

## Exploring the Potency of *Nigella sativa* Seed in Inhibiting SARS-CoV-2 Main Protease Using Molecular Docking and Molecular Dynamics Simulations

Ari Hardianto\*, Muhammad Yusuf, Ika Wiani Hidayat, Safri Ishmayana, and Ukun Mochammad Syukur Soedjanaatmadja

Department of Chemistry, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran,  
Jl. Raya Bandung-Sumedang km 21, Jatinangor 45363, West Java, Indonesia

\* **Corresponding author:**

email: a.hardianto@unpad.ac.id

Received: May 18, 2021

Accepted: July 12, 2021

DOI: 10.22146/ijc.65951

**Abstract:** Coronavirus disease (COVID-19) is a pandemic burdening the global economy. It is caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Black cumin (*Nigella sativa*) seed may contain antivirals for the disease since it was reported to inhibit the human immunodeficiency virus (HIV) and hepatitis C virus (HCV). Main protease ( $M^{pro}$ ) is a vital protein for viral replication and a promising target for COVID-19 drug development. Hence, in this study, we intended to uncover the potency of *N. sativa* seed as the natural source of inhibitors for SARS-CoV-2  $M^{pro}$ . We collected secondary metabolites in *N. sativa* seed through a literature search and employed Lipinski's rule of five as the initial filter. Subsequently, virtual screening campaigns using a molecular docking method were performed, with N3 inhibitor and leupeptin as reference ligands. The top hits were analyzed further using a molecular dynamics simulation approach. Molecular dynamics simulations showed that binding affinities of nigellamine A2 and A3 to  $M^{pro}$  are comparable to that of leupeptin, with median values of  $-43.9$  and  $-36.2$  kcal mol<sup>-1</sup>, respectively. Ultimately, this study provides scientific information regarding *N. sativa* seeds' potency against COVID-19 and helps direct further wet experiments.

**Keywords:** antivirals;  $M^{pro}$ ; *Nigella sativa* seeds; secondary metabolites; virtual screening

### ■ INTRODUCTION

On December 31, 2019, the China World Health Organization (WHO) office reported an outbreak of pneumonia-like cases of unexplained etiology in Wuhan [1]. Later on, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) was identified as the causing agent of such a disease, which are then known as Coronavirus Disease (COVID-19) [2]. The disease quickly spreads around the globe, and on March 11, 2020, WHO announced the pandemic of COVID-19 [3]. To slow the massive spread of COVID-19 many countries implement social distancing, work from home, and lockdown. Nowadays, vaccines have been discovered and administered to many people around the world [4]. However, the COVID-19 pandemic is still growing in many countries and burdens the global economy. By May

16, 2021, COVID-19 has reached 163,188,527 cases and caused 3,383,853 deaths worldwide [5].

Antivirals to cure COVID-19 are an urgent need. One promising inhibition target for SARS-CoV-2 antiviral is the main protease, hereafter called  $M^{pro}$ . It plays a crucial role in cleaving eleven sites on two viral polyproteins (pp1a and pp1ab). The cleavage produces smaller functional viral proteins. Therefore,  $M^{pro}$  inhibition will prevent the formation of virions. Furthermore, since the human proteases do not share the same cleavage specificity with  $M^{pro}$ , inhibiting the viral enzyme would be safe for humans [6].

Black cumin (*Nigella sativa*) seed has long been applied as a folk medicine in Arabian countries, Africa, Europe, and Far East Asia. According to the Prophet Muhammad (PBUH), *N. sativa* seed has curative properties for all illnesses. Moreover, the Bible addresses

the healing properties of *N. sativa* seed [7]. This herbal also has been investigated intensively for its bioactivities, such as anti-inflammatory, antioxidant, anticancer, antifungal, antibacterial, and antiviral [8]. The antiviral properties of the *N. sativa* seed have been reported against cytomegalovirus, hepatitis C virus (HCV), and human immunodeficiency virus (HIV) [7]. Interestingly, in COVID-19 patients, the mixture of honey and *N. sativa* seed substantially improved viral clearance and symptoms, leading to reduced mortality [9]. Therefore, *N. sativa* seed may possess antiviral property against SARS-CoV-2, targeting the M<sup>Pro</sup>.

Many secondary metabolites in *N. sativa* seed have been elucidated and published in peer-reviewed journals [7-8,10-15]. The data benefits *in silico* approach as a rapid method to investigate *N. sativa* seed's potency in inhibiting M<sup>Pro</sup>. *In silico* approach allows us to screen the secondary metabolites against M<sup>Pro</sup> without synthesizing the compounds [16]. Several *in silico* investigations have been reported regarding the potency of *N. sativa* seed as the natural source of COVID-19 antivirals [17-21]. Khan and his co-workers [17] performed a molecular docking study on seven secondary metabolites in *N. Sativa* seed. These are stigmasterol glucoside,  $\alpha$ -hederin, nigellidine, nigellidine-4-O-sulfite, sterol-3- $\beta$ -D-glucoside,  $\beta$ -sitosterol, and dithymoquinone. Their study suggested that  $\alpha$ -hederin is the strongest ligand binding to M<sup>Pro</sup>. Similarly, Bouchentouf and Missoum [18] proposed  $\alpha$ -hederin and nigellidine as the best potential inhibitors for M<sup>Pro</sup> from their molecular docking study of nine secondary metabolites in *N. sativa* seed. In a separate molecular docking study of four secondary metabolites to M<sup>Pro</sup>, Maiti and colleagues [19] also suggested nigellidine as the most potent inhibitor. Based on their molecular docking study of three secondary metabolites in *N. sativa* seed, Sumaryada and Pramudita also proposed nigellidine as the best inhibitor against M<sup>Pro</sup> [20]. Meanwhile, using a more extensive data set of 24 secondary metabolites in their molecular docking study, Ferdian and his co-workers found 3-[(4-methylphenyl)sulfanyl]-1,3-diphenyl-1-propanone as the best potential inhibitor against M<sup>Pro</sup> [21].

