



Targeting the C-Terminal Domain Small Phosphatase 1

Harikrishna Reddy Rallabandi, Palanivel Ganesan^D and Young Jun Kim *^D

Department of Medicinal Biosciences and Nanotechnology Research Center, Konkuk University, Chungju 27478, Korea; harikrishna@kku.ac.kr (H.R.R.); palanivel67@gmail.com (P.G.)

* Correspondence: ykim@kku.ac.kr; Tel.: +82-043-840-3569

Received: 29 March 2020; Accepted: 7 May 2020; Published: 8 May 2020



Abstract: The human C-terminal domain small phosphatase 1 (CTDSP1/SCP1) is a protein phosphatase with a conserved catalytic site of DXDXT/V. CTDSP1's major activity has been identified as dephosphorylation of the 5th Ser residue of the tandem heptad repeat of the RNA polymerase II C-terminal domain (RNAP II CTD). It is also implicated in various pivotal biological activities, such as acting as a driving factor in repressor element 1 (RE-1)-silencing transcription factor (REST) complex, which silences the neuronal genes in non-neuronal cells, G1/S phase transition, and osteoblast differentiation. Recent findings have denoted that negative regulation of CTDSP1 results in suppression of cancer invasion in neuroglioma cells. Several researchers have focused on the development of regulating materials of CTDSP1, due to the significant roles it has in various biological activities. In this review, we focused on this emerging target and explored the biological significance, challenges, and opportunities in targeting CTDSP1 from a drug designing perspective.

Keywords: CTDSP1; drug design; allosteric docking; ensemble docking

1. Introduction

RNA polymerase II (RNAP II) is the crucial component of a transcription apparatus, and orchestrates the post-transcriptional regulation and modification of mRNA. The largest subunit of RNAP II is the C-terminal domain (CTD), which is essential for the recruitment and assembly of transcription complexes. CTD contains consensus tandem heptapeptide $(Y_1S_2P_3T_4S_5P_6S_7)$ repeats, which vary from 26 to 52 from yeast to humans, respectively [1,2]. CTD phosphatase 1 (CTDP1/FCP1) is the first member of a C-terminal domain phosphatase (CTDP) family, which was identified by dephosphorylating the Ser2 of heptapeptide [3,4]. Later studies revealed that CTD small phosphatase 1 (CTDSP1/SCP1) preferentially reverses the phosphorylation of the Ser5 in consensus peptide repeats of RNAP II CTD [5,6]. CTD small phosphatase 2 (CTDSP2/SCP2) and CTD small phosphatase-like (CTDSP1/SCP3) were identified as structurally and functionally closer paralogs of CTDSP1 [5,7]. CTDSP1 hydrolyzes the phosphoryl group on the Ser5 of CTD through the conserved catalytic site of DXDXT/V and Mg²⁺ ions, which is highly conserved throughout CTD small phosphatases (CTDSP).

The catalytic activity of CTDSP1 extends its role into various vital biological activities, and has made it an anticipated drug target [8]. CTDSP1's novel signature mechanism is essential for numerous signaling pathways and cellular activities, such as neuronal gene silencing, cell cycle regulation, and regulation of certain cell signal transductions. Repressor element 1 (RE-1)-silencing transcription factor/neuron-restrictive silencer factor (REST/NRSF) complex is responsible for silencing the neuronal genes, and some studies have illustrated that CTDSP1 is co-precipitated with the REST complex in non-neuronal cells [9,10]. A recent study suggested that CTDSP1 controls the stability of the REST complex through dephosphorylation [11]. Phosphorylated receptor-regulated small mothers against decapentaplegic (SMAD) proteins (R-SMADs) play a significant role in the orchestration of

transforming growth factor β (TGF β) signal transduction and bone morphogenetic protein (BMP) signaling. The R-SMADs could be dephosphorylated by the nuclear CTDSP1 and attenuate the BMP signal at the required level [12]. Further studies have demonstrated CTDSP1's role in cell cycle arrest at the G1/S transition phase by C-Myc mediated dephosphorylation of the retinoblastoma protein (RB) [13,14]. A new study has suggested that cell division associated 3 protein (CDCA3) could also be an enzymatic substrate for CTDSP1 [15]. Apart from these known functionalities, a recent study related to cancer invasion exhibited the exciting results that CTDSP1 dephosphorylates Twist-related protein (TWIST) and regulates cancer cell migration [16]. Given all these characteristics, we are seeking to prove that CTDSP1 can be considered as a potential drug target.

There are a few challenges in targeting CTDSP1 despite its potential. In previous studies [17–19], researchers tried to identify chemical inhibitors targeting human CTDSP1 and emphasized structure-related bottlenecks, which must be solved to develop very specifically targeting drugs for CTDSP1. Structure-based drug design has become a handy tool to solve many complex problems, such as handling the big data of combinatorial synthesis, improving the structure of multitarget drug molecules, dealing with the structural similarity of target proteins, and finding a perfect fit of drug molecules for target proteins in the drug discovery process [20–35]. Molecular docking is a standard modeling tool, and various docking techniques, such as allosteric docking, ensemble docking, and composite docking, have evolved to execute structurally complex issues [24,36–62]. Allosteric docking is a developing method to find drug molecules which are inhibiting or modulating through specific allosteric sites on the enzyme. Ensemble docking uses the generation of an ensemble of target protein, often gained from molecular dynamics simulation. Composite docking is a combined approach with several docking techniques. These docking techniques could provide a way of probing the druggability on the interface of protein–protein interactions [63], searching for the allosteric regulation site of target proteins [64,65], screening of target proteins with flexible structures [66,67], calculating the off-target binding of drug molecules [68,69], and overcoming the weaknesses associated with computer screening methods [70,71]. Therefore, we are going to explore the addressed issues and previous techniques followed by other researchers. We will focus on contemporary techniques of molecular docking, which will help make progress towards successful targeting of CTDSP1.

