

Targeted Covalent Inhibitors Allosterically Deactivate the DEDDh Lassa Fever Virus NP Exonuclease from Alternative Distal Sites

Kuan-Wei Huang,[†] Jing-Wen Chen,[†] Tzu-Yu Hua, Yu-Yu Chu, Tsai-Yuan Chiu, Jung-Yu Liu, Chun-I Tu, Kai-Cheng Hsu, Ya-Ting Kao, Jhih-Wei Chu,* and Yu-Yuan Hsiao*



Cite This: *JACS Au* 2021, 1, 2315–2327



Read Online

ACCESS |

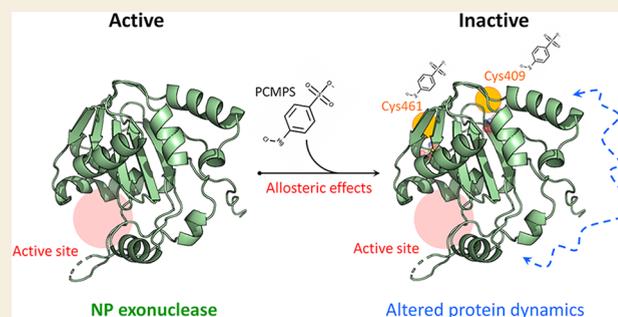
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: For using targeted covalent inhibitors (TCIs) as anticancer and antiviral drugs, we establish that the model compounds PCMPS (*p*-chloromercuriphenyl sulfate) and PCMB (*p*-chloromercuribenzoate) are inhibitors of the DEDDh family of exonucleases. The underlying mechanism is analyzed by X-ray crystallography, activity/nucleic acid-binding assays, and all-atom molecular dynamics (MD) simulations. The first TCI-complexed structures of a DEDDh enzyme, the Lassa fever virus NP exonuclease (NPexo), are resolved to elucidate that the Cys409 binding site is away from the active site and the RNA-binding lid. The NPexo C409A structures indicate Cys461 as the alternative distal site for obstructing the equally active mutant. All-atom MD simulations of the wild type and mutant NPexos in explicit solvent uncover an allosteric inhibition mechanism that the local perturbation induced by PCMPS sulfonate propagates to impact the RNA-binding lid conformation. Binding assay studies confirm that PCMPS does affect the RNA binding of NPexo. The predicted relative potency between PCMPS and PCMB is also in line with experiments. The structural data and inhibition mechanism established in this work provide an important molecular basis for the drug development of TCIs.

KEYWORDS: organomercurial, NP exonuclease, DEDDh exonuclease, inhibitor, molecular dynamics



INTRODUCTION

TCIs are small molecules that block the protein activity by reacting with specific amino acids. Currently, TCIs account for ~30% of the marketed therapeutics¹ and are gaining recognition as a promising strategy for treating complex diseases such as cancer and viral infection. A recent example is the ARS-1620 anticancer drug that induces tumor regression with high potency and selectivity by targeting the G12C cysteine in KRAS.^{2,3} The structural information on protein-TCI complexes primarily reveals binding at a ligand-binding or catalytic site.⁴ As these loci are usually conserved, a disadvantage of TCIs is the potential off-target toxicity.¹ As an alternative mechanism for medical applications, a fundamental question is whether the chemical modification of TCIs at a distal site could exhibit functional impact allosterically.

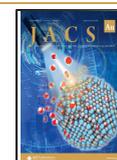
In this regard, we consider the sulfhydryl reactive PCMPS and PCMB.⁵ Such organomercurials target cysteine residues and have been shown to reduce the activities of more than 10 classes of enzymes,^{6–13} including nucleases.^{14,15} The structural data of protein-organomercurial complexes are limited to the Charcot-Leyden protein¹⁶ and carbonic anhydrase 2,^{17,18} and interestingly, they indicate that the activity reduction does not necessarily come from a change in the overall conformation.

Protein oligomerization affected by the TCIs has been proposed but does not explain their inhibition on monomeric proteins.^{7–9} Because organomercurials can competitively block other chemical modifications such as oxidation and alkylation,⁹ they have been applied to evaluate the importance of cysteine residues in a target for stability, oligomerization,^{7–9} and substrate binding.¹⁹ Interpretation of the results, though, is complicated by the unknown and site-specific effects of TCI on the enzyme conformational dynamics.

To investigate the potential of applying TCIs as antiviral or anticancer therapies, we employed PCMPS and PCMB here as model systems to study the inhibition by covalent linkage on DEDDh exonuclease family members in the activities of digesting various DNA and RNA substrates. Enzymes in this family participate in various pathways of DNA repair, genotoxic stress response, DNA degradation, RNA maturation, RNA interference, and RNA turnover.²⁰ In addition to RNase

Received: September 24, 2021

Published: November 16, 2021



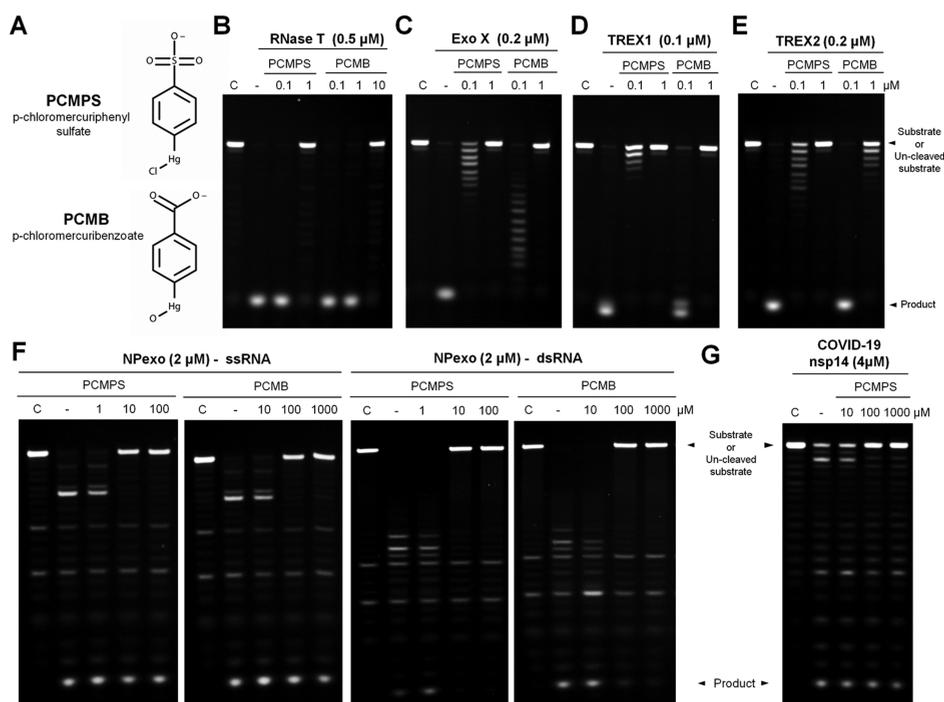


Figure 1. Inhibition efficiency of PCMPS and PCMB on various DEDDh exonucleases. (A) Chemical structures of PCMPS and PCMB. (B–E) Inhibition efficiency of PCMPS and PCMB on 0.5 μM RNase T, 0.2 μM Exo X, 0.1 μM TREX1, and 0.2 μM TREX2. The substrates are ssDNA. PCMPS and PCMB fully inhibit the nuclease activity of these DEDDh exonucleases at an inhibitor concentration between 1 and 10 μM . Capital C stands for the control group. (F) Inhibition efficiency of PCMPS and PCMB on 2 μM NPexo in digesting ssRNA or dsRNA substrates. PCMPS and PCMB fully inhibit the nuclease activity of NPexo at an inhibitor concentration between 10 and 100 μM . (G) Inhibition efficiency of PCMPS on 4 μM SARS-CoV-2 nsp14. The substrate is ssRNA. PCMPS fully inhibits the nuclease activity of nsp14 at an inhibitor concentration of 100 μM .

T, CRN-4, and NPexo in Lassa fever virus,^{14,15} PCMPS and PCMB are established here as inhibitors that can broadly affect DEDDh exonucleases, also inhibiting Exo X, TREX1, TREX2, and SARS-CoV-2 nsp14. Although TREX1 is a promising route for cancer immunotherapy,^{21,22} NPexo is an important target for blocking the Lassa fever virus replication.^{23–26} Similarly, nsp14 is essential for COVID-19 pathogenesis because of the RNA proofreading function and a focus of drug development.^{27–30} nsp14 inhibitors potentially have broad applications in coronavirus (CoV) family related therapies.^{28–32}

In this work, the first TCI-complexed structures of a DEDDh family member are solved by crystallizing the C-terminal domain of NPexo in the apo form and in the presence of PCMPS or PCMB for both the native sequence and the C409A mutant. The overall NPexo conformation is well preserved, and TCI ligation only slightly perturbs local structures. The first targeted site is found to be Cys409 distal to the active site. NPexo C409A retains the exonuclease activity, but PCMPS or PCMB still inhibits the enzyme by targeting Cys461, which is also away from the DEDDh active site region.

To elucidate the underlying mechanism of such TCI inhibition from distal sites, a series of all-atom molecular dynamics (MD) simulations in explicit solvent were conducted starting from the newly solved NPexo structures.

Our results show that through the interaction network in the DEDDh exonuclease structure, local perturbation mediated by the ligated PCMPS can propagate to affect the conformation at the substrate-binding region. To test the predicted mechanism of allosteric inhibition, we used gel shift assays to confirm that the distal bonding of PCMPS or PCMB indeed affects the

substrate binding ability in both the wild-type and mutant NPexo.

RESULTS

DEDDh Exonucleases Are Inhibited by PCMPS and PCMB

The chemical structures of PCMPS and PCMB are illustrated in Figure 1A. As a systematic test whether these organomercurial TCIs can broadly inhibit DEDDh exonucleases, we purified six classical members, involving *E. coli* RNase T, *E. coli* Exo X, mouse TREX1, mouse TREX2, SARS-CoV-2 nsp14, and Lassa fever virus NPexo (Figure S1A), to conduct inhibitor-coupled nuclease activity assay. Sequences of the DNA/RNA substrates are listed at Table S1. For inhibiting RNase T, Exo X, TREX1, and TREX2, the 0.1 μM and 1 μM PCMPS concentration was found to reduce and abolish the exonuclease activity, respectively, in digesting a FAM-labeled ssDNA 20mer substrate (Figure 1B–E). To achieve similar inhibition over these enzymes, the required PCMB concentration is 1 to 10 μM (Figure 1B–E). For SARS-CoV-2 nsp14, 100 μM PCMPS inhibits the exonuclease activity in digesting FAM-labeled ssRNA 20mer (Figure 1G).