The molecular docking approach has some limitations [22]. This computational method applies a

rigid receptor and a ligand with fixed bond lengths and angles to decrease conformational space. Additionally, molecular docking simplifies a scoring function for rapid pose evaluation. Hence, another method, such as molecular dynamics (MD) simulation, should be implemented after the molecular docking step to obtain a more accurate prediction of energy and conformation. MD simulations have been employed in many studies of protein-ligand interactions [23-25].

In this present study, we employed a larger data set of secondary metabolites than the previous studies [16-20] to discover the potency of *N. sativa* seed as the source of M<sup>Pro</sup> inhibitors. Initially, Lipinski's rule of five [26] was implemented to filter the secondary metabolites as oral drugs. Subsequently, we virtually screen the secondary metabolites through molecular docking using Autodock Vina [22]. The best hits were further subjected to molecular dynamics simulation to obtain more realistic binding affinity scores and more understanding regarding the structural dynamics of M<sup>Pro</sup> after ligand binding. This study is valuable to reveal *N. sativa* seeds' potency against COVID-19 and guide further wet laboratory experiments.

## ■ EXPERIMENTAL SECTION

### Secondary Metabolite Structure Preparation

Secondary metabolites in *N. sativa* seeds were collected from literature searches using PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) and Google Scholar (<https://scholar.google.com/>). The 2D structures of secondary metabolites were manually drawn using BIOVIA Draw Version 19.1 NET 64 bit (<https://discover.3ds.com/biovia-draw-academic>) or retrieved from PubChem (<https://pubchem.ncbi.nlm.nih.gov>). The predictions of Ro5 were performed on the webserver of SwissADME (<http://www.swissadme.ch>). Protonation states of the secondary metabolites at physiological pH (7.4) were predicted using MarvinView 18.21.0 (<https://chemaxon.com/products/marvin>). The conversion of 2D to 3D structures was performed using BIOVIA Discovery Studio 2019 Visualizer (<https://www.3dsbiovia.com>). Subsequently, the 3D structures were subjected to a geometry

optimization stage using a semi-empirical method of Parameter Model 6 (PM6) in Gaussian09 [27].

### Molecular Docking and Virtual Screening Campaigns

The 3D structures of M<sup>Pro</sup> (PDB ID 6LU7) were prepared BIOVIA Discovery Studio 2019 Visualizer by separating the respective ligands and discarding ions and water molecules. Additionally, a DockPrep module in Chimera 1.14rc (<https://www.cgl.ucsf.edu/chimera/>) was employed to remove alternate conformers in the 3D structure of M<sup>Pro</sup>. The apo form of M<sup>Pro</sup> and the ligand, N3, were saved as separated pdb files. An additional M<sup>Pro</sup> ligand, leupeptin, was prepared from its crystal structure (6XCH) and utilized as the second reference. All pdb files, including secondary metabolites, were converted into pdbqt using AutoDockTools-1.5.6 (<http://autodock.scripps.edu/resources/adt>). Autodock Vina was utilized to perform molecular docking since its default method was well tested and highly optimized [22]. Molecular docking validation was conducted by redocking the crystal structures of M<sup>Pro</sup>-N3 (PDB ID 6LU7) and -leupeptin (6XCH). For the virtual screening step, PaDEL-ADV (<http://padel.nus.edu.sg/software/padeladv/index.html>) was employed for automatization. All molecular docking experiments conducted in this study used the default method of Autodock Vina [22]. Firstly, the global search exhaustiveness was eight. Secondly, the maximum number of docking poses was nine. Thirdly, the best and worst binding modes were set to have an energy difference up to 3 kcal mol<sup>-1</sup>.

### Molecular Dynamics Simulation

To identify potential inhibitors of M<sup>Pro</sup>, previous studies [17-21] used the molecular docking approach, which might be insufficient since the method has some limitations [22]. Therefore, the current study employed MD for such a purpose. MD simulations were performed on Ubuntu 20.04.2.0 LTS computer with Intel Xeon® CPU E5-2678 v3 @2.5 GHz × 24, GPU NVIDIA GeForce RTX 2080Ti 6 GB, and 16 GB of RAM. The MD simulation procedure was adapted from our previous research [25]. Briefly, we calculated partial charges of secondary metabolite molecules by using the Austin Model 1 - Bond

Charge Corrections (AM1-BCC) protocol in the *antechamber* program (AmberTools20) [28] (<https://ambermd.org/AmberTools.php>). The other parameters for the secondary metabolites were derived from Generalized Amber Force Fields 2 (GAFF2) [28]. For all MD simulations, ff14SB [29] was assigned to amino acid residues of M<sup>Pro</sup>. Every M<sup>Pro</sup> and secondary metabolite complex system was prepared using the *tleap* program (AmberTools20). In the MD preparation step, we used the explicit water model of SPC/E to solvate each complex of the M<sup>Pro</sup>-ligand complex. We set the boundary box to 10 Å. Using the *tleap* program in AmberTools20, a few Na<sup>+</sup> and Cl<sup>-</sup> ions were added to accomplish a physiological salt concentration of 0.15 M.

We employed GPU-accelerated Particle-Mesh Ewald Molecular Dynamics (PMEMD) and periodic boundary conditions, as implemented in Amber20 [28], for each protein-ligand complex. Initially, we performed two sequential steps of energy minimization. In the first step, the protein-ligand complex was restrained by 25 kcal mol<sup>-1</sup> Å<sup>-2</sup>. In the following step, a restraint of 5 kcal mol<sup>-1</sup> Å<sup>-2</sup> was applied. Under a 50-ps NVT condition, the system temperature was raised to 300 K. The system was switched to an NPT simulation, and the density was adjusted to 1 g cm<sup>-3</sup> over 50 ps. In the following simulations of the NVT condition, the solute's restraint was gradually reduced every 50 ps by 1 kcal mol<sup>-1</sup> Å<sup>-2</sup> until it was removed entirely.

To acquire a 100-ns MD trajectory for every system, we simulated each system at 300 K under the NPT condition. Particle-mesh Ewald (PME) method was employed to treat long-range electrostatic interactions, whereas, for short-range non-bonded interactions, we used a 10 Å cut-off. By using a SHAKE algorithm, we constrained all bonds involving hydrogen atoms. Constant temperature and pressure of each system were maintained by Langevin thermostat and Berendsen barostat, respectively.