2. CTDSP1 as an Emerging Target

When selecting a new drug target, we must consider the activity of the target protein depending on various factors, such as expression level, subcellular localization, post-translational modification, and the protein–protein interaction network [72–76]. Previously, many experiments have acknowledged the various roles of CTDSP1, such as neuronal gene silencing, cell cycle regulation, and regulation of TGF β and BMP signal transduction. The biological roles of CTDSP1 are summarized in Figure 1, Table 1, and another review paper [8]. Interestingly, recent studies have proclaimed new functional insights related to cancer, which is a new dimension of potential activity. Hence, here we are concentrating on how CTDSP1 is implicated in different cancer-related mechanisms and binding partners, as one of the new drug targets.

Cancer progression involves different stages of migration, invasion, and metastasis, and each step is carried out by different transcription factors. Among them, Twist-related protein (TWIST) is commonly identified in many cancers, such as breast, gastric, squamous cell carcinoma, cervical, ovarian, esophageal, and gastric cancers. TWIST is associated with different signaling cascades, like TGF β and v-akt oncogene homolog (AKT)/phosphoinositide 3-kinase (PI3K), to promote cancer cell invasion and metastasis with the help of B lymphoma Mo-MLV insertion region 1 homolog (BMI1) and v-akt oncogene homolog 2 (AKT2) [77,78]. It has been observed that a mitogen-activated protein kinase (MAPK) phosphorylates the Twist-related protein 1 (TWIST1) on N-terminal Ser68, and stabilizes it to enhance cancer cell migration, invasion, and metastasis in breast cancer cells. Further, in contrast to the findings of previous studies, CTDSP1 has been shown to dephosphorylate the Ser68 residue of TWIST1 through the interaction of the N-terminus of TWIST1 and amino acid residues 43–63 of

CTDSP1. Additionally, no other similar phosphatases were able to constitute this reaction with the same efficiency, and the inverse concentration of TWIST1 and CTDSP1 has been observed in several cancer cell lines. This suggests that CTDSP1 can potentially attenuate TWIST1 and antagonize cancer progression [16].



Figure 1. Scheme illustrating the biological substrates and roles of CTDSP1. The dashed arrows represent subcellular translocation of proteins, and the solid arrows represent phosphorylation (black) and dephosphorylation (blue) of proteins. The red lines indicate the biological results of dephosphorylation by CTDSP1.

CTDSP1 has been known as a nuclear phosphatase, but in a recent study, researchers found that it is localized to the plasma membrane by lipid modifications [79]. This is also fashionably related to the tumor growth and N-terminal modification abilities of CTDSP1. AKT activation requires Ser473 and Thr308 phosphorylation, and consequently is involved in angiogenesis and tumor growth [80–85]. Many pieces of research have focused on the C-terminal activity of CTDSP1, but the understanding of N-terminal residues of CTDSP1 is evolving, with new activities identified in impeding tumor growth, invasion, and metastasis. Myristoylated AKT was found to be localized to the plasma membrane, and at the same time, CTDSP1 was also identified in the plasma membrane [86,87]. The reason for the membrane localization of CTDSP1 does not originate from a transmembrane domain, because CTDSP1 does not have the transmembrane domain [6]. N-terminal Cys44 and Cys45 of CTDSP1 (Figure 2) are the key residues involved in plasma membrane localization, which undergo the most dynamic posttranslational modification—palmitoylation [79]. CTDSP1 could be palmitoylated through the thioester bond between palmitate and N-terminal Cys, and the modification anchors CTDSP1 to the plasma membrane [79]. Mutational studies have demonstrated that only palmitoylated CTDSP1 performs selective dephosphorylation of AKT [79]. The palmitoylation-unable mutant of CTDSP1 failed in the removal of the phosphate group from the Ser473 of AKT [79]. The direct interaction of AKT and CTDSP1 could negatively modulate angiogenesis by tumor growth suppression through dephosphorylation of Ser473. These observations were illustrated by wound healing assay and in vitro experiments in HeLa Cells [79]. In summary, CTDSP1 can be a tumor suppressor through inhibiting cancer cell migration and invasion by the dephosphorylation of TWIST and AKT. The tumor suppressor properties of CTDSP2 and CTDSPL has also been reported in recent studies [88-91]. CTDSP2 and CTDSPL also inhibit tumor growth and angiogenesis through dephosphorylation of RB and tumor-suppressor protein promyelocytic leukemia (PML).