In addition to the wild type NPexo, its two mutant forms of the C409A single mutation and the C409A/C461A double mutation were also purified, and both mutants demonstrate exonuclease activities and substrate preference similar to those of the wild type. At the 0.5 μM protein concentration, all three NPexos completely degrade full-length dsRNA but not full-length ssRNA (Figure S1B, C). NPexo digesting dsRNA more effectively is consistent with the observation in previous studies.^{24,25}

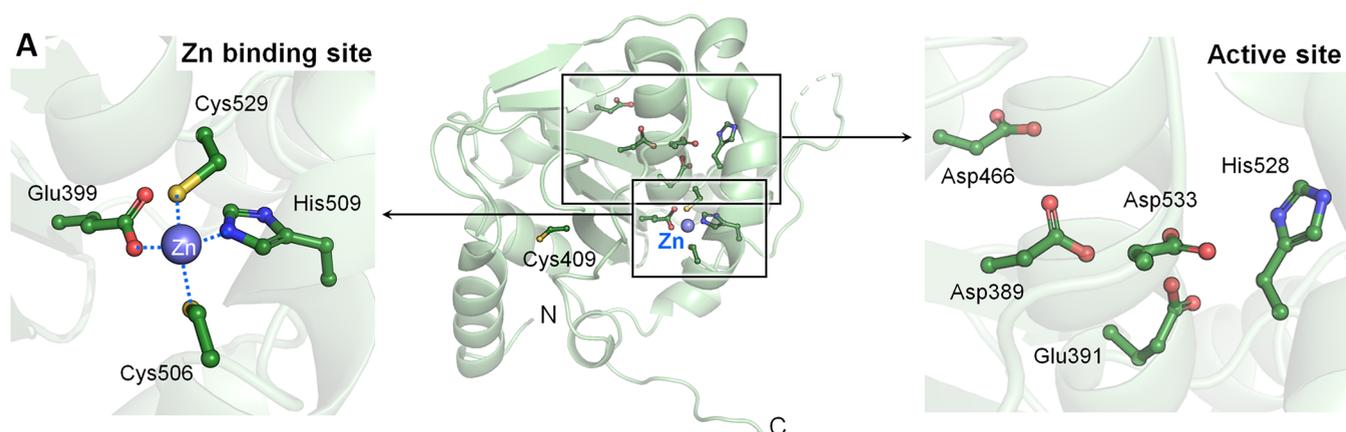


Figure 2. Apo-NPexo crystal structure. Left, overall structure of the nuclease domain of NPexo. Right, active site structure of NPexo, which displays a classical conformation similar to other DEDDh exonucleases. Middle, structure of the Zn binding site. The zinc ion has a structural role bound by Glu399, Cys506, His509, and Cys529.

Table 1. Data Collection and Refinement Statistics^a

| | apo-NPexo | NPexo-PCMPS complex | NPexo-PCMB complex | apo-NPexo C409A | NPexo C409A-PCMPS complex | NPexo C409A-PCMB complex |
|--|-----------------------|---|---|---|---|---|
| data collection | PDB: 7V37 | PDB: 7V38 | PDB: 7V39 | PDB: 7V3A | PDB: 7V3B | PDB: 7V3C |
| wavelength (Å) | 0.97622 | 0.99311 | 1.00000 | 1.00000 | 0.99258 | 0.99263 |
| space group | I222 | P2 ₁ 2 ₁ 2 ₁ |
| asymmetric unit | 2 monomers | 2 monomers | 2 monomers | 2 monomers | 2 monomers | 2 monomers |
| cell dimensions | | | | | | |
| <i>a</i> , <i>b</i> , <i>c</i> (Å) | 74.13, 121.27, 137.72 | 50.23, 74.01, 140.81 | 50.19, 76.29, 140.78 | 59.43, 71.53, 127.32 | 58.90, 69.19, 122.70 | 58.81, 68.85, 123.85 |
| α , β , γ (deg) | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 | 90.0, 90.0, 90.0 |
| resolution (Å) | 30.0–2.4 (2.49–2.4) | 30.0–2.4 (2.49–2.4) | 30.0–2.2 (2.28–2.2) | 30.0–2.1 (2.18–2.1) | 30.0–1.8 (1.86–1.8) | 30.0–1.9 (1.97–1.9) |
| <i>R</i> _{sym} or <i>R</i> _{merge} | 10.1 (49.4) | 8.5 (49.6) | 8.9 (39.5) | 8.9 (44.7) | 6.3 (49.7) | 5.8 (47.6) |
| <i>I</i> / σ <i>I</i> | 15.0 (2.0) | 36.2 (2.6) | 28.9 (4.7) | 25.5 (2.6) | 41.1 (3.2) | 41.6 (2.3) |
| completeness (%) | 98.8 (98.3) | 98.9 (92.5) | 99.9 (100.0) | 99.8 (99.5) | 99.8 (99.8) | 99.6 (99.5) |
| redundancy | 4.6 (3.7) | 11.8 (10.2) | 6.3 (6.1) | 6.5 (5.5) | 11.1 (10.4) | 9.2 (6.7) |
| phasing method | molecular replacement | Hg-SAD | molecular replacement | molecular replacement | Hg-SAD | molecular replacement |
| refinement | | | | | | |
| resolution (Å) | 28.7–2.4 | 29.8–2.4 | 28.8–2.2 | 29.8–2.1 | 29.9–1.8 | 26.4–1.9 |
| No. reflections | 23991 | 20782 | 28191 | 32331 | 47464 | 40356 |
| <i>R</i> _{work} / <i>R</i> _{free} | 20.1/25.0 | 21.2/26.1 | 20.4/23.5 | 20.6/24.1 | 17.8/21.8 | 18.8/22.4 |
| rms deviations | | | | | | |
| bond lengths (Å) | 0.002 | 0.002 | 0.002 | 0.002 | 0.007 | 0.004 |
| bond angles (deg) | 0.495 | 0.490 | 0.430 | 0.403 | 0.889 | 0.636 |
| Ramachandran plot statistics (%) | | | | | | |
| favoured region | 98.2 | 97.52 | 97.51 | 98.01 | 97.07 | 97.46 |
| allowed region | 1.8 | 2.48 | 2.49 | 1.99 | 2.93 | 2.54 |
| outlier region | 0 | 0 | 0 | 0 | 0 | 0 |

^aEach structure was obtained from a single crystal. Values in parentheses are for highest-resolution shell.

PCMPS and PCMB can inhibit the activities of the wild type as well as the mutant NPexos. With the enzyme at 2 μ M, 10 μ M PCMPS and 100 μ M PCMB are sufficient to reduce the exonuclease activity as most of FAM-labeled ssRNA or dsRNA are retained (Figure 1F and Figure S4). Refined analysis shows that 4 μ M PCMPS completely blocks the ssRNA digestion at 2 μ M NPexo, while for the more susceptible dsRNA, 0.6 μ M PCMPS blocks the activity at 0.3 μ M NPexo (Figure S1D). Together, PCMPS and PCMB are indeed inhibitors of DEDDh exonucleases with the former exhibiting stronger inhibition power.

Crystal Structures of NPexo Complexed with PCMPS or PCMB

To uncover the targeted residue by which the organomercurial TCI inhibits a DEDDh exonuclease, we crystallized and determined the X-ray structures of the nuclease domain of NPexo (342–569) in the apo form and in complexation with PCMPS or PCMB. The PCMPS or PCMB were incubated with the NPexo protein for 10 min at 4 °C before cocrystallization experiments. In the absence of an inhibitor, the nuclease domain of NPexo exhibits an identical fold as in the structures of a full-length NPexo²³ and a dsRNA bound nuclease domain of NPexo^{24,25} (Figure 2 and Figure S2).