### Binding Energy Calculation

We utilized the Molecular Mechanics Generalized Born Surface Area (MMGBSA) method, which is implemented in MMPBSA.py (AmberTools20), to

estimate binding energy values between ligands and  $M^{pro}$ . The MMGBSA binding free energy ( $\Delta G_{MMGBSA}^{\circ}$ ) is calculated as follows in Eq. (1):

$$\Delta G_{MMGBSA}^{\circ} = \langle G_{com} \rangle_i - \langle G_{rec} \rangle_i - \langle G_{lig} \rangle_i \quad (1)$$

where  $\langle G_{com} \rangle_i$ ,  $\langle G_{rec} \rangle_i$ , and  $\langle G_{lig} \rangle_i$  denote the average energy value for complex, enzyme, and ligand, respectively. The breakdown of every  $G_x$  is as follow in Eq. (2):

$$G_x = E_{MM} + G_{solv}^{GB} + G_{solv}^{SA} \quad (2)$$

where  $E_{MM}$  is the gas phase energy,  $G_{solv}^{GB}$  is the electrostatic contribution to the solvation energy determined using the Generalized Born (GB) implicit solvent model, and  $G_{solv}^{SA}$  is the hydrophobic portion of the solvation energy. The hydrophobic portion is computed by the Linear Combination of Pairwise Overlaps (LCPO) method.  $E_{MM}$  consists of internal energy terms, van der Waals ( $E_{vdw}$ ), and electrostatic interactions ( $E_{el}$ ). The internal energy terms constitute angle ( $E_{angle}$ ), bond ( $E_{bond}$ ), and torsion energies ( $E_{torsion}$ ). Since we single-trajectory MD simulations, internal energy terms were excluded [25].

## MD Trajectory Analysis

We utilized *cptraj* program in AmberTools20 to analyze MD trajectories. The analysis includes computation of Root-Mean-Square Fluctuation (RMSF), Root-Mean-Square Deviation (RMSD), and H-bond conservation. All plots were produced by using a *plotnine* Python library (<https://plotnine.readthedocs.io/en/stable/#>). Statistical significance tests were computed by using R (<https://cran.r-project.org/>) and RStudio (<https://www.rstudio.com/>).

## RESULTS AND DISCUSSION

### Data Collection and Lipinski's Rule of Five

Several 67 secondary metabolites in *N. sativa* seed were curated from the literature [7,8,10–13] and tabulated in Table S1. This number of secondary metabolites is larger than the similar works previously reported [17–21]. The secondary metabolites consist of seventeen monoterpenoids, one diterpenoid, two sesquiterpenoids, six phenylpropanoids, three vitamin E compounds, ten phytosterols, eight triterpenoids, three saponins, three flavonols, and fourteen alkaloids. The two-dimensional

(2D) structures of these secondary metabolites were retrieved from PubChem or sketched using BIOVIA Draw Version 19.1 NET 64 bit.

As the initial screening, we used the SwissADME web server [30] to apply Lipinski's rule of five (Ro5) to all secondary metabolites. Ro5 helped to define the probability of the secondary metabolites as oral drugs through simple physicochemical properties [26]. The first rule is that molecular weight is not greater than 500 g mol<sup>-1</sup> (MW ≤ 500). The second rule is the calculated LogP not greater than 5 (CLogP ≤ 5) or Moriguchi LogP not over than 4.15 (MLogP ≤ 4.15). The third rule is that the number of hydrogen-bond donors is not greater than 5 (H-bond donors ≤ 5). The fourth rule is that the number of H-bond acceptors is not greater than 10 (H-bond acceptors ≤ 10). Of 67 secondary metabolites, eight compounds were excluded since they violated more than one criteria of Ro5 (Table S2) [26]. They are nigellamine A1, nigellamine B1, nigellamine B2,  $\alpha$ -hederin, 3-O- $[\beta$ -D-xylopyranosyl-(1→2)- $\alpha$ -L-rhamnopyranosyl-(1→2)- $\beta$ -D-glucopyranosyl]-11-methoxy-16,23-dihydroxy-28-methyl olean-12-enoate, quercetin 3-(6'''-feruloylglucosyl)-(1→2)-galactosyl-(1→2)-glucoside, quercetin 3-glucosyl-(1→2)-galactosyl-(1→2)-glucoside, and kaempferol 3-glucosyl-(1→2)-galactosyl-(1→2)-glucoside. Interestingly,  $\alpha$ -hederin, which is suggested by some studies [17–18] as the potential secondary metabolite, violates three rules of Ro5 (Table S2). It has an MW of 750.96 g mol<sup>-1</sup>, seven H-bond donors, and 12 H-bond acceptors. Therefore,  $\alpha$ -hederin may not be active as an oral drug.

### Secondary metabolite structure preparation

The protonation state of a chemical structure, which is contributed by any titratable functional group, depends on environmental pH [24]. Meanwhile, the correct assignment of protonation state has a crucial role in producing meaningful studies of molecular docking [31] and dynamics simulation [24]. Thus, we predicted charge states of secondary metabolites in *N. sativa* seeds at physiological pH (7.4) using MarvinView 18.21.0. This program successfully assisted our previous study [24] in assigning correct protonation states of chemical structures. In the present study, the program generated the predictions of major microspecies (data not shown)

and charge states of all secondary metabolites (Table S1). We assigned hydrogen atoms carefully during the 2D to 3D conversions of the secondary metabolite structures based on these prediction data.

The geometry optimization of ligand structures is an essential step in molecular docking since it significantly improves the binding energy values of the resulting poses [32]. Hence before conducting molecular docking, we optimized the geometry of all 3D structures using a semi-empirical method of Parameter Model 6 (PM6) in Gaussian09. This semi-empirical approach is accurate in reproducing geometries of simple organic crystals and small molecules [33]. Representative secondary metabolite structures optimized by the use of the PM6 method are depicted in Fig. S1.