CTD Small Phosphatase	Aliases	Substrates	PDB IDs	Biological Roles [References]
CTDSP1	SCP1 NIF3	RNAP II CTD REST R-SMADs CDCA3, RB TWIST1, AKT PML, c-Myc	4YGY, 4YH1, 3PGL, 3L0B, 3L0C, 3L0Y, 2GHQ, 2GHT, 1T9Z, 1TA0, 3PGL	Transcription factor recruitment [5,93] Neuronal gene silencing [9,10] Osteoblast differentiation [12] Cell cycle regulation [13,14] Tumor suppressor [88–91]
CTDSP2	SCP2 OS4 NIF2	RNAP II CTD R-SMADs PML	2Q5E	Transcription factor recruitment [5,94] Promoter clearance [94] Ras activation [95]
CTDSPL	SCP3 HYA22 NIF1	RNAP II CTD R-SMADs PML	2HHL	Transcription factor recruitment [5] Tumor suppressor [90]

Table 1. Biochemical characteristics and biological roles of CTD small phosphatases, Modified from another review paper [92].



Figure 2. Structural similarity of CTD small phosphatases. (**A**) The sequence alignment of human CTDSP1 1-261 (NCBI accession number: NP_067021.1) with human CTDSP2 1-271 (NCBI accession number: NP_005721.3) and human CTDSPL 1-276 (NCBI accession number: NP_001008392.1), showing the active site (yellow square), palmitoylation residues (blue square), N-terminal insertion residues (black square), and representative residues consisting of the hydrophobic pocket of human CTDSP1 written in black below the alignment. (**B**) 3D structure alignment of human CTDSP1 (green, PDB ID: 3PGL), human CTDSP2 (yellow, PDB ID: 2Q5E), and human CTDSPL (red, PDB ID: 2HHL).

Given the above findings, CTDSP1 could be a novel target for cancer therapy. Other biological activities from earlier research have been based on the C-terminal domain and catalytic site of CTDSP1. Present experiments are showing that the N-terminal domain of CTDSP1 also has an inevitable role in CTDSP1 activity [79], which might be a unique feature compared to other CTDSPs. Thus, the regulation of CTDSP1 expression and activity may be related to the development of0 novel strategies for cancer treatment. However, there are still many challenges in targeting CTDSP1, such as solving the structural complexity among CTDSPs, finding the unique characteristics of CTDSP1, and looking for novel druggable materials to target CTDSP1 in cancer-related research and development.

3. Challenges and Opportunities in Targeting CTDSP1

There are seven human CTD phosphatases: CTDP1, CTDSP1, CTDSP2, CTDSPL, CTD small phosphatase like 2 (CTDSPL2), CTD nuclear envelope phosphatase 1 (CTDNEP1), and ubiquitin-like domain-containing CTD phosphatase 1 (UBLCP1). Based on the phylogenetic analysis of human CTD phosphatases presented in a previous article [96], CTDSP1, CTDSP2, CTDSPL, and CTDNEP1 are closely related, but CTDP1, UBLCP1, and CTDSPL2 are distantly related. Similar phenomena have been observed in a sequence alignment analysis presented in another review paper [95]. Currently, several structures of CTD phosphatases have been deposited into the protein data bank [6,7,17,97–102]. Among them, CTDSP1 (PDB ID: 2GHT, 2GHQ, 3PGL, 4YH1, 4YGY, 3L0B, 3L0C, 3L0Y, 2GHT, 1T9Z, 1TA0), CTDSP2 (PDB ID: 2Q5E), and CTDSPL (PDB ID: 2HHL) are structurally much conserved [5]. The structural similarity of CTDSP1, CTDSP2, and CTDSPL observed through the sequence and 3D structure alignment is presented in Figure 2. Additionally, the functionality of CTDSP2 and CTDSPL, such as their catalytic activity against RNAP II CTD peptide and their tumor-suppressing role, also overlaps with that of CTDSP1 [5,7,10,88]. However, CTDSP2 and CTDSPL have 10 or 12 residue insertions on the N-terminal end adjacent to the active sites, respectively, but CTDSP1 has no insertion on the N-terminal end, based on our structure pairwise comparison (Figure 2). Although they are operating on similar substrates, such as RNAP II CTD and R-SMADs with conserved signature active sites [5,7,10], they might be highly specific towards their binding partners due to N-terminal differences.

In earlier studies [17,18], researchers emphasized the structural complexity in targeting CTDSP as the significant bottleneck; hence, in this section, we are going to address the structural barriers and privileges of CTDSP1 as a successful biological target. The major challenge in targeting CTDSP1 is the conserved structures of CTDSP (Figure 2). CTDSP1, CTDSP2, and CTDSPL are approximately 40% conserved among their full sequences, and are highly conserved around the catalytic site [5,7,10]. Attacking the catalytic site of CTDSP1 may also impact other CTDSPs with identical domains [7,17]. The second complication is that CTDSP1 is implicated in different signaling events, such as neuronal gene silencing, negative regulation of cancer, and cell cycle regulation; hence, nonspecific targeting may impact many different biological activities of CTDSP1 [8]. However, CTDSP1 shows such different cellular localization traits that understanding the particular purpose of CTDSP1's location in a cell is mandatory.