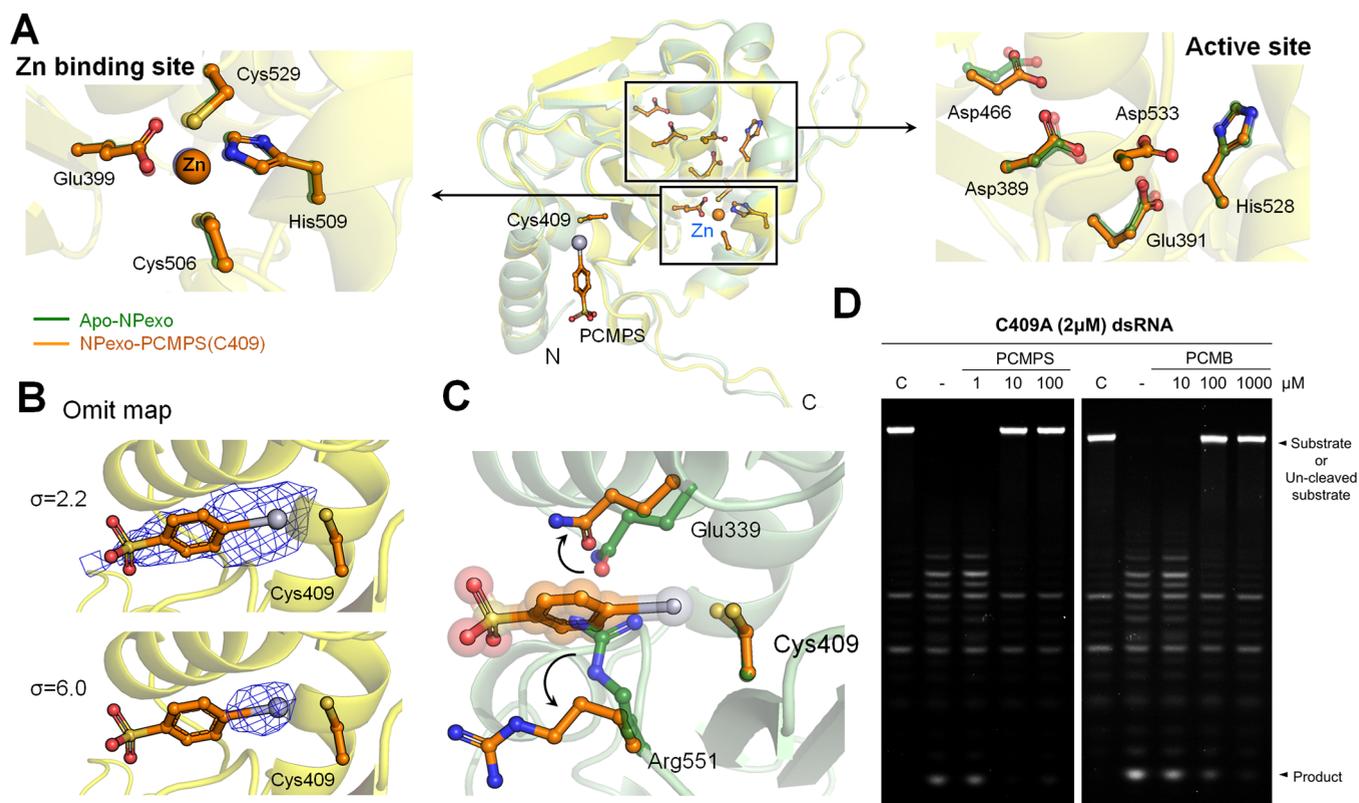


Figure 3. NPexo-PCMPS complex crystal structure. (A) Overall structure of the NPexo-PCMPS complex. The residues of apo- and PCMPS-bound NPexo are displayed by transparent green and orange, respectively. The overall structure of the NPexo-PCMPS complex is very similar to that of apo-NPexo. The right and left panels show that the active site and Zn binding site structures of the PCMPS-bound NPexo are very similar to those in the apo-NPexo structure. (B) Omitted electron density map ($F_o - F_c$, 2.2 σ and 6.0 σ) of PCMPS in the NPexo-PCMPS complex. (C) PCMPS-induced side chain movement. The side chains of apo-NPexo and PCMPS-bound NPexo are colored in green and orange, respectively. (D) Inhibitor-coupled nuclease activity assays of mutated NPexo, C409A. The 2 μM NPexo C409A is inhibited by PCMPS or PCMB in a dose-dependent manner. Most dsRNA substrate is not degraded with 10 μM PCMPS or 100 μM PCMB in the reaction mixture. Capital C stands for the control group.

Therefore, truncation of the N-terminal domain and substrate binding do not affect the NPexo structure.

The 3'-5' exonuclease activity in the DEDDh family has a common nuclease domain of a mixed $\alpha\beta$ fold with four conserved acidic residues (DEDD) and a histidine (h) general base to bind two magnesium ions at the active site.^{20,33} The NPexo active-site residues Asp389, Glu391, Asp466, His528, and Asp533 in the apo form (apo-NPexo) adopt the same arrangement as in other DEDDh exonucleases (Figure 2). The apo-NPexo structure also contains a zinc ion coordinated by Glu399, Cys506, His509, and Cys529 near the active site (Figure 2). The zinc ion is not involved in catalysis and ought to play a structural role.²³

Crystals of the NPexo-PCMPS complex were obtained at the crystallization reagents with 2 mM PCMPS. This 1.85- to 2.4-fold excess of inhibitor over NPexo in the solution phase of crystal growing is within the range conducted for the inhibitor-coupled nuclease activity assays. The mercury signals were screened by a fluorescence energy scan before the diffraction data we collected, and the NPexo-PCMPS structure was determined by single-wavelength anomalous dispersion (SAD) and refined to the 2.4 Å resolution. For the NPexo-PCMPS structure, the crystallization conditions, diffraction, and refinement statistics are reported in Table S2 and Table 1.

As the first structure for a DEDDh family member complexed with a TCI, the NPexo-PCMPS structure illustrates

that PCMPS forms a mercury–sulfur bond with Cys409 (Figure 3A), which is on the protein surface and far away from the active site. The omit map analysis shows that the electron density map in the sulfonate region of PCMPS is rather broad and provides less stringent restraints in fitting to the three-dimensional structure (Figure 3B), suggesting that the sulfonate area of Cys409-PCMPS is somewhat flexible in the crystal. This result is in line with the lack of hydrogen bonding or specific interactions between PCMPS and the nearby protein atoms in the NPexo-PCMPS structure. The conformation of the active site, the Zn binding site, and the overall fold in the TCI-complexed structure remains close to those in apo-NPexo (Figure 3A), showing that the PCMPS attachment does not alter the structural integrity of NPexo. The ligation of PCMPS, though, does affect the side chain orientation of amino acids near Cys409, such as Glu339 and Arg551, to accommodate for the volume of the TCI (Figure 3C). We also determined the structure of NPexo complexed with PCMB at 2.2 Å resolution (NPexo-PCMB), which shows the same binding mode as in the NPexo-PCMPS structure (Figure S3).

NPexo C409A Structures Reveal the TCI Targeting on Alternative Cysteine Residues

With the exonuclease activity of NPexo C409A unaffected (Figure S1B, C), PCMPS and PCMB still inhibit the mutant enzyme with similar efficacies as in the case of the wild-type (Figure 3D and Figure S4A). To elucidate the underlying

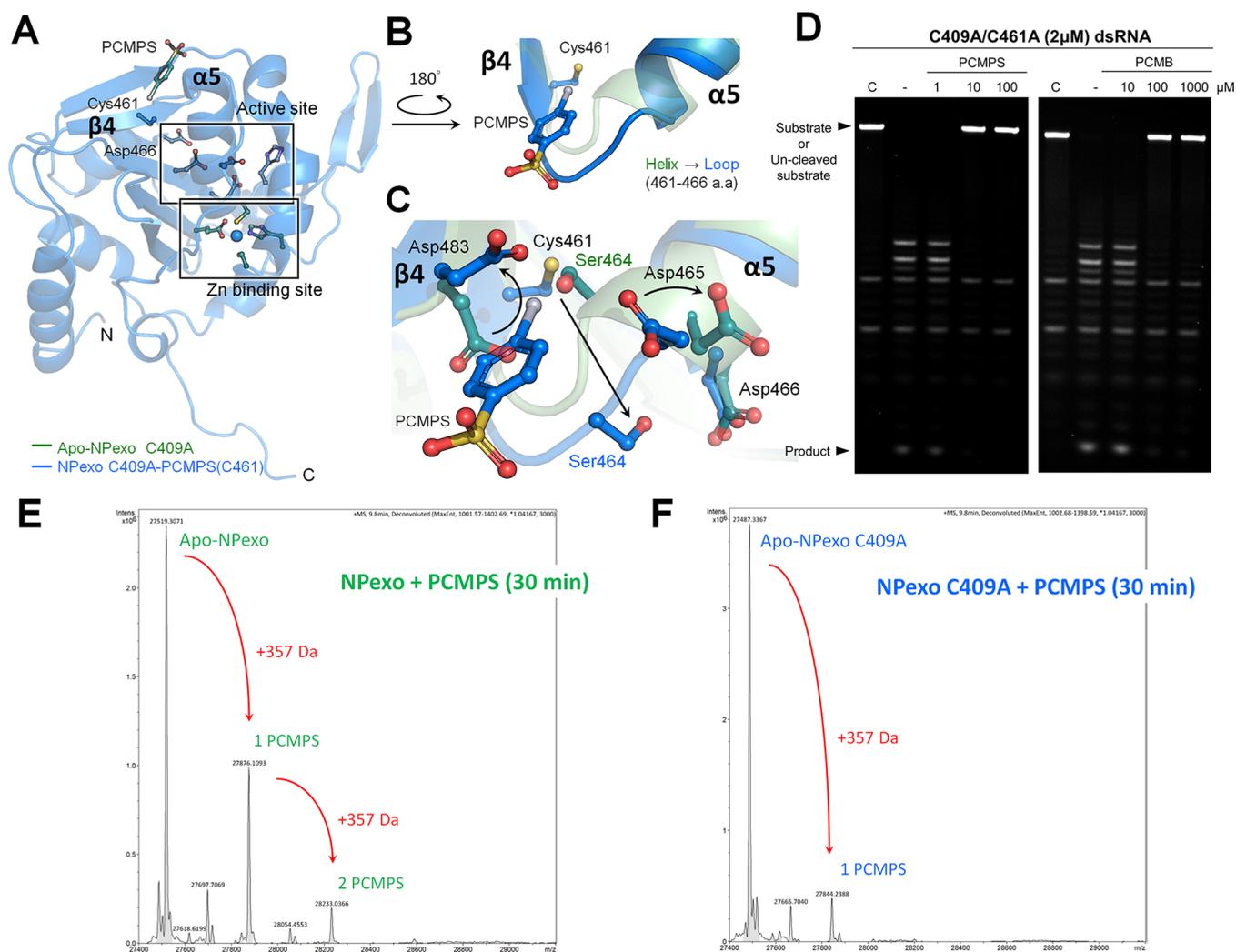


Figure 4. NPexo C409A-PCMPS complex crystal structure and liquid chromatography–mass spectrometry (LC-MS) analysis. (A) Overall structure of the NPexo C409A-PCMPS complex. PCMPS binds to Cys461. (B, C) Closer view and structural comparison of the PCMPS binding site. The apo-NPexo C409A and the PCMPS bound structures are displayed in green and blue, respectively. Black arrows indicate the moving paths of structurally changed residues. (D) PCMPS or PCMB coupled nuclease activity assay of the double mutant, 2 μ M C409A/C461A. Capital C stands for the control group. (E, F) LC-MS spectrometry of 30 min PCMPS-treated wild-type NPexo or NPexo C409A. There are two and one major PCMPS-bound protein peaks in wild-type NPexo and NPexo C409A, respectively. The molecular weight is increased by about 357 Da after being bound by one PCMPS molecule.

mechanism, we determined the crystal structures of NPexo C409A in the apo form (apo-NPexo C409A) and in the complexed form with PCMPS (NPexo C409A-PCMPS) by molecular replacement and SAD. The apo-NPexo C409A and the wild-type apo-NPexo structures are essentially identical (Figure S5), indicating that the enzyme structure is not affected by the C409A mutation.

The NPexo C409A-PCMPS structure was resolved at 1.8 Å resolution. There are two enzyme monomers in the asymmetric unit as in the other PCMPS- or PCMB-bound NPexo structures obtained in this work. The attached PCMPS does not lead to a significant change in the asymmetric unit, the duration for crystallization, and diffraction resolution of the wild-type and mutated NPexo crystals. One of the monomers is complexed with PCMPS (Figure S6A) and the other is inhibitor free with an identical fold as in the apo-NPexo C409A structure. With Cys409 mutated to alanine, PCMPS forms a mercury–sulfur bond with Cys461 instead (Figure 4A, B). The omit maps of PCMPS for this structure are shown in

Figure S7A. This result is a first demonstration that a TCI can target alternative cysteine residues with specific preference.