### Virtual screening campaigns

Employing the molecular docking approach as a virtual screening tool, we continued to explore *N. sativa* seeds' potency in inhibiting M<sup>pro</sup>. We used the crystal structures with PDB ID 6LU7, which contains N3 inhibitor and M<sup>pro</sup> [34]. Beforehand, the redocking procedure was conducted on the complex of M<sup>pro</sup> and N3 as a validation step. The redocking procedure ensures the stochastic search method of molecular docking finding a global minimum [22]. The redocking result with a low RMSD of 1.50 Å (Fig. S2) was achieved by using a grid box size of 20 × 28 × 18 and spacing 1 Å. Such a redocking result suggested that N3 binds to M<sup>pro</sup> with a -9.1 kcal/mol binding energy score.

Additionally, we used leupeptin as another reference ligand, where its complex with M<sup>pro</sup> is available as a crystal structure with PDB ID 6XCH. Leupeptin is a natural

protease inhibitor that has inhibition activities to the main proteases of SARS-CoV [35] and SARS-CoV-2 [6]. Redocking of leupeptin to M<sup>pro</sup> gave the best hit with an RMSD of 1.47 Å (Fig. S2) and a binding energy score of 7.6 kcal mol<sup>-1</sup>. These redocking results reflect the agreement between *in silico* and experiment studies, where the inhibition activity of N3 [34] is higher than that of leupeptin [35]. The use of N3 and leupeptin as the reference ligands allows a depiction of binding affinity levels of secondary metabolites in *N. sativa* seeds against M<sup>pro</sup>.

In the subsequent step, the 59 secondary metabolites in *N. sativa* seeds were subjected to virtual screening against M<sup>pro</sup>. The results suggested that all secondary metabolites bind to M<sup>pro</sup> weaker than N3 (Table S3). Nevertheless, six secondary metabolites exhibit binding energy scores stronger than that of leupeptin (Table 1). They are nigellidine-4-O-sulfite, taraxerol, nigellidine, nigellamine A2, nigellamine A3, and melanthigenin. Therefore, these six secondary metabolites are considered as the potential ligands inhibiting M<sup>pro</sup>.

Most molecular docking methods, such as using Autodock Vina, implement fixed bond angles and a rigid receptor to reduce conformational space. Moreover, at every conformational search, the methods use a simplified scoring function for rapid energy evaluation. These limitations require further refinement methods, like molecular dynamics simulation, to achieve more realistic energy prediction and conformational search [22]. Hence, we subjected the seven secondary metabolites bound to M<sup>pro</sup>, as listed in Table 1, to molecular dynamics simulation.

**Table 1.** Binding energy values of reference ligands and potential M<sup>pro</sup> inhibitor candidates from *N. sativa* seeds

Ligand	Binding Energy (kcal mol <sup>-1</sup> )	Information
N3	-9.1	Reference ligand
Nigellidine-4-O-sulfite	-8.2	Secondary metabolites from <i>N. sativa</i> seeds
Taraxerol	-7.8	Secondary metabolites from <i>N. sativa</i> seeds
Nigellidine	-7.8	Secondary metabolites from <i>N. sativa</i> seeds
Nigellamine A2	-7.7	Secondary metabolites from <i>N. sativa</i> seeds
Nigellamine A3	-7.7	Secondary metabolites from <i>N. sativa</i> seeds
Melanthigenin	-7.7	Secondary metabolites from <i>N. sativa</i> seeds
Leupeptin	-7.6	Reference ligand

\*The binding energy values were resulted of molecular docking procedure using AutoDock Vina

### Binding Affinity Refinement Using Molecular Dynamics Simulation

We performed MD simulations for all complex of ligands (Table 1) bound to  $M^{pro}$  to yield a 100-ns trajectory. Subsequently, we computed  $\Delta G^{\circ}_{MMGBSA}$  every 10 ns and visualized the resulting values from each trajectory, as shown in Fig. S3. Additionally, we calculated the descriptive statistics (Table 2) and created the Box-plot (Fig. 1). The  $\Delta G^{\circ}_{MMGBSA}$  results are considerably different from binding energy scores obtained from molecular docking (Table 1) and, thus, re-rank the ligands' order in terms of binding affinities. While the N3 inhibitor maintains its ranking as the most potent ligand for  $M^{pro}$ , leupeptin exhibits a stronger  $\Delta G^{\circ}_{MMGBSA}$  value than all secondary metabolites, except nigellamine A2. Nigellidine, taraxerol, and nigellidine-4-O-sulfite are far weaker than leupeptin in binding to  $M^{pro}$ . Meanwhile,  $\Delta G^{\circ}_{MMGBSA}$  value of A3 is quite similar to that of leupeptin.

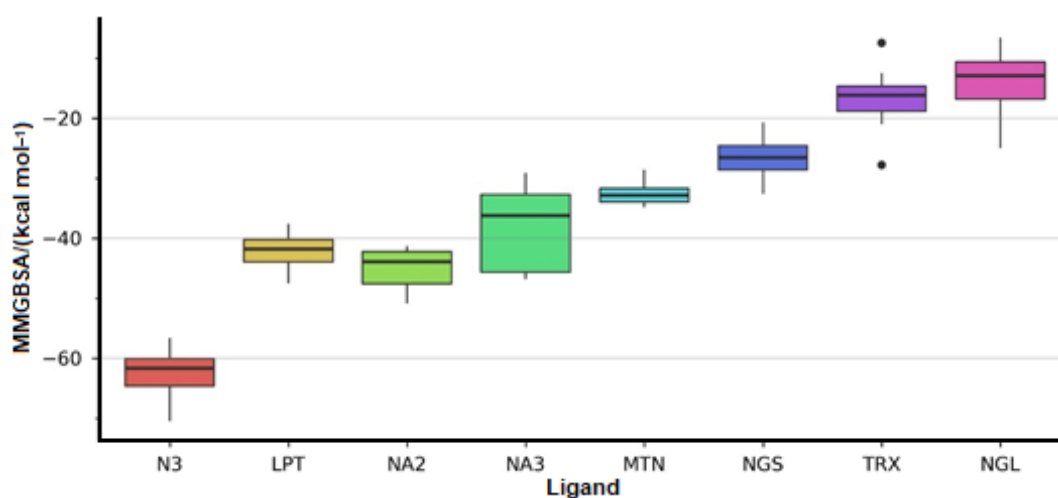
We utilized inferential statistics approaches to evaluate the significant difference among values and of leupeptin, nigellamine A2, nigellamine A3, and melanthigenin. In advance, we checked assumptions for the parametric method, analysis of variance (ANOVA). The data is normally distributed, according to the Shapiro test ( $p$ -value =  $9.34 \times 10^{-2}$ ;  $\alpha = 5\%$ ), but lack variance homogeneity ( $p$ -value =  $2.33 \times 10^{-3}$ ). Thus, we run a non-parametric significance test, Kruskal-Wallis rank-sum