A thorough review of the literature demonstrated the history of selective inactivation of protein phosphatases with small molecules [103]. These studies have helped us learn about vital targetable sites, and will improve the precision in targeting CTDSP1. One successful instance occurred in impeding the CTDSP1 effect using rabeprazole, as a commercially available drug [17]. The co-crystallization experiment exhibited a precise targeting of CTDSP1 by using its unique features (Figure 3). The architecture of the CTDSP active site domain illustrates that the first aspartate residue (Asp96 in CTDSP1) acts as a nucleophile in the general acid/base mechanism. The consecutive one (Asp98 in CTDSP1) helps in the nucleophilic attack, along with metal ion Mg^{2+} , but these residues are not involved in substrate binding. CTDSP1's active site is surrounded by a unique hydrophobic pocket composed of Phe106, Val118, Leu155, Tyr158, and Arg178 residues (Figure 3B,D). It forms a binding groove with proline residues of a heptapeptide repeat of RNAP II CTD [6]. Rabeprazole blocks the binding groove and interrupts the dephosphorylation of CTD peptide [17]. The above result emphasizes that the hydrophobic pocket, shown in Figure 3, helps with the unique targeting of CTDSP1. Further, it can be inferred from previous studies [5,79] that the substrate selectivity of CTDSP1 will be initially decided by cellular localization and further binding groove residues. These characteristics will help with finding a novel targeting strategy to precisely regulate CTDSP1's activity.



Figure 3. Structural representation of human CTDSP1. (**A**) The hydrophobic pocket of human CTDSP1 along with the co-crystallized ligand rabeprazole, marked with an orange circle, adapted from the PDB database (PDB ID: 3PGL). (**B**) Close-up view of the hydrophobic pocket in A. (**C**) The active site and hydrophobic pocket of human CTDSP1, along with rabeprazole and CTD phosphopeptide, showing the related amino acid residues. (**D**) Close-up view of the active site and hydrophobic pocket in C. (**E**) Clustered structures of human CTDSP1 along with natural and chemical ligands adapted from the PDB databank (PDB ID: 2GHT, 2GHQ, 3PGL, 4YH1, 4YGY).

The field of drug designing has been rapidly advancing in recent decades, and many new state-of-the-art technologies and methods have been discovered [20,104–114]. Well rationalized structural information of receptor proteins and ligand chemicals is the essential requirement for drug designing. Based on this principle, the structure-based drug discovery (SBDD) process has evolved. In recent decades, the structure-based drug design technique has been considered the best tool to attack molecular targets with high accuracy. Highly ambiguous targets like enzymes need to be considered carefully and precisely. Doing so requires extensive structural information about the various ensembles, like pockets, folds, and hydrogen bonds. Focusing proteins in drug discovery direct the way to finding the best lead compound, with high efficacy at the lowest micromolar concentrations. Numerous software tools for protein-ligand dockings, such as Autodock, DOCK 4.0, GOLD, Flex-X, ICM, SLIDE, GLIDE, and DeepBindRG, were developed based on various scoring functions, including empirical, knowledge-based, and machine learning, and are summarized in Table 2 [50,115–130].

C-terminal domain small phosphatases have evolved with conserved structures and similar functions [5,96]. Proper structures for SBDD are available only for CTDSP1, CTDSP2, and CTDSPL. Finely determined X-ray structures (PDB ID: 2GHT, 2GHQ, 3PGL, 4YH1, 4YGY) are available for CTDSP1 in the PDB database. The essential requirement for structure-based drug design [131], which has also been considered in a few earlier studies [17,18], is the need to target precisely the successive regulation of CTDSP1. Previous studies can also help us to understand the essence of the combinatorial approach in the successful identification of novel small molecules for CTDSP1 as a drug target. The recent technology of allosteric docking has been tested successfully to find regulatory probe molecules for distinct molecules, like receptor tyrosine kinases (RTK) [132] and protein tyrosine phosphatases (PTP) [133]. Allosteric docking techniques to develop PTP inhibitors have been summarized in previous review papers [134,135]. We suggest the application of selective

targeting techniques, such as allosteric targeting and ensemble docking, to overcome the obstacles in modulating the activity of CTDSP1.

Docking Tool	Algorithm	Scoring Function	Representative References
Autodock	Genetic algorithm, Simulated annealing, Lamarckian algorithm	Force field, Empirical	[55]
DOCK 4.0	Incremental search, Shape fitting	Force field	[54]
GOLD	Genetic algorithm	Force field	[57]
Flex-X	Multiple copy, Simultaneous search, Incremental search	Empirical	[58]
ICM	Monte Carlo sampling	Empirical	[59]
SLIDE	Incremental construction	Force field, Empirical	[60]
GLIDE	Monte Carlo sampling	Empirical	[61,62]
GOLD/ASP	Genetic algorithm	Knowledge-based	[127,128]
DeepBindRG	Deep learning algorithm	Machine learning	[129]

Table 2. Comparison of some frequently-used molecular docking tools.

