systems.

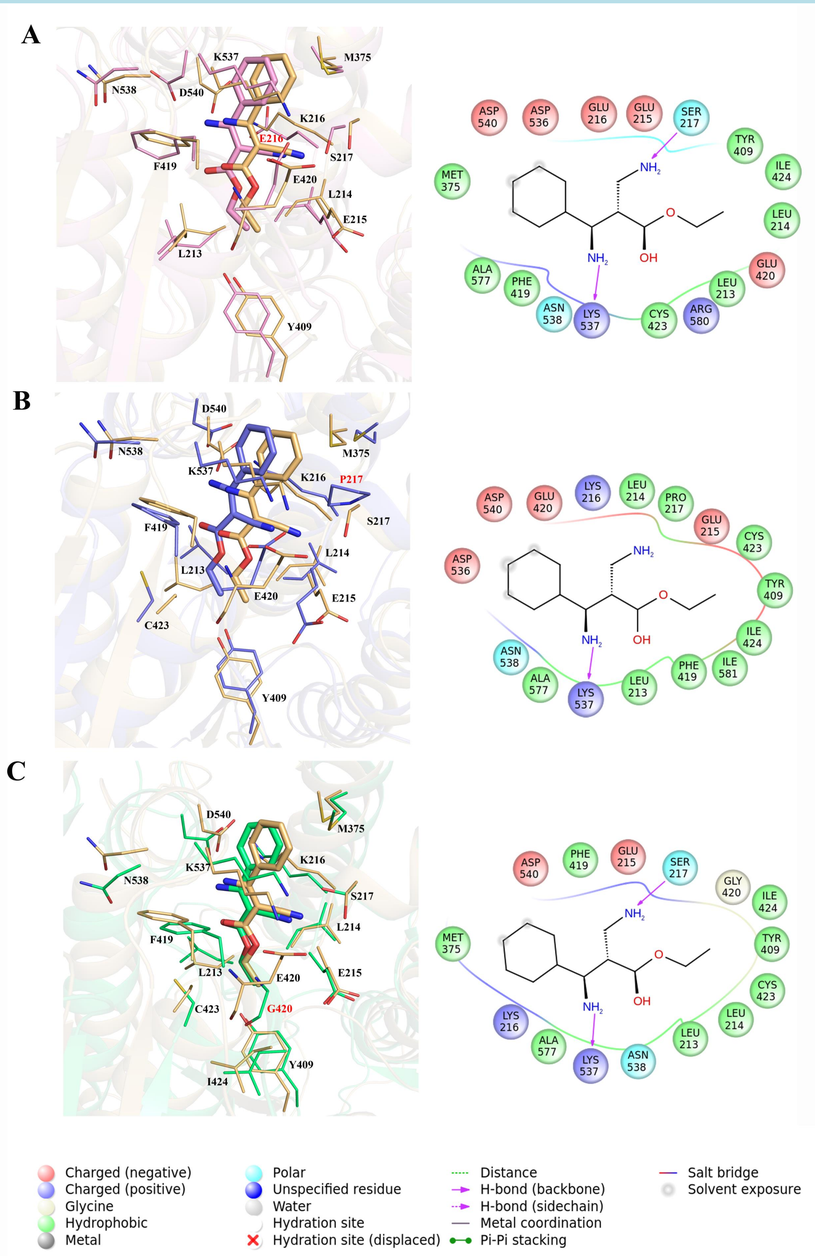


Fig.2 The representative structure of FgMyoI^{K216E}, FgMyoI^{S217P}, FgMyoI^{E420G} and FgMyoI^{WT} system extracted from the last 20 ns of the molecular dynamics simulation.