test. The test was followed by Dunn's multiple comparison test with the Bonferroni method as the post hoc test. The test (Supporting Information: Note) suggested that there is no statistical difference among  $\Delta G^{\circ}_{MMGBSA}$  values of leupeptin, nigellamine A2, and nigellamine A3. Therefore, nigellamine A2 and A3 are potential candidates as  $M^{pro}$  inhibitors, particularly the first compound.

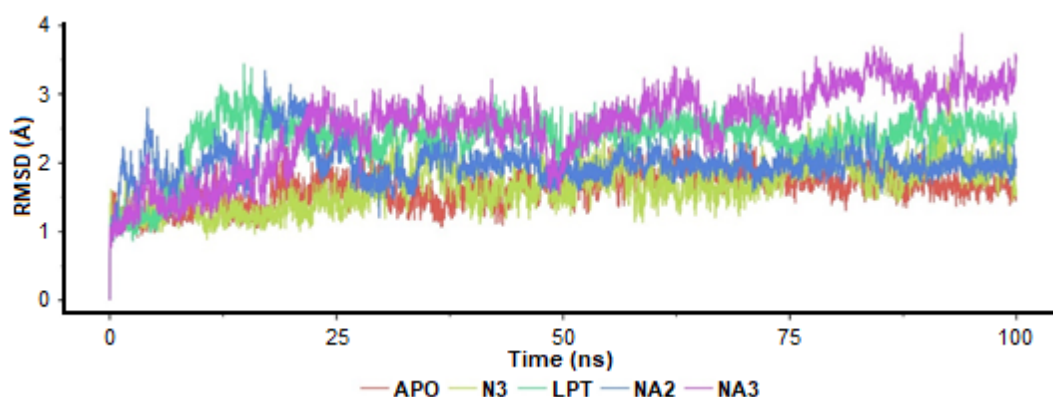
### The Influence of Ligand Binding on the Structural Dynamics of $M^{pro}$

The binding of every ligand to  $M^{pro}$  affects differently on the structural dynamics of the protein. Ligands with weak binding affinities tend to increase conformational changes of  $M^{pro}$  (Fig. S4: top panels). Nigellidine-4-O-sulfite, taraxerol, and nigellidine cause the increase of conformational changes on  $M^{pro}$  structure. The increase also occurs due to the binding of melanthigenin, nigellamine A2, and nigellamine A3 to  $M^{pro}$  (Fig. 2 and Fig. S4). Nevertheless, some ligands, including taraxerol, nigellidine, and melanthigenin, show low conformational changes when binding to  $M^{pro}$  (Fig. S4: lower panels).

As the reference ligand, the binding of leupeptin to  $M^{pro}$  causes a low conformational change on the protein, particularly after 25 ns (Fig. 2). Such binding also shows a relatively stable trajectory of  $\Delta G^{\circ}_{MMGBSA}$  values, around



**Fig 1.** Box-plot of  $\Delta G^{\circ}_{MMGBSA}$  values of several ligands binding to  $M^{pro}$ . The ligands include N3 inhibitor, leupeptin (LPT), and secondary metabolites in *N. sativa* seeds. These secondary metabolites are nigellamine A2 (NA2), nigellamine A3 (NA3), melanthigenin (MTN), nigellidine-4-O-sulfite (NGS), taraxerol (TRX), and nigellidine (NGL)



**Fig 2.** RMSD plots of  $M^{pro}$  in apo and ligand-bound forms. The ligands are N3 inhibitor, leupeptin (LPT), and secondary metabolites in *N. sativa* seeds, including nigellamine A2 (NA2), and A3 (NA3)

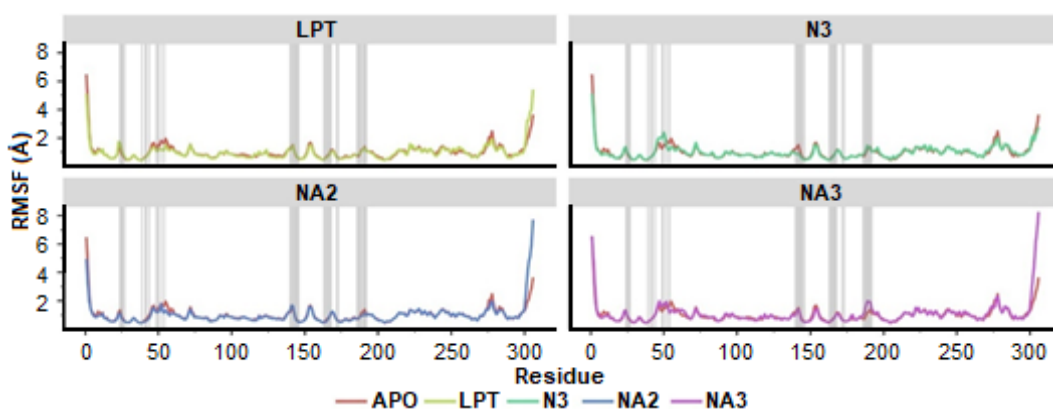
-40 kcal mol<sup>-1</sup> (Fig. S3). In contrast, the binding of N3 results in a fluctuation in the protein RMSD, where the  $\Delta G^{\circ}_{MMGBSA}$  trajectory of the ligand exhibits a dramatic change to become weaker, about -30 kcal mol<sup>-1</sup> (Fig. S3). Meanwhile, the secondary nigellamine A2, which displays a slightly stronger binding affinity than leupeptin to  $M^{pro}$ , stabilizes the protein conformation, as shown with a low RMSD fluctuation after 40 ns. During the first 40-ns, the binding of nigellamine A2 to  $M^{pro}$  experiences a strengthen. But then it weakens to around -40 kcal mol<sup>-1</sup> after 50 ns and remains until the end of the trajectory (Fig. S3).