Generally, people target the active site of a receptor protein to modify or inhibit the functionality of the receptor protein in drug discovery. However, it is very difficult to find a distinguished method for developing new molecules for biological or metabolic target proteins with structural similarities. Traditional inhibition techniques will harm the system by targeting active sites [105]. Other functionally imperative sites can be identified through allosteric targeting, apart from the regular active site of a receptor protein. The identified allosteric sites can help restrain the activity of a protein. Allosteric targeting can be achieved in numerous ways, such as targeting protein–protein interactions, focusing on multivalent inhibitors, and targeting disulfide bonds [136–140]. In the process of allosteric targeting, pharmacophore modeling of a ligand could apply for targeting both a directed specific site and other distinctive sites of a protein. Therefore, controlling protein activity by selecting allosteric sites of a receptor protein could be one of the targeting methods for CTDSP1.

In recent drug discovery history, with the advancement of technology, robust tools have been evolving, and some researchers are aiming to control activities at the metabolic level by targeting checkpoint signaling molecules in cancer [141,142]. The allosteric docking technique gained importance, since these vital targets were operated effectively. Allosteric docking involves a few crucial steps like allosteric binding site identification, analysis, and optimization [104,105,143]. The allosteric binding sites are identified through the literature and database searches, cavity finding, blind docking, coevolution, and so forth. Optimized allosteric binding sites are obtained from analysis and optimization, with the help of molecular dynamics simulation through perturbation response scanning, trajectory analysis, course-grained residue network analysis, and calculating Gibbs free energy [144]. Allosteric docking involves similar docking tools and algorithms to the other docking techniques summarized in Table 2. A careful and precise selection of allosteric binding sites based on understanding biochemical and biological characteristics of target proteins, and choosing proper docking tools, could increase the success rate of targeting [145,146]. A study on protein tyrosine phosphatase 1B (PTP1B), which is a crucial member in various biological activities and has the conserved catalytic site among PTP, showed that an allosteric inhibitor causes conformational changes in the protein outside the active site of PTP1B, and abolishes the enzyme activity of PTP1B extrinsically [147].

Similarly, CTDSP1 also has a conserved structure and similar activity to other CTDSPs. We observed that CTDSP1 also possesses characteristics from co-crystallized ligand experiments and a unique hydrophobic site adjacent to the active site, shown in Figure 3, which helps the inhibitor to bind

to CTDSP1 without affecting catalytic residues [17,18]. Interestingly, CTDNEP1, as one of the conserved paralogs, was not affected by the ligand rabeprazole as much as CTDSP1 [17]. This report suggests that the hydrophobic pocket of CTDSP1 is highly specific for CTDSP1's substrate recognition, and has importance in dephosphorylation activity against natural substrates. This site could be a potential binding target for allosteric drugs, which is satisfied with the fundamental requirement for non-conventional targeting. Furthermore, finding an allosteric targeting site for CTDSP1 is so recent that there might be many chances to identify novel binding sites to target CTDSP1 allosterically.

Observations from recent studies [16,79] suggest that N-terminal palmitoylated CTDSP1 is required for TWIST1 protein dephosphorylation. This whole mechanism is constituted through palmitoylation of the Cys44 and Cys45 residues on an N-terminal domain of CTDSP1 (Figures 1 and 2). The mechanism is also specific to CTDSP1, since no other paralogs were evidenced in the experiment. This observation could be one of the other potential targetable allosteric sites. Although the binding possibility of allosteric drug molecules of CTDSP1 to other CTDSP has not been totally excluded, we may develop a way of targeting the N-terminal region of CTDSP1 based on the difference of N-terminal sequences among CTDSPs. In this case, the ligand might be bound away from the catalytic site of CTDSP1, and the targeting chemicals might not influence all other CTDSP1 using several predicted sites or regions to be found through a careful investigation.

Ensemble docking is an emerging method that works based on the structure-based drug design technique. It is a high-throughput and robust drug discovery technique, where multiple structures can be validated simultaneously and accurately. This docking method uses all discrete ensembles of a protein structure (EPS) [49,148]. In the ensemble docking process, generating various conformations of a receptor protein is the first step, and then, degrees of freedom will be calculated based on ligand binding. GOLD docking software is extensively used for implementing this technique. Initially, the structural library of a target protein with its conformational structures is prepared, and then, it is compared with a highly converged stable reference structure. Root mean square deviation (RMSD) differences of molecular dynamics trajectories based on the comparison with the reference structure are often calculated in vital regions of a target protein [149]. The structures are docked with the desired ligand, and conformational binding poses are predicted for complexes. Based on the degree the experimental structure is conserved, predicted poses would be ranked and continued for post-processing. This technique has been successfully applied with the p38 MAP kinase system, and outperformed conventional methods [150]. This method could be the best for targeting CTDSP1 due to the structural similarities among CTDSPs, the existence of multiple structures that can be compared simultaneously, and the availability of experimental data of co-crystallized ligands for accurate pose prediction and analysis.

Comparison of the chemical and natural ligand structure's discretion against the apo-protein demonstrates the stability of the receptor protein and differences in conformational changes, which are useful for validating other paralog structures [46,151]. The RMSD values of the backbone and side-chain atoms denote the fundamental differences among protein structures, and extend insights into developing selective inhibitors [152]. Ensemble docking is one of the most promising techniques for modulating CTDSP1 activity, since CTDSP1, CTDSP2, and CTDSPL have the well-built X-ray structures (Figure 2B) required for ensemble prediction and analysis. Moreover, CTDSP1 has multiple 3D structures (Figure 3E) incorporated with natural and chemical ligands, for use as a reference structure. Therefore, ensemble docking can elicit the information of unique sites on CTDSP1 compared to other CTDSPs, which is vital for drug discovery targeting CTDSP1.