The overall topology, active site, and zinc binding region structures of NPexo C409A-PCMPS also show close similarities with those in the apo form (Figure S7B). Cys461 is less exposed compared to Cys409 and is surrounded by $\alpha 5$ (Ser464–Gln474) and a β -strand (Lys480–Ile484) (Figure S5C). Upon PCMPS attaching to Cys461, the Ser464 to Asp466 region of $\alpha 5$ is slightly perturbed to move away from Cys461, and the side chain of Asp483 also rotated away from Cys461 (Figure 4B, C). These local structural changes revealed in the NPexo C409A-PCMPS structure provide unprecedented information on how the conformation near the region around Cys461 would change to accommodate PCMPS. Similarly, in the NPexo C409A-PCMB structure was resolved at 1.9 Å resolution, the overall topology, active site, and the zinc binding region remain close to those in the apo-NPexo C409A structure. In NPexo C409A-PCMB, the inhibitor also ligates to

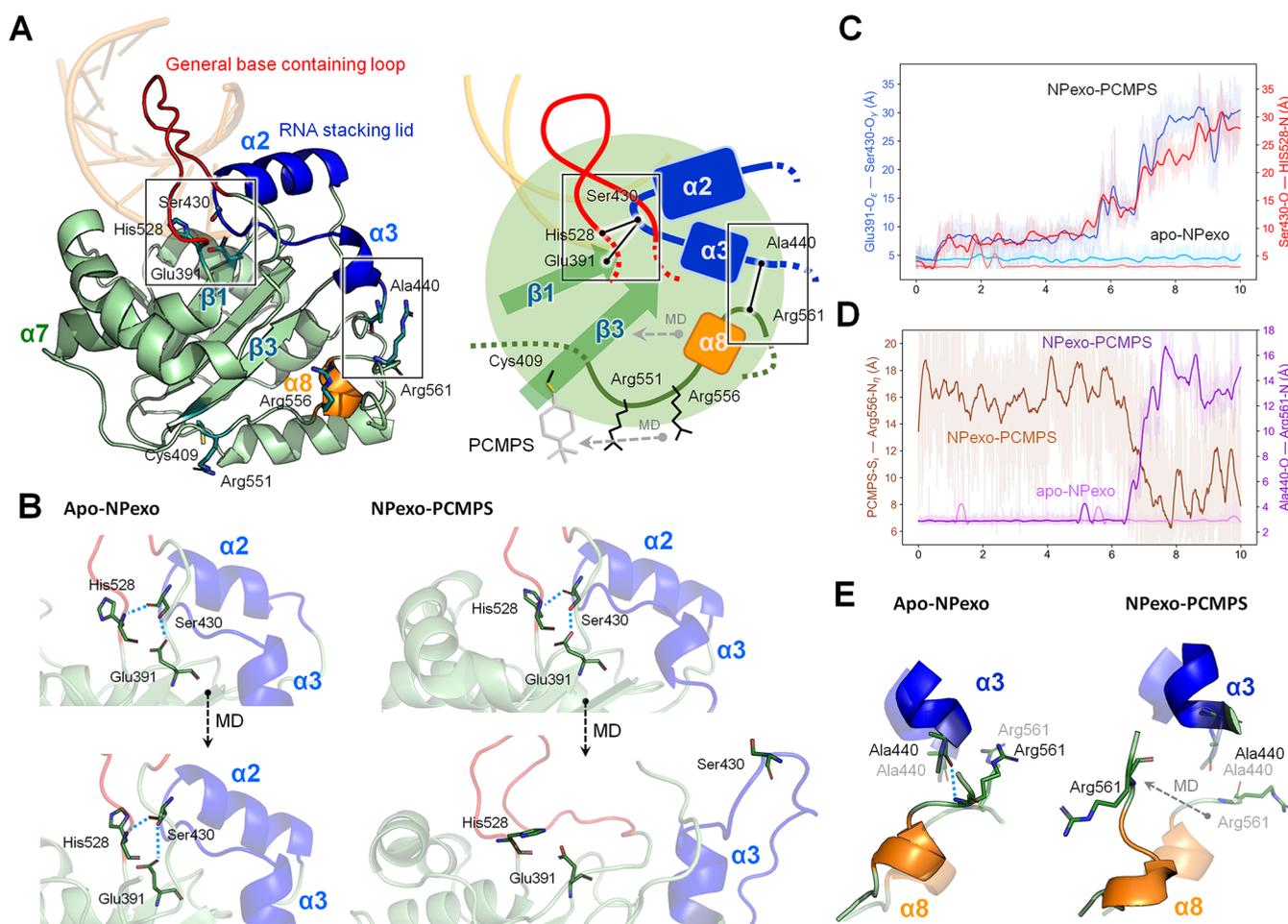


Figure 5. Structural variation of apo- and PCMPS-bound NPexo during the all-atom MD simulations in explicit solvent. (A) Overall structure and schematic representation of NPexo structural components with the key residues labeled. The structural changes during the NPexo-PCMPS simulation are highlighted by gray arrows. (B) Conformational change as indicated by the Glu391-O_{ε1}-Ser430-O_γ and His528-N-Ser430-O distances during the apo-NPexo and NPexo-PCMPS simulations. Left and right panels represent the comparison with the initial crystal structure for the representative conformations in the apo-NPexo and NPexo-PCMPS simulations, respectively. (C) Temporal variation of the Glu391-O_{ε1}-Ser430-O_γ and His528N-Ser430-O distances during the MD simulations. (D) Temporal variation of the PCMPS-S_γ-Arg556-N_η and Ala440-O-Ala440-O-Arg561-N distances during the MD simulations. (E) Conformational change as indicated by the Ala440-Arg561 distance during the MD simulations. Left and right panels represent the comparison with the initial crystal structures for the representative conformations in the apo-NPexo and NPexo-PCMPS simulations, respectively. The initial crystal structure is displayed as a transparent mode.

Cys461 and induces similar local perturbation as observed in the NPexo C409A-PCMPS structure (Figure S8).

Aligning the apo-NPexo, NPexo-PCMPS, NPexo C409A-PCMPS, and NPexo-dsRNA structures resolved here with the dsRNA-NPexo structure (PDB code: 4FVU) illustrates that Cys409 and Cys461 would not overlap or contact with the bound dsRNA (Figure 7A). The orientations of the substrate contacting residues as observed in the 4FVU structure, including Asp426, Gln462, Asp465, and Arg492 also have high similarities with those in the structures resolved here. A slight but noticeable structural variation is at Asp465 in the NPexo C409A-PCMPS structure (Figure 4C). In the 4FVU substrate-bound structure, Asp465 near the TCI-ligation site Cys461 interacts with the non-scissile strand of dsRNA.

When both Cys409 and Cys461 are mutated to alanine, the exonuclease activity of the C409A/C461A double mutant remains similar to that of the wild type as discussed earlier (Figure S1). Furthermore, PCMPS and PCMB still inhibit the activity of the double mutant (Figure 4D and Figure S4B).

Therefore, it is expected that PCMPS and PCMB target other Cys residues, such as Cys506 or Cys529 for inhibition.

As an independent test for the sequential attachment of PCMPS on Cys residues starting from Cys409, liquid chromatography–mass spectrometry (LC-MS) experiments were conducted on apo-wild-type NPexo, apo-NPexo C409A, PCMPS-bound NPexo, and PCMPS-bound NPexo C409A. The results are listed in Table S3. With 30 min of reaction time, wild-type NPexo has two additional PCMPS bound peaks, whereas NPexo C409A has only one, indicating that Cys409 is indeed a more reactive site for PCMPS binding (Figure 4E, F and Figure S9A, B). When the reaction time was increased to 60 min, more PCMPS bound peaks are observed in wild-type NPexo (four) and in NPexo C409A (three) (Figure S9C, D), showing the sequential binding of PCMPS after targeting Cys409. The multiple PCMPS bound peaks in NPexo also demonstrate the vulnerability of other cysteine sites in NPexo to PCMPS, consistent with its inhibition on the C409A/C461A mutant (Figure 4D and Figure S4B).

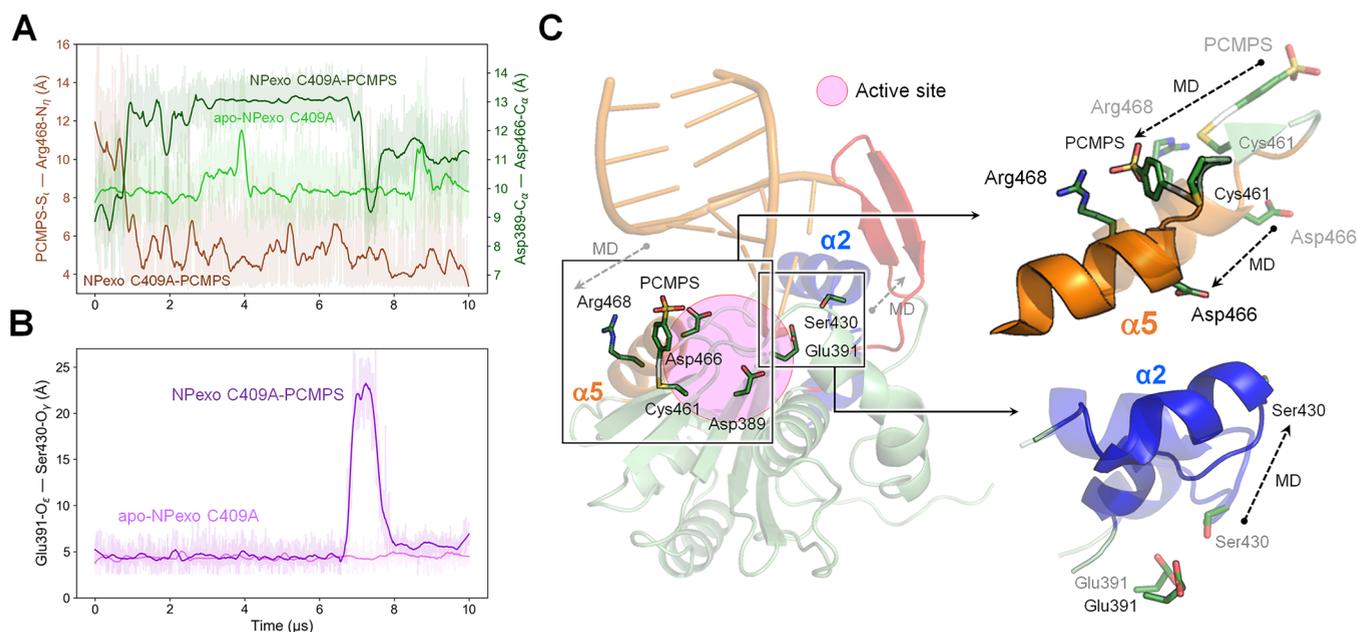


Figure 6. Structural variation of apo- and PCMPS-bound NPexo C409A during the all-atom MD simulations in explicit solvent. (A, B) Temporal variation of the PCMPS-S_γ-Arg468-N_η, Asp389-C_α-Asp466-C_ω, and Glu391-O_ε-Ser430-O_γ distances in the apo-NPexo C409A and NPexo C409A-PCMPS simulations. (C) Structural comparison around the PCMPS binding region for the representative conformations in the apo-NPexo C409A and NPexo C409A-PCMPS. Up and down panels show the conformational change around the α5 and α2 components, respectively. The initial crystal structure is displayed as a transparent mode.