Besides RMSD, we also computed the RMSF for the ligand-bound forms of  $M^{pro}$  (Fig. 3 and Fig. S5). RMSF complements RMSD in providing detailed information regarding conformation changes of  $M^{pro}$ . The ligand binding primarily fluctuates Met49, Leu50, Pro52, Tyr54, Phe140, Leu141, Asn142, Gly143, Ser144, Cys145,

Gly146, Val186, Asp187, Arg188, Gln189, Thr190, Ala191, and Gln192. They are residues forming the binding site of  $M^{pro}$ .

The binding of reference ligands mainly lowers the conformational changes of  $M^{pro}$ . The N3 inhibitor reduces the fluctuation of Pro52, Tyr54, and Glu55, but it increases the conformational changes on Glu47, Asp48, Met49, and Leu50 as compensation. Additionally, such binding also decreases the fluctuation of Phe140, Leu141, Asn142, Gly143, Ser144, and Cys145. The binding of leupeptin causes a slight increase of fluctuation around Asn142, Gly143, Ser144, Cys145, Gly146.

Two secondary metabolites with similar  $\Delta G^{\circ}_{MMGBSA}$  values to leupeptin, nigellamine A2, and A3, affect differently on the fluctuation of  $M^{pro}$ . While nigellamine A3 increases the conformational changes of amino acid



**Fig 3.** RMSF plots of  $M^{pro}$  in apo and ligand-bound forms. The ligands are N3 inhibitor, leupeptin (LPT), and secondary metabolites in *N. sativa* seeds, including nigellamine A2 (NA2) and A3 (NA3). The binding residues of  $M^{pro}$  are shaded by grey color

**Table 2.** Non-bonded Interactions between M<sup>Pro</sup> and ligands before and after MD simulations. For MD simulations, the non-bonded interactions were extracted from the last frame of the MD trajectories

Interaction	LPT		N3		NA2		NA3	
	D*	MD^	D	MD	D	MD	D	MD
H-bond	5	4	4	7	2	1	4	1
Unconventional H-bond	5	5	6	9	5	3	1	1
Alkyl	2		2	2		6	2	1
Pi-Alkyl	1	1	3	1	2	1	3	2
Pi-Pi					1	1		1
Amide-Pi						1		
Pi-sulphur		1				1		
Unfavorable	2		2					

\*D: molecular docking

^MD: molecular dynamics simulation

residues at the binding site of M<sup>Pro</sup>, nigellamine A2 reduces the fluctuation of the protein entirely. Such influence of nigellamine A2 on protein structural dynamics suggests that the compound is the most potent antiviral for COVID-19 among other secondary metabolites in *N. sativa* seeds, targeting M<sup>Pro</sup>.

### Non-Bonded Interactions between M<sup>Pro</sup> and Ligands

As discussed above, the order of ligand binding affinity to M<sup>Pro</sup> alters from molecular docking (Table 1) to MD simulation results (Fig. 1). Unlike molecular docking, MD simulations allow ligands and M<sup>Pro</sup> to move and interplay, adjusting their intermolecular interactions [23,25]. According to molecular docking results, N3 and leupeptin display two unfavorable interactions with M<sup>Pro</sup> (Table 2, Fig. S6). After MD simulations, these interactions disappear, and N3 increases its intermolecular H-bond, while leupeptin decreases such interaction with M<sup>Pro</sup>. Nigellamine A3 experiences some reductions regarding intermolecular interactions with M<sup>Pro</sup> (Table 2), which may explain the decrease of its binding affinity to M<sup>Pro</sup> during MD trajectory. Meanwhile, nigellamine A2 reduces conventional and unconventional H-bonds, the ligand creates new non-bonded interactions of six alkyl-alkyl, one amide-Pi, and one Pi-sulphur (Table 2). These various interactions of nigellamine A2 to M<sup>Pro</sup> may describe the comparable binding of the ligand and

leupeptin to the enzyme.

### CONCLUSION

In this study, we have explored the potency of *N. sativa* seed in inhibiting M<sup>Pro</sup>, based on virtual screening campaigns of 67 secondary metabolites. In the beginning, the secondary metabolites in *N. sativa* seeds were filtered using Lipinski's rule of five, which defines their probabilities as oral drugs. The filtered secondary metabolites were subjected to virtual screening campaigns using a molecular docking approach. Using N3 and leupeptin as the reference ligands, the virtual screening stage screened seven secondary metabolites as inhibitor candidates targeting M<sup>Pro</sup>. Further binding energy refinement, *i.e.*,  $\Delta G^{\circ}_{MMGBSA}$ , using a molecular dynamics method suggested three secondary metabolites as the potential inhibitor candidates. They are nigellamine A2 and A3. With  $\Delta G^{\circ}_{MMGBSA}$ , RMSD, and RMSF data, this study suggests nigellamine A2 as the most potent inhibitor candidate for inhibiting M<sup>Pro</sup>. This study's results are beneficial in uncover the potency *N. sativa* seed, combating COVID-19, and assist further wet laboratory experiments. Furthermore, secondary metabolites in *N. sativa* seed may have potential inhibition activities against other functional proteins in SARS-CoV-2, such as RNA-dependent RNA polymerase (RdRp) and non-structural protein 16/10 (NSP16/10). Therefore, other investigations can be directed to these proteins.



## ■ SUPPORTING INFORMATION

Supporting information (SI) contains Table S1-S4, Fig. S1-S6, and Note.

## ■ ACKNOWLEDGMENTS

We would like to Universitas Padjadjaran for financial supports. This research was funded by *Hibah Riset Data Pustaka dan Daring* (RDPD), grant number 1735/UN6.3.1/LT/2020, and supported by Universitas Padjadjaran Academic Leadership Grant (ALG), grant number 1959/UN6.3.1/PT.00/2021. We also acknowledge ChemAxon for providing a free trial license of MarvinSketch 18.21.0.