We present the comparison of conventional docking, allosteric docking, and ensemble docking of targeting CTDSP1 in Figure 4. Conventional docking target the active site of CTDSP1, disturbing the diverse biological roles of it and the catalytic activities of CTDSP2 and CTDSPL. Thus, conventional docking is not a desirable method of SBDD. Instead of the conventional approach, we suggest allosteric docking and ensemble docking for targeting CTDSP1 as better methods of SBDD. Allosteric docking

targets a binding site other than the active site, which can, therefore, produce much more specific drug molecules for CTDSP1 without interrupting the activities of other CTDSPs. Ensemble docking uses the generated conformational library of CTDSP1 and produces the most vital regions of it. The produced essential regions could be tested through biochemical screening, and may contribute optimization when targeting CTDSP1. However, we agree that there could be yet unexplored, better docking methods than allosteric docking and ensemble docking in the targeting of CTDSP1.



Figure 4. Comparison of some molecular docking methods. The receptor protein shows the structure of human CTDSP1, and the ligand library shows several candidate compounds [17,18,153] for targeting CTDSP1. The left square represents conventional docking, which targets the active site of human CTDSP1. The middle square shows allosteric docking, which finds functionally significant regions as allosteric sites, shown with black circles, of human CTDSP1. The right square presents the ensemble generation of human CTDSP1, which is the primary technique used in ensemble docking.

4. Conclusions

CTDSP1's protein activity is vital for various biological activities. To date, the C-terminal catalytic domain of CTDSP1 has been studied, and limited activities of CTDSP1 identified. Interestingly, recent studies have extended the knowledge we have of CTDSP1's functionality by focusing on the N-terminal region of CTDSP1. This new functional insight, as a novel protein regulator, has increased the significance of CTDSP1. The highly similar structures of CTDSPs have become barriers to targeting CTDSP1. Therefore, we focused on finding contemporary and robust methods and tools to confront this structural problem. The previous studies offered information about a critical functional site of a hydrophobic pocket adjacent to the active site of CTDSP1, which is essential for accommodating natural substrates. The N-terminal region of CTDSP1 is required for palmitoylation, to anchor proteins in the plasma membrane, and was identified as a potential targetable allosteric site of CTDSP1. A novel allosteric site may also be found through the investigation on protein-protein interaction analysis of CTDSP1. Targeting the hydrophobic pocket, N-terminal region, and a novel site of CTDSP1 could provide a new chance to develop cancer-related treatments. Allosteric docking has been identified as a suitable technique for this, as it can precisely target desired sites on receptor proteins. Therefore, allosteric docking could be an approach to target the hydrophobic pocket, N-terminal region, and novel site of CTDSP1. Ensemble docking produces numerous structural confirmations of receptors through molecular dynamics. The generated ensembles help to exclude false positives and discover highly specific inhibitors. CTDSP1 has close structural analogues, like CTDSP2 and CTDSPL. Rich data

are available on CTDSP1's structure, which will be very useful for the design of potential novel drugs through ensemble docking with CTDSP1. Thus, the application of ensemble docking for CTDSP1 could potentially outperform conventional methods. Finally, selectivity could be achieved by associating allosteric docking and ensemble docking as robust and high-throughput methods for targeting CTDSP1. In order to develop particular drug molecules for CTDSP1, more detailed investigations searching for biochemical characteristics among CTDSPs, revealing the protein–protein interaction network of CTDSP1, and finding a novel allosteric site of CTDSP1 are required. We hope that more researchers will continue to disclose more precise ways of targeting CTDSP1, with the help of structure-based drug design.

Author Contributions: Writing—original draft preparation, H.R.R.; writing—review and editing, P.G. and Y.J.K.; supervision, Y.J.K.; funding acquisition, Y.J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Research Foundation of Korea, grant number 2019R1F1A1059216 and 2016R1D1A1B03932519.