Structural Components of NPexo Exhibit Diverse Flexibilities

With the NPexo structures unaffected after binding PCMPS or PCMB, a potential cause for the inhibition from a distal site is the local perturbation propagating through the interaction network and allosterically modulate substrate binding and/or catalysis. To investigate how PCMPS binding affects the conformational dynamics, we performed all-atom MD simulations starting from the newly solved X-ray structures (apo-NPexo, NPexo-PCMPS, apo-NPexo C409A, NPexo C409A-PCMPS) for a 10 μs production run. In all simulations, an explicit solvent model is used for the aqueous environment, and each sampled protein configuration is aligned to the initial X-ray structure by the C_α atoms of β-core, including β1, β2, and β3 as defined in Figure S12A. For other segments, the C_α RMSDs (root-mean-square difference) to the X-ray coordinates were also calculated, and their averaged values in MD trajectories as summarized in Table S5 show diverse conformational flexibilities.

In the apo-NPexo simulation, the β-core remains very close to the initial X-ray structure (RMSD 0.9 ± 0.1 Å, Figure S12B and Movie S1) and the DEDDh active site also has low and plateaued values (1.3 ± 0.3 Å, Figure S12B). On the other hand, the RNA-stacking lid (Leu420-Ala440) containing α2 (Leu420-Ser430) and α3 (Val435-Ala440) is more mobile (RMSD 2.6 ± 0.5 Å) and the general-base containing loop (Val513-His528) exhibits very significant fluctuation (RMSD 8.0 ± 2.6 Å, Figure S12B and Movie S1). The RNA stacking lid contains Asp426 involved in substrate binding²⁴ and forms a lidlike motif around the active site with the general-base containing loop (Figure 5A). Ser430 in the same helix α2 with Asp426 hydrogen bonds to active site Glu391 in β1 and His528 in β7, and they remain stable in the apo-NPexo simulation (Figure 5A and 5B). On the other side of the RNA-stacking lid, the 6-residue helix α3 (Val435-Ala440) couples to

the nearby α8 (Arg556-Val559) as indicated by the backbone hydrogen bonding between Ala440 and Arg561 (Figure 5E). In summary, the flexible RNA-stacking lid is specifically coupled to the nearby general base-containing loop and α8 in the apo-NPexo simulation.

TCI-Induced Perturbation Propagates Across the NPexo Interaction Network

The effects of PCMPS ligating to Cys409 are studied by the NPexo-PCMPS simulation starting from the X-ray structure. The β-core remains close to the initial coordinates as they do in the apo-NPexo simulation (RMSD 0.9 ± 0.2 Å, Figure S12C and Movie S2). The DEDDh cluster of the active site, though, exhibits higher RMSD (2.6 ± 0.8 Å). Furthermore, the RNA stacking lid detaches from β1 and experiences a drastic conformational change, leading to α2 helix turning into a loop (RMSD 11.6 ± 6.0 Å, Figure 5B, C and Movie S2). As a result, the hydrogen bonds of Ser430 with Glu391 and His528 that remain stable in the apo-NPexo simulation are disrupted in the NPexo-PCMPS simulation (Figure 5B, C).

With the TCI ligated to Cys409 at β3, the side-chain orientation of the nearby Arg551 is altered by PCMPS as shown in the X-ray structure (Figure 3C). In the NPexo-PCMPS simulation, the negatively charged PCMPS sulfate indeed strongly interacts with the Arg551 side chain (Figure 5A). Second, the side chain of nearby Arg556 in α8 is also drawn closer to PCMPS in the NPexo-PCMPS simulation, causing α8 to migrate toward β3 in the core (Figure 5A). This shift disrupts the Ala440-Arg561 coupling as in the apo-NPexo simulation since Arg561 is in α8 (Figure 5D, 5E). Although Ala440 is in the RNA-stacking region, losing this coupling with α8 leads to the drastic conformational change of the RNA-stacking lid with α2 melting in the NPexo-PCMPS simulation (Figure 5B). As mentioned earlier, the α2 helix contains Asp426 that participates in RNA binding. Therefore, the perturbation of PCMPS ligating to C409 can propagate to

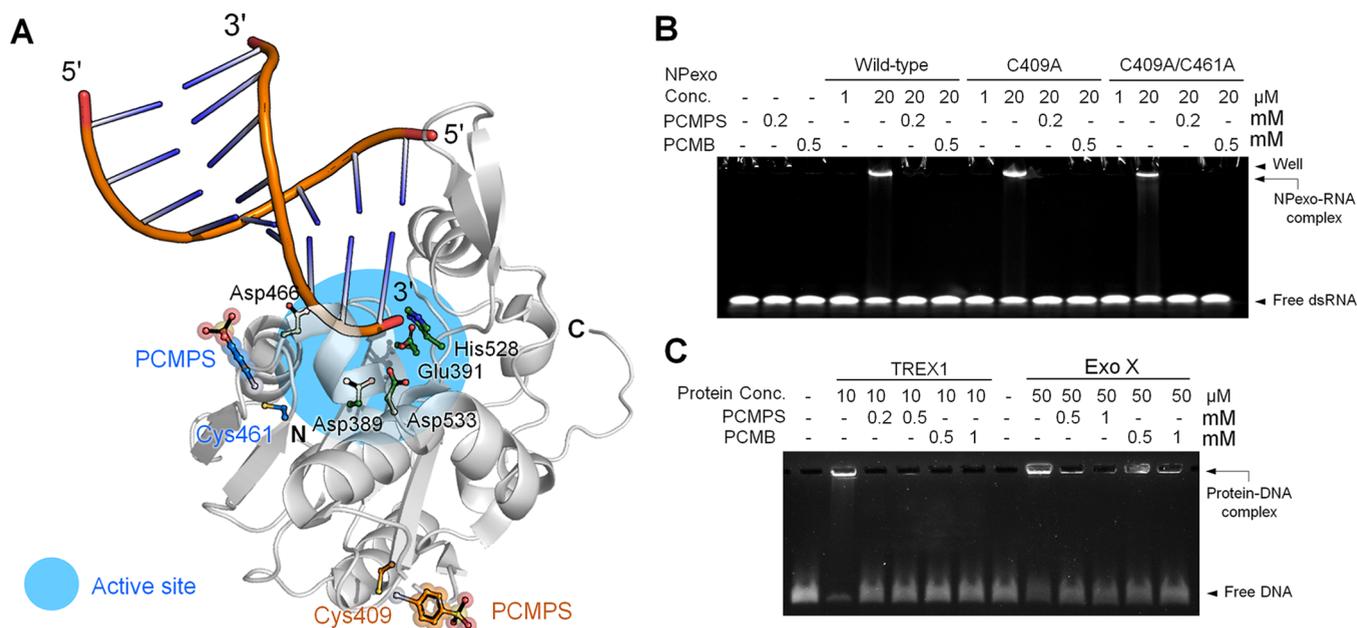


Figure 7. Inhibitor coupled DNA/RNA binding assays. (A) Structural superposition of the NPexo-dsRNA complex (PDB accession code 3MWP), the NPexo-PCMPS complex, and the NPexo C409A-PCMPS complex. The structure of the NPexo-dsRNA complex is colored light gray. PCMPS and the attached cysteine in the NPexo-PCMPS complex and in the NPexo C409A-PCMPS complex are colored orange and blue, respectively. PCMPS does not overlap or have contact with dsRNA. (B) Substrate binding of 20 μ M wild-type and mutated NPexo is abolished by PCMPS or PCMB as the bands of the protein–RNA complex disappear in their presence. (C) Substrate binding of 10 μ M TREX1 and 50 μ M Exo X is reduced by PCMPS or PCMB. The bands of the protein–DNA complex disappear or are lowered in the presence of PCMPS and PCMB. Capital C stands for the control group.

affect the RNA-stacking lid and in turn its ability to recruit and couple to substrates. In several independent MD simulations, Arg556 shifting to interact with PCMPS is consistently observed.

NPexo C409A-PCMPS Simulation Reveals an Alternative Long-Range Effect

After mutating Cys409 to alanine, PCMPS targets Cys461 as illustrated in the newly solved apo-NPexo C409A and NPexo C409A-PCMPS structures. These X-ray coordinates were used to initiate the apo-NPexo C409A and NPexo C409A-PCMPS simulations to reveal the impact of PCMPS attaching to Cys461. In the apo-NPexo C409A simulation, the β -core and the active site exhibit small RMSD as they do in the apo-NPexo simulation. Behaviors of the RNA-stacking region and the general base containing loop are also similar to those of the wildtype (Figures S12B and S13A and Movie S3). In the NPexo C409A-PCMPS simulation, although the β -core remains stable, the DEDDh cluster and α 5 (Ser464-Gln474) exhibit significantly larger RMSDs (Figure S13B). The RNA-stacking lid and the general-base containing loop also have high RMSD values with a significant conformational shift occurring around 6.7 μ s of the NPexo C409A-PCMPS trajectory.

Cys461 locates in a strand at the N-terminal side of α 5 that contains the active site residue Asp466 and the RNA contacting residue Asp465. In the NPexo C409A-PCMPS simulation, the PCMPS sulfonate interacts with the side chain of Arg468 in α 5, causing the helix to shift (Figure 6A, C and Movie S4). Since Asp466 in the active site is also in α 5, the DEDDh cluster has much higher RMSD in the NPexo C409A-PCMPS simulation comparing to that in the apo-NPexo C409A simulation. Such active site deformation (Figure 6A, C) due to Asp466 moving toward PCMPS can be further

illustrated by the longer distance between the Asp466 C α and that of Asp389, another DEDDh residue (Figure 6A). In addition to perturbing the structure of the DEDDh cluster, the PCMPS induced variation can also propagate to affect substrate binding because the Cys461 containing α 5 is in contact with the loop between β 1 and β 2 that is next to α 2 in the RNA stacking lid. Indeed, the conformational shift starting at 6.7 μ s of the NPexo C409A-PCMPS simulation involves α 5, the DEDDh cluster, the RNA stacking lid, and the general base containing loop (Figure S13B). The aforementioned hydrogen bonding between Ser430 (α 2) in the RNA stacking region and Glu391 in the active site is also lost upon this conformational shift, whereas this interaction is well maintained in the apo-NPexo C409A simulation (Figure 6B, C). Compared to the wildtype case of targeting C409, the long-range effects of PCMPS ligating to Cys461 follows a distinct route, but in both cases, the impact of local perturbation can propagate to the RNA stacking lid of NPexo and affect the conformation therein.