## ■ AUTHOR CONTRIBUTIONS

AH conducted the experiment, AH and MY analyzed the data, AH, MY, IWH, SI, and UMSS wrote and revised the manuscript. All authors agreed to the final version of this manuscript.

## ■ REFERENCES

- [1] WHO, 2020, *Pneumonia of unknown cause – China*, <https://www.who.int/csr/don/05-january-2020-pneumonia-of-unknown-cause-china/en/>, accessed on May 16, 2021.
- [2] Gorbalenya, A.E., Baker, S.C., Baric, R.S., de Groot, R.J., Drosten, C., Gulyaeva, A.A., Haagmans, B.L., Lauber, C., Leontovich, A.M., Neuman, B.W., Penzar, D., Perlman, S., Poon, L.L.M., Samborskiy, D.V., Sidorov, I.A., Sola, I., and Ziebuhr, J., 2020, The species *Severe acute respiratory syndrome-related coronavirus*: Classifying 2019-nCoV and naming it SARS-CoV-2, *Nat. Microbiol.*, 5 (4), 536–544.
- [3] WHO, 2020, *WHO announces COVID-19 outbreak a pandemic*, <http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic>, accessed on May 16, 2021.
- [4] Mathieu, E., Ritchie, H., Ortiz-Ospina, E., Roser, M., Hasell, J., Appel, C., Giattino, C., and Rod s-Guirao, L., 2021, A global database of COVID-19 vaccinations, *Nat. Hum. Behav.*, 5 (7), 947–953.
- [5] Worldometers, 2021, *COVID-19 coronavirus pandemic*, <https://www.worldometers.info/coronavirus/>, accessed on May 16, 2021.
- [6] Kneller, D.W., Galanie, S., Phillips, G., O’Neill, H.M., Coates, L., and Kovalevsky, A., 2020, Malleability of the SARS-CoV-2 3CL M<sup>pro</sup> active-site cavity facilitates binding of clinical antivirals, *Structure*, 28 (12), 1313–1320.e3.
- [7] Yimer, E.M., Tuem, K.B., Karim, A., Ur-Rehman, N., and Anwar, F., 2019, *Nigella sativa* L. (black cummin): A promising natural remedy for wide range of illnesses, *Evidence-Based Complementary Altern. Med.*, 2019, 1528635.
- [8] Khan, M.A., and Afzal, M., 2016, Chemical composition of *Nigella sativa* Linn: Part 2 Recent advances, *Inflammopharmacology*, 24 (2-3), 67–79.
- [9] Ashraf, S., Ashraf, S., Ashraf, M., Imran, M.A., Kalsoom, L., Siddiqui, U.N., Farooq, I., Habib, Z., Ashraf, S., Ghufuran, M., Akram, M.K., Majeed, N., Zain-ul-Abdin, Akmal, R., Rafique, S., Nawaz, K., Yousaf, M.I.K., Ahmad, S., Shahab, M.S., Nadeem, M.F., Azam, M., Zheng, H., Malik, A., Ayyaz, M., Mahmud, T., Saboor, Q.A., Ahmad, A., Ashraf, M., Izhar, M., Hilal, A., Muhammad, A., Shaukat, Z., Khaqan, A., Hayat, K., Arshad, S., Hassan, M., Abeer-bin-Awais, Ahmad, A., Mughal, T., Virk, A.R., Umer, M., Suhail, M., Zulfiqar, S., Sarfraz, S., Anwar, M.I., Humayun, A., Khokhar, R.A., and Siddique, S., 2020, Honey and *Nigella sativa* against COVID-19 in Pakistan (HNS-COVID-PK): A multi-center placebo-controlled randomized clinical trial, *medRxiv*, Preprint article, has not been reviewed.
- [10] Nickavar, B., Mojab, F., Javidnia, K., and Amoli, M.A.R., 2003, Chemical composition of the fixed and volatile oils of *Nigella sativa* L. from Iran, *Z. Naturforsch., C: Biosci.*, 58 (9-10), 629–631.
- [11] Ahmad, A., Husain, A., Mujeeb, M., Khan, S.A., Najmi, A.K., Siddique, N.A., Damanhour, Z.A., and Anwar, F., 2013, A review on therapeutic potential of *Nigella sativa*: A miracle herb, *Asian Pac. J. Trop. Biomed.*, 3 (5), 337–352.
- [12] Mehta, B.K., Pandit, V., and Gupta, M., 2009, New