Acknowledgments: We acknowledge the contribution of Dongsun Lee and Jinmo Sung to preparing this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Harlen, K.M.; Churchman, L.S. The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 263–273. [CrossRef] [PubMed]
- 2. Jeronimo, C.; Collin, P.; Robert, F. The RNA Polymerase II CTD: The Increasing Complexity of a Low-Complexity Protein Domain. *J. Mol. Biol.* **2016**, *428*, 2607–2622. [CrossRef] [PubMed]
- Mandal, S.S.; Cho, H.; Kim, S.; Cabane, K.; Reinberg, D. FCP1, a phosphatase specific for the heptapeptide repeat of the largest subunit of RNA polymerase II, stimulates transcription elongation. *Mol. Cell. Biol.* 2002, 22, 7543–7552. [CrossRef] [PubMed]
- 4. Archambault, J.; Pan, G.; Dahmus, G.K.; Cartier, M.; Marshall, N.; Zhang, S.; Dahmus, M.E.; Greenblatt, J. FCP1, the RAP74-interacting subunit of a human protein phosphatase that dephosphorylates the carboxyl-terminal domain of RNA polymerase IIO. *J. Biol. Chem.* **1998**, 273, 27593–27601. [CrossRef]
- 5. Yeo, M.; Lin, P.S.; Dahmus, M.E.; Gill, G.N. A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *J. Biol. Chem.* **2003**, *278*, 26078–26085. [CrossRef]
- Zhang, Y.; Kim, Y.; Genoud, N.; Gao, J.; Kelly, J.W.; Pfaff, S.L.; Gill, G.N.; Dixon, J.E.; Noel, J.P. Determinants for dephosphorylation of the RNA polymerase II C-terminal domain by Scp1. *Mol. Cell* 2006, 24, 759–770. [CrossRef]
- 7. Kamenski, T.; Heilmeier, S.; Meinhart, A.; Cramer, P. Structure and mechanism of RNA polymerase II CTD phosphatases. *Mol. Cell* **2004**, *15*, 399–407. [CrossRef]
- 8. R, H.R.; Kim, H.; Noh, K.; Kim, Y.J. The diverse roles of RNA polymerase II C-terminal domain phosphatase SCP1. *BMB Rep.* **2014**, *47*, 192–196.
- 9. Nesti, E.; Corson, G.M.; McCleskey, M.; Oyer, J.A.; Mandel, G. C-terminal domain small phosphatase 1 and MAP kinase reciprocally control REST stability and neuronal differentiation. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E3929–E3936. [CrossRef]
- 10. Yeo, M.; Lee, S.K.; Lee, B.; Ruiz, E.C.; Pfaff, S.L.; Gill, G.N. Small CTD phosphatases function in silencing neuronal gene expression. *Science* 2005, *307*, 596–600. [CrossRef]
- Burkholder, N.T.; Mayfield, J.E.; Yu, X.; Irani, S.; Arce, D.K.; Jiang, F.; Matthews, W.L.; Xue, Y.; Zhang, Y.J. Phosphatase activity of small C-terminal domain phosphatase 1 (SCP1) controls the stability of the key neuronal regulator RE1-silencing transcription factor (REST). *J. Biol. Chem.* 2019, 293, 16851–16861. [CrossRef] [PubMed]
- Sapkota, G.; Knockaert, M.; Alarcon, C.; Montalvo, E.; Brivanlou, A.H.; Massague, J. Dephosphorylation of the linker regions of Smad1 and Smad2/3 by small C-terminal domain phosphatases has distinct outcomes for bone morphogenetic protein and transforming growth factor-beta pathways. *J. Biol. Chem.* 2006, 281, 40412–40419. [CrossRef] [PubMed]