Testing the Mechanism of TCI-Induced Allosteric Inhibition

After ligating to NPexo, PCMPS electrostatically interacts with the surrounding Arg residues and hence their original couplings to other amino acids in the protein interaction network. Concomitantly, the domino-like propagation of local perturbation can lead to changes in the RNA-binding lid conformation as indicated by all-atom MD simulations. A key prediction of this allosteric inhibition is that substrate binding would be affected by PCMPS as supported by the cases of CRN-4 and RNase T DEDDh family members.¹⁴ To test this mechanism on NPexo, we used inhibitor-coupled DNA/RNA binding assays to measure the impact of TCI on substrate binding.¹⁴

Figure 7B shows that in the presence of 0.2 mM PCMPS or 0.5 mM PCMB, the band of NPexo-RNA complexes disappears. Clearly, the organomercurial attachment decreases the ability of the wild-type and the mutant NPexos in binding RNA as suggested by the mechanism of allosteric inhibition. PCMPS and PCMB reducing substrate binding in DEDDh exonucleases was also tested and confirmed for TREX1 and Exo X (Figure 7C). Our binding assay measurements show that PCMPS and PCMB inhibit the substrate binding in DEDDh exonucleases.

DISCUSSION AND CONCLUSION

Our demonstration of the broad inhibition on DEDDh exonucleases by PCMPS and PCMB, including TREX1, SARS-CoV-2 nsp14, and NPexo, illustrates the evident potential of TCIs in antiviral and/or anticancer therapies. The first TCI-complexed structures for a DEDDh family member were resolved here and uncover that the ligated cysteine residues are distal to the active and substrate-binding sites. Mutating the first binding site Cys409 in NPexo leads to PCMPS or PCMB targeting Cys461 instead. Switching of the ligated residues is also observed in the PMB bound hemoglobin.⁷ Such a sequential preference of TCI targeting alternative cysteine residues shows the existence of multiple pathways for exonuclease inhibition, which is a potentially very useful property in therapeutic applications.

With PCMPS or PCMB not blocking the access to the active and substrate-binding sites, how TCI binding at a distal residue inhibits the exonuclease activity is via a missing mechanistic connection between the structural and activity data. The NPexo all-atom MD simulations conducted here elucidate the diverse flexibilities of different structural motifs in an aqueous environment and the domino propagation of the TCI-induced local perturbation to affect the RNA-binding lid conformation. It has been anticipated that the sulfonate moiety of PCMPS might affect the protein function by interacting with nearby amino acids,^{34–36} and our MD simulations and analysis provide a first illustration for the long-range propagation of such local perturbation. This allosteric inhibition mechanism by TCI is consistent with a number of experimental observations. The prediction of PCMPS or PCMB affecting the RNA binding in NPexo is confirmed by our inhibitor-coupled DNA/RNA binding assays. The behaviors of organomercurial enzyme inhibition being distinct to that of HgCl₂³⁴ are also in line with the primary role coming from the TCI sulfonate but not the mercury atom as revealed by our all-atom MD simulations. Another prediction is that the inhibition capability of PCMPS would be stronger than that of PCMB for its larger and more extended sulfonate, which agrees with our activity measurements.

The allostery of TCI ligation also calls for caution in its usage of probing protein functional properties, because the observed changes may not reflect the intrinsic behaviors of the target. For example, using PCMPS to study the importance of cysteine residues in protein oligomerization^{7,9} is not recommended because TCI could bind a residue at the protein–protein interaction zone.³⁷ Even if the ligation site is not at the interface, the potential allosteric effects of TCI may modulate the conformation of the structural motif responsible for oligomerization. Similarly, using an organomercurial in combination with mutagenesis to identify the important residues for substrate binding and catalysis likely reveals the

allosteric effects of the TCI instead of the residue's intrinsic functional property, and ultracautious is hence recommended.

EXPERIMENTAL SECTION

Protein Expression and Purification

Recombinant RNase T, Exo X, TREX1, and TREX2 were prepared as described in earlier works.^{14,38,39} The gene of SARS-CoV-2 nsp14 was cloned into the NdeI/XhoI sites of pET-28a (Novagen, USA). SARS-CoV-2 nsp14 protein was expressed in the *E. coli* BL21-CodonPlus (DE3)-RIPL strain and sequentially purified by cobalt-based chromatography column (HiTrap TALON crude 5 mL, GE Healthcare), HiTrap Q column (GE Healthcare, USA), HiTrap SP column (GE Healthcare, USA), and gel filtration chromatography column (Superdex 75, GE Healthcare, USA). Purified nsp14 was stored in 50 mM Tris, pH 7, 300 mM NaCl. The exonuclease domain (342–569 a.a.) of NPexo from *Lassa fever Virus* was cloned into expression vectors pET-28c (Novagen, USA) to generate the N-terminal His-tagged fusion protein. The plasmid was transformed into *E. coli* Tuner (DE3) strain cultured in LB medium supplemented with 35 µg/mL Kanamycin. Cells were grown to an O.D.600 of 0.5–0.6 and induced by 1 mM IPTG at 18 °C for 18 h. The harvested cells were lysed through sonication and passed through a Ni-NTA resin affinity column (Qiagen Inc., USA) or a cobalt-based chromatography column (HiTrap TALON crude 5 mL, GE Healthcare) using a standard protocol. The protein sample was further purified by a HiTrap Heparin column (GE Healthcare, USA) and a gel filtration chromatography column (Superdex 200, GE Healthcare, USA). Purified NPexo samples were concentrated to 23–30 mg/mL in 300 mM NaCl and 50 mM Tris-HCl (pH 7.0). Mutated NPexos were purified by the same procedure.

Nuclease Activity Assays

The sequences of DNA or RNA substrates are listed in Table S1. DNA/RNA substrates were labeled with FAM at the 5'-end by MDBio, Inc., Taiwan. Inhibitors including PCMPS (C367750, Toronto Research Chemicals) and PCMB (55540, Sigma-Aldrich) were mixed with target proteins at room temperature for 10 min. The labeled substrates (0.5 µM) were incubated with target proteins or target protein-inhibitor mixture at 37 °C at 120 mM NaCl, 20 mM Tris-HCl, pH 7.0, and 2 mM MgCl₂ for 30–60 min. The protein concentrations used in the nuclease activity are listed as following: RNase T (0.5 µM), ExoX (0.2 µM), TREX1 (0.1 µM), TREX2 (0.2 µM), NPexo (2 µM), and nsp14 (4 µM). The reaction was stopped by adding 2× TBE-urea sample loading buffer (G-Biosciences, USA) at 95 °C for 5 min. The DNA digestion patterns were resolved on 20% denaturing polyacrylamide gels and visualized by ultraviolet light.

Crystallization and Crystal Structural Determination

Wild-type and mutated NPexos (23–30 mg/mL) were mixed with inhibitors for 10 min before crystallization. Detailed information regarding crystallization conditions of the six structures resolved here are shown in Table S2. All crystals were cryoprotected by Paratone-N (Hampton Research, USA) for data collection at BL13C1, BL13B1, BL15A1, and TPS05A in NSRRRC, Taiwan. All diffraction data were processed by HKL2000, and diffraction statistics are listed in Table 1. Structures were solved by the molecular replacement method (MR) or single-wavelength anomalous diffraction (SAD). The crystal structure of *Lassa* nucleoprotein with dsRNA (PDB: 4G9Z) was used as the search model for MR method, and phasing was done by MOLREP of CCP4.⁴⁰ The programs SHELX C/D/E, Phaser and DM in CCP4⁴⁰ or AutoSol of PHENIX 1.13_2998 were used for SAD phasing. Only one Hg site was identified and used in solving the phase problem. The FOM value is higher than 0.3 in Phaser and higher than 0.6 in DM. The models were built by Coot-0.8.1 and refined by PHENIX-1.9–1692.⁴¹ Diffraction structure factors and structural coordinates were deposited in the RCSB Protein Data Bank with the PDB ID code of 7V37 (apo-NPexo), 7V38 (NPexo-PCMPS complex), 7V39 (NPexo-PCMB complex), 7V3A (apo-NPexo

C409A), 7V3B (NPexo C409A-PCMPS complex), and 7V3C (NPexo C409A-PCMB complex).

Organomercurials can be used as additives in X-ray crystallography to improve crystal quality,⁴² to promote growth of highly ordered crystals,⁴³ or to be used as heavy-atom references in the soaking methods for solving the phasing problem⁴⁴ by multiwavelength anomalous diffraction.⁴⁵ The structural analysis conducted in work indeed demonstrates that PCMPS and PCMB have sufficient power to solve the phasing problem for the NPexo protein of 250 amino acids by single-anomalous diffraction. Furthermore, superposition of our wild-type and PCMPS- or PCMB-bound structures of NPexo illustrates that formation of the mercury–sulfur bond does not affect protein structures and only perturb local conformation in the crystalline environment. Applying these organomercurials in crystal growth or for solving the phase problem thus appears to be a reasonable strategy concerning the structure of a monomeric protein.

Molecular Dynamics (MD) Simulation

The all-atom models of apo-NPexo, NPexo-PCMPS, apo-NPexo C409A, and NPexo-C409A-PCMPS simulations were developed from the 7 V37, 7 V38, 7 V3A, and 7 V3B X-ray structures resolved in this work. The CHARMM36 all-atom protein force field⁴⁶ was used to describe the molecular interaction energetics. The molecular mechanics force field of PCMPS-modified cysteine side chain was constructed following the CHARMM parametrization protocol⁴⁷ using the ftk toolkit⁴⁸ implemented in VMD.⁴⁹ The details of this parametrization and the resulting force field is reported in the [Supplementary Methods](#). Each protein system is solvated in a cubic box of TIP3P water⁵⁰ with counterions at the 1.5 mM ionic strength. After 6 ns of heating and equilibration, the production run is conducted at 310 K and 1.103 bar for 10 μ s. The sampled configurations are saved every 1 ps for analysis. The other details of the all-atom MD simulations are reported in the [Supplementary Methods](#), [Table S4](#), and [Figures S10 and S11](#).