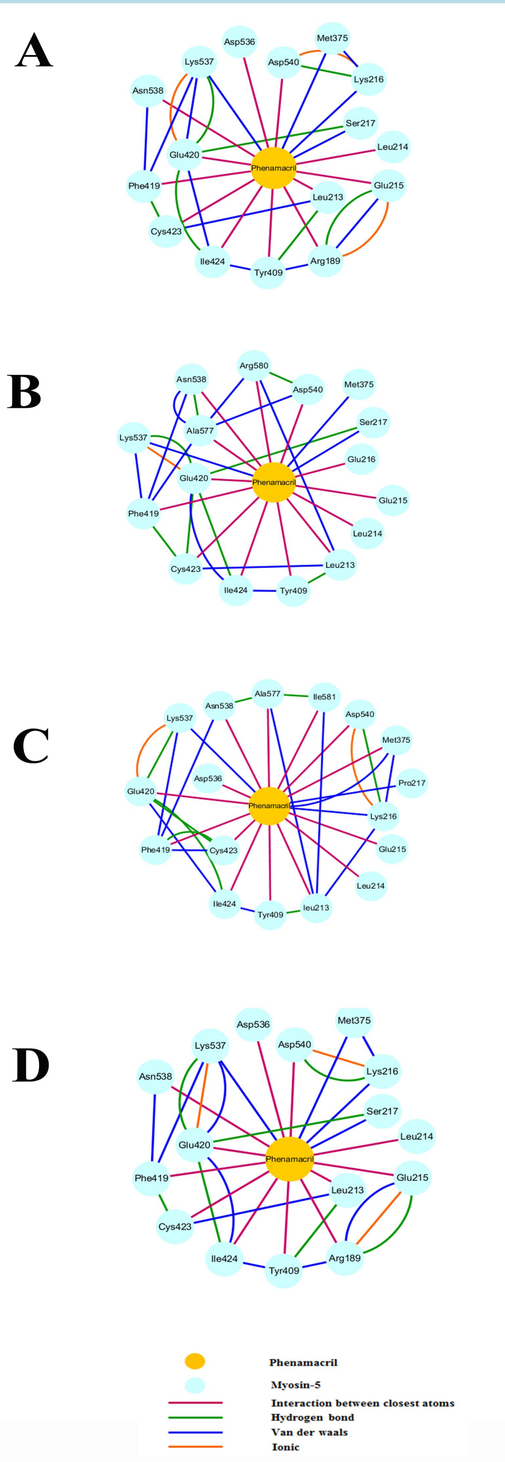


Fig.3 The residues interaction network of Phenamacril binding with b-tubulin in wild type and mutant systems within 4 Å. .

Table.3 Binding free energy for Phenamacril bound to FgMyoI by MM-GB/PBSA methods.

system	$\Delta G_{binding}$		ΔE_{ele}		ΔE_{vdw}		ΔG_{PB-SA}		ΔG_{PB}		ΔG_{GB-SA}		ΔG_{GB}		ΔG_{bind}	
	average	STD	average	STD	average	STD	average	STD	average	STD	average	STD	average	STD	average	STD
FgMyoI^{WT}	-29.20	±0.14	-16.37	±0.18	-36.65	±0.11	-4.29	±0.003	28.11	±0.15	-4.76	±0.007	16.30	±0.11	-36.72	±0.10
FgMyoI^{K216E}	-23.74	±0.15	-11.45	±0.19	-35.44	±0.11	-4.30	±0.003	27.45	±0.16	-4.78	±0.07	12.22	±0.13	-34.67	±0.10
FgMyoI^{S217P}	-26.00	±0.14	-6.90	±0.13	-37.74	±0.11	-4.19	±0.004	22.83	±0.12	-4.91	±0.07	9.43	±0.09	-35.21	±0.10
FgMyoI^{E420G}	-26.42	±0.16	-17.41	±0.22	-37.29	±0.11	-4.29	±0.004	32.57	±0.22	-4.81	±0.007	19.21	±0.17	-35.49	±0.11

Conclusion

1. The residue interaction network analysis, binding free energy calculation indicated that the interaction and binding free energy between Bcb-tubulin and Phenamacril is stronger in the wild type than that in the mutated forms (K216E, S217P and E420G).
2. In summary, the results obtained in this study are benefit to understand the interaction mechanism between FgMyoI and Phenamacril, and provides valuable reference for future structure-based fungicide design.

References

1. Zheng, Z., Hou, Y., Cai, Y., Zhang, Y., Li, Y., and Zhou, M. (2015) Wholegenome sequencing reveals that mutations in myosin-5 confer resistanceto the fungicide phenamacril in Fusarium graminearum. Sci. Rep. 5, 8248 .
2. Wollenberg RD, Taft MH, Giese S, et al. Phenamacril is a reversible and noncompetitive inhibitor of Fusarium class I myosin. J Biol Chem. 2019;294(4):1328-1337. doi:10.1074/jbc.RA118.005408.
3. Zhou Y, Zhou XE, Gong Y, et al. Structural basis of Fusarium myosin I inhibition by phenamacril. PLoS Pathog. 2020;16(3):e1008323. Published 2020 Mar 12. doi:10.1371/journal.ppat.1008323.