- principles from seeds of *Nigella sativa*, *Nat. Prod. Res.*, 23 (2), 138–148.
- [13] Merfort, I., Wray, V., Barakat, H.H., Hussein, S.A.M., Nawwar, M.A.M., and Willuhn, G., 1997, Flavonol triglycosides from seeds of *Nigella sativa*, *Phytochemistry*, 46 (2), 359–363.
- [14] Yuan, T., Nahar, P., Sharma, M., Liu, K., Slitt, A., Aisa, H.A., and Seeram, N.P., 2014, Indazole-type alkaloids from *Nigella sativa* seeds exhibit antihyperglycemic effects via AMPK activation *in vitro*, *J. Nat. Prod.*, 77 (10), 2316–2320.
- [15] Morikawa, T., Xu, F., Ninomiya, K., Matsuda, H., and Yoshikawa, M., 2004, Nigellamines A3, A4, A5, and C, new dolabellane-type diterpene alkaloids, with lipid metabolism-promoting activities from the Egyptian medicinal food black cumin, *Chem. Pharm. Bull.*, 52 (4), 494–497.
- [16] Yusuf, M., Hardianto, A., Muchtaridi, M., Nuwarda, R.F., and Subroto, T., 2019, “Introduction of Docking-Based Virtual Screening Workflow Using Desktop Personal Computer” in *Encyclopedia of Bioinformatics and Computational Biology*, Volume 2, Eds. Ranganathan, S., Gribskov, M., Nakai, K., and Schönbach, B., Academic Press, Oxford, 688–699.
- [17] Khan, S.L., Siddiqui, F.A., Jain, S.P., and Sonwane, G.M., 2021, Discovery of potential inhibitors of SARS-CoV-2 (COVID-19) Main Protease (Mpro) from *Nigella Sativa* (black seed) by molecular docking study, *Coronaviruses*, 2 (3), 384–402.
- [18] Bouchentouf, S., and Missoum, N., 2020, Identification of compounds from *Nigella Sativa* as new potential inhibitors of 2019 novel Coronasvirus (Covid-19): Molecular docking study, *Preprints*, 2020, 2020040079.
- [19] Maiti, S., Banerjee, A., Nazmeen, A., Kanwar, M., and Das, S., 2020, Active-site molecular docking of Nigellidine with nucleocapsid-NSP2-M<sup>Pro</sup> of COVID-19 and to human IL1R-IL6R and strong antioxidant role of *Nigella-sativa* in experimental rats, *J. Drug Targeting*, 0 (ja), 1–23.
- [20] Sumaryada, T., and Pramudita, C.A., 2021, Molecular docking evaluation of some Indonesian’s popular herbals for a possible COVID-19 treatment, *Biointerface Res. Appl. Chem.*, 11 (3), 9827–9835.
- [21] Ferdian, P.R., Elfirta, R.R., Emilia, Q., and Ikhwan, A.Z.N., 2020, Inhibitory potential of black seed (*Nigella sativa* L.) bioactive compounds towards main protease of SARS-CoV-2: *In silico* study, *Ann. Bogor.*, 24 (2), 81–94.
- [22] Forli, S., Huey, R., Pique, M.E., Sanner, M.F., Goodsell, D.S., and Olson, A.J., 2016, Computational protein–ligand docking and virtual drug screening with the AutoDock suite, *Nat. Protoc.*, 11 (5), 905–919.
- [23] Hardianto, A., Khanna, V., Liu, F., and Ranganathan, S., 2019, Diverse dynamics features of novel protein kinase C (PKC) isozymes determine the selectivity of a fluorinated balanol analogue for PKC $\epsilon$ , *BMC Bioinf.*, 19 (13), 187–197.
- [24] Hardianto, A., Yusuf, M., Liu, F., and Ranganathan, S., 2017, Exploration of charge states of balanol analogues acting as ATP-competitive inhibitors in kinases, *BMC Bioinf.*, 18 (16), 19–29.
- [25] Hardianto, A., Liu, F., and Ranganathan, S., 2018, Molecular dynamics pinpoint the global fluorine effect in balanoids binding to PKC $\epsilon$  and PKA, *J. Chem. Inf. Model.*, 58 (2), 511–519.
- [26] Lipinski, C.A., 2016, Rule of five in 2015 and beyond: Target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions, *Adv. Drug Delivery Rev.*, 101, 34–41.
- [27] Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Petersson, G.A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A., Bloino, J., Janesko, B.G., Gomperts, R., Mennucci, B., Hratchian, H.P., Ortiz, J.V., Izmaylov, A.F., Sonnenberg, J.L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V.G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K.,

- Montgomery, Jr., J.A., Peralta, J.E., Ogliaro, F., Bearpark, M., Heyd, J.J., Brothers, E., Kudin, K.N., Staroverov, V.N., Keith, T., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Millam, J.M., Klene, M., Adamo, C., Cammi, R., Ochterski, J.W., Martin, R.L., Morokuma, K., Farkas, O., Foresman, J.B., and Fox, D.J., 2016, *Gaussian 09 Revision A.02*, Gaussian, Inc., Wallingford CT.
- [28] Case, D.A., Aktulga, H.M., Belfon, K., Ben-Shalom, I.Y., Brozell, S.R., Cerutti, D.S., Cheatham, III, T.E., Cruzeiro, V.W.D., Darden, T.A., Duke, R.E., Giambasu, G., Gilson, M.K., Gohlke, H., Goetz, A.W., Harris, R., Izadi, S., Izmailov, S.A., Jin, C., Kasavajhala, K., Kaymak, M.C., King, E., Kovalenko, A., Kurtzman, T., Lee, T.S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Machado, M., Man, V., Manathunga, M., Merz, K.M., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., O'Hearn, K.A., Onufriev, A., Pan, F., Pantano, S., Qi, R., Rahnamoun, A., Roe, D.R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C.L., Skrynnikov, N.R., Smith, J., Swails, J., Walker, R.C., Wang, J., Wei, H., Wolf, R.M., Wu, X., Xue, Y., York, D.M., Zhao, S., and Kollman, P.A., 2020, *Amber 2020*, University of California, San Francisco.
- [29] Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C., 2015, ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB, *J. Chem. Theory Comput.*, 11 (8), 3696–3713.
- [30] Egbert, M., Whitty, A., Keserű, G.M., and Vajda, S., 2019, Why some targets benefit from beyond rule of five drugs, *J. Med. Chem.*, 62 (22), 10005–10025.
- [31] Torres, P.H.M., Sodero, A.C.R., Jofily, P., and Silva, F.P., 2019, Key topics in molecular docking for drug design, *Int. J. Mol. Sci.*, 20 (18), 4574.
- [32] Singh, N., Villoutreix, B.O., and Ecker, G.F., 2019, Rigorous sampling of docking poses unveils binding hypothesis for the halogenated ligands of L-type Amino acid Transporter 1 (LAT1), *Sci. Rep.*, 9 (1), 15061.
- [33] Stewart, J.J.P., 2009, Application of the PM6 method to modeling proteins, *J. Mol. Model.*, 15 (7), 765–805.
- [34] Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., Yang, X., Bai, F., Liu, H., Liu, X., Guddat, L.W., Xu, W., Xiao, G., Qin, C., Shi, Z., Jiang, H., Rao, Z., and Yang, H., 2020, Structure of M<sup>Pro</sup> from SARS-CoV-2 and discovery of its inhibitors, *Nature*, 582 (7811), 289–293.
- [35] Simmons, G., Gosalia, D.N., Rennekamp, A.J., Reeves, J.D., Diamond, S.L., and Bates, P., 2005, Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry, *Proc. Natl. Acad. Sci. U.S.A.*, 102 (33), 11876–11881.