DNA/RNA-Binding Assays

The FAM-labeled double-stranded 20-bp RNA or 708 bp-long PCR product were used as binding substrates of various DEDDh exonucleases, including NPexo (20 μ M), TREX1 (10 μ M) and Exo X (50 μ M). For each DEDDh exonuclease, mixing with an inhibitor was conducted at room temperature for 10 min. A DNA/RNA substrate was then incubated with an exonuclease/inhibitor mixture in a buffer containing 120 mM NaCl, 20 mM Tris-HCl, pH 8.0 and 10 mM EDTA, at room temperature for 30 min. After incubation, the reaction mixture was applied to 20% TBE gels or agarose gels for visualization by ultraviolet light.

Liquid Chromatography–Mass Spectrometry Analysis

To track the PCMPS reaction evolution with NPexo, we prepared the different time-dependent binding reactions by mixing 50 μ L of 500 μ M purified wild-type or mutated NPexo with 5 μ L of 10 mM PCMPS into 2 mL of 120 mM NaCl, 20 mM Tris-HCl, pH 7.0, 2 mM MgCl₂. The protein–ligand mixtures were then desalted to remove the free PCMPS and salt by discontinuous diafiltration using the Amicon Ultra-15 centrifugal filters (Millipore). The desalted samples were dried by a vacuum concentrator and then resuspended with 50% acetonitriles and 2% formic acid for further LC-MS analysis.

An UltiMate 3000 UHPLC (Dionex) coupled with an Impact HD Q-TOF mass spectrometer (Bruker) was used for LC-MS analysis. The chromatographic separation was carried out on a Thermo Acclaim C18 column (2.1 \times 100 mm, 2.2 μ m) with elution gradient program of 20% to 100% acetonitrile in 20 min. The eluted samples were then protonated by the positive ion mode of electrospray ionization source (ESI+) in mass spectrometer. The mass spectra were collected over the mass range of m/z 50–1500 at a resolving power of 40 000. The collected data were processed and analyzed by Compass Data Analysis 4.1 (Bruker, Germany). The ESI Compass Maximum Entropy deconvolution algorithm over the mass range of 20 000 to 30 000 Da was used for neutral mass spectra analysis.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacsau.1c00420>.

Supplementary methods, Tables S1–S5, Figures S1–S13 (PDF)

Movie S1, apo-NPexo simulation (MP4)

Movie S2, NPexo-PCMPS simulation (MP4)

Movie S3, apo-NPexo C409A simulation (MP4)

Movie S4, NPexo C409A-PCMPS simulation (MP4)

Accession Codes

Coordinates and structure factors were deposited in Protein Data Bank under accession code 7V37 (apo-NPexo), 7V38 (NPexo-PCMPS complex), 7V39 (NPexo-PCMB complex), 7V3A (apo-NPexo C409A), 7V3B (NPexo C409A-PCMPS complex), and 7V3C (NPexo C409A-PCMB complex).

AUTHOR INFORMATION

Corresponding Authors

Jhih-Wei Chu – Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; Institute of Molecular Medicine and Bioengineering, Institute of Bioinformatics and Systems Biology, and Center for Intelligent Drug Systems and Smart Bio-devices (IDS2B), National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan; Email: jwchu@nctu.edu.tw

Yu-Yuan Hsiao – Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; Institute of Molecular Medicine and Bioengineering, Institute of Bioinformatics and Systems Biology, and Center for Intelligent Drug Systems and Smart Bio-devices (IDS2B), National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan; Drug Development and Value Creation Research Center, Center for Cancer Research, Kaohsiung Medical University, Kaohsiung 80708, Taiwan; orcid.org/0000-0002-9648-5879; Email: mike0617@nctu.edu.tw

Authors

Kuan-Wei Huang – Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; Institute of Molecular Medicine and Bioengineering, National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan

Jing-Wen Chen – Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; Institute of Molecular Medicine and Bioengineering and Institute of Bioinformatics and Systems Biology, National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan; orcid.org/0000-0003-2043-7253

Tzu-Yu Hua – Institute of Bioinformatics and Systems Biology, National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan

Yu-Yu Chu – Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; Institute of Molecular Medicine and Bioengineering and Institute of Bioinformatics and Systems Biology, National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan

Tsai-Yuan Chiu – Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; Institute of Molecular Medicine and Bioengineering, National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan

Jung-Yu Liu – Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; Institute of Molecular Medicine and Bioengineering, National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan

Chun-I Tu – Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; Institute of Molecular Medicine and Bioengineering, National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan

Kai-Cheng Hsu – Graduate Institute of Cancer Biology and Drug Discovery, College of Medical Science and Technology, Taipei Medical University, Taipei 11031, Taiwan; Ph.D. Program for Cancer Molecular Biology and Drug Discovery, College of Medical Science and Technology, Ph.D. Program in Biotechnology Research and Development, College of Pharmacy, and Biomedical Commercialization Center, Taipei Medical University, Taipei 11031, Taiwan; orcid.org/0000-0002-9022-6673

Ya-Ting Kao – Department of Biological Science and Technology, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; Institute of Molecular Medicine and Bioengineering, Institute of Bioinformatics and Systems Biology, and Center for Intelligent Drug Systems and Smart Bio-devices (IDS2B), National Yang Ming Chiao Tung University, Hsinchu 30068, Taiwan

Complete contact information is available at:
<https://pubs.acs.org/10.1021/jacsau.1c00420>

Author Contributions

[†]K.-W.H. and J.-W.C. contributed equally.

Funding

This work is financially supported by the Ministry of Science and Technology (MOST) of Taiwan under grant 109-2113-M-009-023- to J.-W.C. and under the Excellent Youth Scholar Grant and Young Scholar Fellowship Program (Columbus Program) though grants 108-2636-B-009-004, 109-2636-B-009-004, and 110-2636-B-009-009 to Y.-Y. H.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Portions of this research were performed at the National Synchrotron Radiation Research Center (BL-13B1, BL-13C1, and TPS-05A), a national user facility supported by the National Science Council of Taiwan, ROC. The Synchrotron Radiation Protein Crystallography Facility is supported by the National Core Facility Program for Biotechnology. This work was financially supported by the “Center For Intelligent Drug Systems and Smart Biodevices (IDS²B)” and “Smart Platform of Dynamic Systems Biology for Therapeutic Development” project from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan. This study is supported partially by NCTU-KMU JOINT RESEARCH PROJECT, Kaohsiung Medical University (NCTUKMU108-

DR-01). Mass spectrometric analysis were performed by the Center for Advanced Instrumentation and Department of Applied Chemistry at National Yang Ming Chiao Tung University, Hsinchu, Taiwan, R.O.C. We thank Dr. Sheng-Cih Huang (Mass Laboratory, National Yang Ming Chiao-Tung University, Hsinchu) for mass spectrometry analysis.

ABBREVIATIONS

NPexo, NP exonuclease; PCMPs, *p*-chloromercuriphenyl sulfate; PCMB or PHMB, *p*-chloromercuribenzoate; PMB, *p*-hydroxymercuribenzoate; MD, molecular dynamics; SAD, single-wavelength anomalous dispersion; RMSD, root-mean-square difference; ER-LBD, estrogen receptor ligand binding domain

REFERENCES

- (1) Sutanto, F.; Konstantinidou, M.; Domling, A. Covalent inhibitors: a rational approach to drug discovery. *RSC Med. Chem.* **2020**, *11* (8), 876–884.
- (2) Janes, M. R.; Zhang, J.; Li, L. S.; Hansen, R.; Peters, U.; Guo, X.; Chen, Y.; Babbar, A.; Firdaus, S. J.; Darjania, L.; Feng, J.; Chen, J. H.; Li, S.; Li, S.; Long, Y. O.; Thach, C.; Liu, Y.; Zarihi, A.; Ely, T.; Kucharski, J. M.; Kessler, L. V.; Wu, T.; Yu, K.; Wang, Y.; Yao, Y.; Deng, X.; Zarrinkar, P. P.; Brehmer, D.; Dhanak, D.; Lorenzi, M. V.; Hu-Lowe, D.; Patricelli, M. P.; Ren, P.; Liu, Y. Targeting KRAS Mutant Cancers with a Covalent G12C-Specific Inhibitor. *Cell* **2018**, *172* (3), 578–589.
- (3) Ghosh, A. K.; Samanta, I.; Mondal, A.; Liu, W. R. Covalent Inhibition in Drug Discovery. *ChemMedChem* **2019**, *14* (9), 889–906.
- (4) Vasudevan, A.; Argiriadi, M. A.; Baranczak, A.; Friedman, M. M.; Gavriluk, J.; Hobson, A. D.; Hulce, J. J.; Osman, S.; Wilson, N. S. Covalent binders in drug discovery. *Prog. Med. Chem.* **2019**, *58*, 1–62.
- (5) Park, C. S.; Doh, P. S.; Lee, C. J.; Han, D. S.; Carraway, R. E.; Miller, T. B. Cellular mechanism of stimulation of renin secretion by the mercurial diuretic mersalyl. *J. Pharmacol. Exp. Ther.* **1991**, *257* (1), 219–24.
- (6) Sen, S. S.; Bhuyan, N. R.; Lakshman, K.; Roy, A. K.; Chakraborty, B.; Bera, T. Membrane bound pyrophosphatase and P-type adenosine triphosphatase of *Leishmania donovani* as possible chemotherapeutic targets: similarities and differences in inhibitor sensitivities. *Biochemistry* **2009**, *74* (12), 1382–7.
- (7) Kan, H. I.; Chen, I. Y.; Zulfajri, M.; Wang, C. C. Subunit disassembly pathway of human hemoglobin revealing the site-specific role of its cysteine residues. *J. Phys. Chem. B* **2013**, *117* (34), 9831–9.
- (8) Fraaije, M. W.; Mattevi, A.; van Berkel, W. J. Mercuriation of vanillyl-alcohol oxidase from *Penicillium simplicissimum* generates inactive dimers. *FEBS Lett.* **1997**, *402* (1), 33–5.
- (9) Murphy, D. M.; Ivanenkov, V. V.; Kirley, T. L. Identification of cysteine residues responsible for oxidative cross-linking and chemical inhibition of human nucleoside-triphosphate diphosphohydrolase 3. *J. Biol. Chem.* **2002**, *277* (8), 6162–9.
- (10) Behnke, C. A.; Le Trong, I.; Godden, J. W.; Merritt, E. A.; Teller, D. C.; Bajorath, J.; Stenkamp, R. E. Atomic resolution studies of carbonic anhydrase II. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66* (5), 616–627.
- (11) Kawabata, S.; Davie, E. W. A microsomal endopeptidase from liver with substrate specificity for processing proproteins such as the vitamin K-dependent proteins of plasma. *J. Biol. Chem.* **1992**, *267* (15), 10331–6.
- (12) Mandel, R.; Ryser, H. J.; Ghani, F.; Wu, M.; Peak, D. Inhibition of a reductive function of the plasma membrane by bacitracin and antibodies against protein disulfide-isomerase. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (9), 4112–6.
- (13) Mitra, S.; Nath, D.; Majumder, G. C. Purification and characterization of a protein kinase from goat sperm plasma membrane. *Biochem. Cell Biol.* **1994**, *72* (5–6), 218–26.

- (14) Huang, K. W.; Hsu, K. C.; Chu, L. Y.; Yang, J. M.; Yuan, H. S.; Hsiao, Y. Y. Identification of Inhibitors for the DEDDh Family of Exonucleases and a Unique Inhibition Mechanism by Crystal Structure Analysis of CRN-4 Bound with 2-Morpholin-4-ylethanesulfonate (MES). *J. Med. Chem.* **2016**, *59* (17), 8019–29.
- (15) Prentice, D. A.; Kitos, P. A.; Gurley, L. R. Effects of phosphatase inhibitors on nuclease activity. *Cell Biol. Int. Rep.* **1985**, *9* (11), 1027–34.
- (16) Ackerman, S. J.; Liu, L.; Kwatia, M. A.; Savage, M. P.; Leonidas, D. D.; Swaminathan, G. J.; Acharya, K. R. Charcot-Leyden crystal protein (galectin-10) is not a dual function galectin with lysophospholipase activity but binds a lysophospholipase inhibitor in a novel structural fashion. *J. Biol. Chem.* **2002**, *277* (17), 14859–68.
- (17) DuBay, K. H.; Iwan, K.; Osorio-Planes, L.; Geissler, P. L.; Groll, M.; Trauner, D.; Broichhagen, J. A predictive approach for the optical control of carbonic anhydrase II activity. *ACS Chem. Biol.* **2018**, *13*, 793.
- (18) Mikulski, R.; Domsic, J. F.; Ling, G.; Tu, C.; Robbins, A. H.; Silverman, D. N.; McKenna, R. Structure and catalysis by carbonic anhydrase II: role of active-site tryptophan 5. *Arch. Biochem. Biophys.* **2011**, *516* (2), 97–102.
- (19) Wolfgang, C. L.; Zhang, Z.; Gabriel, J. L.; Pieringer, R. A.; Soprano, K. J.; Soprano, D. R. Identification of sulfhydryl-modified cysteine residues in the ligand binding pocket of retinoic acid receptor beta. *J. Biol. Chem.* **1997**, *272* (2), 746–53.
- (20) Zuo, Y.; Deutscher, M. P. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic acids research* **2001**, *29* (5), 1017–26.
- (21) Vanpouille-Box, C.; Alard, A.; Aryankalayil, M. J.; Sarfraz, Y.; Diamond, J. M.; Schneider, R. J.; Inghirami, G.; Coleman, C. N.; Formenti, S. C.; Demaria, S. DNA exonuclease Trex1 regulates radiotherapy-induced tumour immunogenicity. *Nat. Commun.* **2017**, *8*, 15618.
- (22) Vanpouille-Box, C.; Formenti, S. C.; Demaria, S. TREX1 dictates the immune fate of irradiated cancer cells. *Oncoimmunology* **2017**, *6* (9), No. e1339857.
- (23) Qi, X.; Lan, S.; Wang, W.; Schelde, L. M.; Dong, H.; Wallat, G. D.; Ly, H.; Liang, Y.; Dong, C. Cap binding and immune evasion revealed by Lassa nucleoprotein structure. *Nature* **2010**, *468* (7325), 779–83.
- (24) Hastie, K. M.; King, L. B.; Zandonatti, M. A.; Saphire, E. O. Structural basis for the dsRNA specificity of the Lassa virus NP exonuclease. *PLoS One* **2012**, *7* (8), No. e44211.
- (25) Hastie, K. M.; Kimberlin, C. R.; Zandonatti, M. A.; MacRae, I. J.; Saphire, E. O. Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3' to 5' exonuclease activity essential for immune suppression. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (6), 2396–401.
- (26) Ilori, E. A.; Frank, C.; Dan-Nwafor, C. C.; Ipadeola, O.; Krings, A.; Ukponu, W.; Womi-Eteng, O. E.; Adeyemo, A.; Mutbam, S. K.; Musa, E. O.; Lasuba, C. L. P.; Alemu, W.; Okogbenin, S.; Ogbaini, E.; Unigwe, U.; Ogah, E.; Onoh, R.; Abejegah, C.; Ayodeji, O.; Ihekweazu, C. Increase in Lassa Fever Cases in Nigeria, January–March 2018. *Emerging Infect. Dis.* **2019**, *25* (5), 1026–1027.
- (27) Minskaia, E.; Hertzog, T.; Gorbalenya, A. E.; Campanacci, V.; Cambillau, C.; Canard, B.; Ziebuhr, J. Discovery of an RNA virus 3'–>5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (13), 5108–13.
- (28) Graham, R. L.; Becker, M. M.; Eckerle, L. D.; Bolles, M.; Denison, M. R.; Baric, R. S. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. *Nat. Med.* **2012**, *18* (12), 1820–6.
- (29) Ogando, N. S.; Zevenhoven-Dobbe, J. C.; van der Meer, Y.; Bredenbeek, P. J.; Posthuma, C. C.; Snijder, E. J. The Enzymatic Activity of the nsp14 Exoribonuclease Is Critical for Replication of MERS-CoV and SARS-CoV-2. *J. Virol.* **2020**, *94* (23), e01246-20.
- (30) Robson, F.; Khan, K. S.; Le, T. K.; Paris, C.; Demirbag, S.; Barfuss, P.; Rocchi, P.; Ng, W. L. Coronavirus RNA Proofreading: Molecular Basis and Therapeutic Targeting. *Mol. Cell* **2020**, *79* (5), 710–727.
- (31) Smith, E. C.; Blanc, H.; Vignuzzi, M.; Denison, M. R. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLoS Pathog.* **2013**, *9* (8), No. e1003565.
- (32) Chen, Y.; Cai, H.; Pan, J.; Xiang, N.; Tien, P.; Ahola, T.; Guo, D. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (9), 3484–9.
- (33) Yang, W. Nucleases: diversity of structure, function and mechanism. *Q. Rev. Biophys.* **2011**, *44* (1), 1–93.
- (34) Schwenker, M. M. Mercuric chloride and p-chloromercuriphenylsulfonate exert a biphasic effect on the binding of the stimulant [3H]methylphenidate to the dopamine transporter. *Synapse (Hoboken, NJ, U. S.)* **1994**, *16* (3), 188–94.
- (35) Santos, K. L.; Vento, M. A.; Wright, J. W.; Speth, R. C. The effects of para-chloromercuribenzoic acid and different oxidative and sulfhydryl agents on a novel, non-AT1, non-AT2 angiotensin binding site identified as neurolysin. *Regul. Pept.* **2013**, *184*, 104–14.
- (36) Jayaraman, A.; Thandeeswaran, M.; Priyadarsini, U.; Sabarathinam, S.; Nawaz, K. A.; Palaniswamy, M. Characterization of unexplored amidohydrolase enzyme-pterin deaminase. *Appl. Microbiol. Biotechnol.* **2016**, *100* (11), 4779–89.
- (37) Pike, A. C.; Brzozowski, A. M.; Walton, J.; Hubbard, R. E.; Thorsell, A. G.; Li, Y. L.; Gustafsson, J. A.; Carlquist, M. Structural insights into the mode of action of a pure antiestrogen. *Structure (Oxford, U. K.)* **2001**, *9* (2), 145–53.
- (38) Huang, K. W.; Liu, T. C.; Liang, R. Y.; Chu, L. Y.; Cheng, H. L.; Chu, J. W.; Hsiao, Y. Y. Structural basis for overhang excision and terminal unwinding of DNA duplexes by TREX1. *PLoS Biol.* **2018**, *16* (5), No. e2005653.
- (39) Cheng, H. L.; Lin, C. T.; Huang, K. W.; Wang, S.; Lin, Y. T.; Toh, S. I.; Hsiao, Y. Y. Structural insights into the duplex DNA processing of TREX2. *Nucleic Acids Res.* **2018**, *46* (22), 12166–12176.
- (40) Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59* (7), 1131–7.
- (41) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkoczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L. W.; Jain, S.; McCoy, A. J.; et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **2019**, *75* (10), 861–877.
- (42) Menchise, V.; De Simone, G.; Alterio, V.; Di Fiore, A.; Pedone, C.; Scozzafava, A.; Supuran, C. T. Carbonic anhydrase inhibitors: stacking with Phe131 determines active site binding region of inhibitors as exemplified by the X-ray crystal structure of a membrane-impermeant antitumor sulfonamide complexed with isozyme II. *J. Med. Chem.* **2005**, *48* (18), 5721–7.
- (43) Guerri, A.; Briganti, F.; Scozzafava, A.; Supuran, C. T.; Mangani, S. Mechanism of cyanamide hydration catalyzed by carbonic anhydrase II suggested by cryogenic X-ray diffraction. *Biochemistry* **2000**, *39* (40), 12391–7.
- (44) Birck, C.; Poch, O.; Romier, C.; Ruff, M.; Mengus, G.; Lavigne, A. C.; Davidson, I.; Moras, D. Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. *Cell* **1998**, *94* (2), 239–49.
- (45) Jin, X.; Ballicora, M. A.; Preiss, J.; Geiger, J. H. Crystal structure of potato tuber ADP-glucose pyrophosphorylase. *EMBO J.* **2005**, *24* (4), 694–704.
- (46) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. J. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8* (9), 3257–3273.
- (47) Foloppe, N.; MacKerell, A. D., Jr. All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **2000**, *21* (2), 86–104.

(48) Mayne, C. G.; Saam, J.; Schulten, K.; Tajkhorshid, E.; Gumbart, J. C. Rapid parameterization of small molecules using the force field toolkit. *J. Comput. Chem.* **2013**, *34* (32), 2757–2770.

(49) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14* (1), 33–38.

(50) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.