- 13. Wang, W.; Liao, P.; Shen, M.; Chen, T.; Chen, Y.; Li, Y.; Lin, X.; Ge, X.; Wang, P. SCP1 regulates c-Myc stability and functions through dephosphorylating c-Myc Ser62. *Oncogene* **2015**, *35*, 491–500. [CrossRef] [PubMed]
- Liu, Y.; Liu, W.B.; Liu, K.J.; Ao, L.; Cao, J.; Zhong, J.L.; Liu, J.Y. Overexpression of miR-26b-5p regulates the cell cycle by targeting CCND2 in GC-2 cells under exposure to extremely low frequency electromagnetic fields. *Cell Cycle* 2015, 15, 357–367. [CrossRef] [PubMed]
- 15. Kim, Y.J.; Bahk, Y.Y. A study of substrate specificity for a CTD phosphatase, SCP1, by proteomic screening of binding partners. *Biophys. Res. Commun.* **2014**, *448*, 189–194. [CrossRef]
- Sun, T.; Fu, J.; Shen, T.; Lin, X.; Liao, L.; Feng, X.H.; Xu, J. The Small C-terminal Domain Phosphatase 1 Inhibits Cancer Cell Migration and Invasion by Dephosphorylating Ser(P)68-Twist1 to Accelerate Twist1 Protein Degradation. J. Biol. Chem. 2016, 291, 11518–11528. [CrossRef]
- 17. Zhang, M.; Cho, E.J.; Burstein, G.; Siegel, D.; Zhang, Y. Selective inactivation of a human neuronal silencing phosphatase by a small molecule inhibitor. *ACS Chem. Biol.* **2011**, *6*, 511–519. [CrossRef]
- Park, H.; Lee, H.S.; Ku, B.; Lee, S.R.; Kim, S.J. Two-track virtual screening approach to identify both competitive and allosteric inhibitors of human small C-terminal domain phosphatase 1. *J. Comput. Aided Mol. Des.* 2017, *31*, 743–753. [CrossRef]
- Yoshida, T.; Yamazaki, K.; Imai, S.; Banno, A.; Kaneko, A.; Furukawa, K.; Chuman, Y. Identification of a Specific Inhibitor of Human Scp1 Phosphatase Using the Phosphorylation Mimic Phage Display Method. *Catalysts* 2019, *9*, 842. [CrossRef]
- 20. Bera, I.; Payghan, P.V. Use of Molecular Dynamics Simulations in Structure-Based Drug Discovery. *Curr. Pharm. Des.* **2019**, *25*, 3339–3349. [CrossRef]
- Lane, J.R.; Chubukov, P.; Liu, W.; Canals, M.; Cherezov, V.; Abagyan, R.; Stevens, R.C.; Katritch, V. Structure-based ligand discovery targeting orthosteric and allosteric pockets of dopamine receptors. *Mol. Pharm.* 2013, *84*, 794–807. [CrossRef]
- 22. Majewski, M.; Ruiz-Carmona, S.; Barril, X. Dynamic Undocking: A Novel Method for Structure-Based Drug Discovery. *Methods Mol. Biol.* 2018, *1824*, 195–215.
- Morra, G.; Genoni, A.; Neves, M.A.; Merz, K.M., Jr.; Colombo, G. Molecular recognition and drug-lead identification: What can molecular simulations tell us? *Curr. Med. Chem.* 2009, 17, 25–41. [CrossRef] [PubMed]
- 24. Batool, M.; Ahmad, B.; Choi, S. A Structure-Based Drug Discovery Paradigm. *Int. J. Mol. Sci.* 2019, 20, 2783. [CrossRef] [PubMed]
- 25. Bajusz, D.; Ferenczy, G.G.; Keseru, G.M. Structure-based Virtual Screening Approaches in Kinase-directed Drug Discovery. *Curr. Top. Med. Chem.* **2017**, *17*, 2235–2259. [CrossRef]
- 26. Barril, X.; Hubbard, R.E.; Morley, S.D. Virtual screening in structure-based drug discovery. *Mini Rev. Med. Chem.* **2004**, *4*, 779–791. [CrossRef] [PubMed]
- 27. Ghosh, S.; Nie, A.; An, J.; Huang, Z. Structure-based virtual screening of chemical libraries for drug discovery. *Curr. Opin. Chem. Biol.* **2006**, *10*, 194–202. [CrossRef]
- 28. Tavousi, P.; Amin, R.; Shahbazmohamadi, S. Assemble-And-Match: A Novel Hybrid Tool for Enhancing Education and Research in Rational Structure Based Drug Design. *Sci. Rep.* **2018**, *8*, 849. [CrossRef]
- 29. Anderson, A.C. The process of structure-based drug design. Chem. Biol. 2003, 10, 787–797. [CrossRef]
- 30. Verlinde, C.L.; Hol, W.G. Structure-based drug design: Progress, results and challenges. *Structure* **1994**, 2, 577–587. [CrossRef]
- Wang, X.; Song, K.; Li, L.; Chen, L. Structure-Based Drug Design Strategies and Challenges. *Curr. Top. Med. Chem.* 2018, 18, 998–1006. [CrossRef] [PubMed]
- 32. Grey, J.L.; Thompson, D.H. Challenges and opportunities for new protein crystallization strategies in structure-based drug design. *Expert Opin. Drug Discov.* **2010**, *5*, 1039–1045. [CrossRef] [PubMed]
- Lounnas, V.; Ritschel, T.; Kelder, J.; McGuire, R.; Bywater, R.P.; Foloppe, N. Current progress in Structure-Based Rational Drug Design marks a new mindset in drug discovery. *Comput. Struct. Biotechnol. J.* 2013, 5, e201302011. [CrossRef] [PubMed]
- 34. Cheng, T.; Li, Q.; Zhou, Z.; Wang, Y.; Bryant, S.H. Structure-based virtual screening for drug discovery: A problem-centric review. *Aaps J.* **2012**, *14*, 133–141. [CrossRef]
- Van Montfort, R.L.M.; Workman, P. Structure-based drug design: Aiming for a perfect fit. *Essays Biochem*. 2017, *61*, 431–437. [CrossRef]

- 36. Bajusz, D.; Ferenczy, G.G.; Keseru, G.M. Ensemble docking-based virtual screening yields novel spirocyclic JAK1 inhibitors. *J. Mol. Graph. Model.* **2016**, *70*, 275–283. [CrossRef]
- 37. Cavasotto, C.N. Normal mode-based approaches in receptor ensemble docking. *Methods Mol. Biol.* **2012**, *819*, 157–168.
- 38. Cowan-Jacob, S.W.; Jahnke, W.; Knapp, S. Novel approaches for targeting kinases: Allosteric inhibition, allosteric activation and pseudokinases. *Future Med. Chem.* **2014**, *6*, 541–561. [CrossRef]
- 39. Ellingson, S.R.; Miao, Y.; Baudry, J.; Smith, J.C. Multi-conformer ensemble docking to difficult protein targets. *J. Phys. Chem. B* **2014**, *119*, 1026–1034. [CrossRef]
- 40. Da, C.; Mooberry, S.L.; Gupton, J.T.; Kellogg, G.E. How to deal with low-resolution target structures: Using SAR, ensemble docking, hydropathic analysis, and 3D-QSAR to definitively map the alphabeta-tubulin colchicine site. *J. Med. Chem.* **2013**, *56*, 7382–7395. [CrossRef]
- 41. Erol, I.; Aksoydan, B.; Kantarcioglu, I.; Salmas, R.E.; Durdagi, S. Identification of novel serotonin reuptake inhibitors targeting central and allosteric binding sites: A virtual screening and molecular dynamics simulations study. *J. Mol. Graph. Model.* **2017**, *74*, 193–202. [CrossRef] [PubMed]
- Evangelista, W.; Weir, R.L.; Ellingson, S.R.; Harris, J.B.; Kapoor, K.; Smith, J.C.; Baudry, J. Ensemble-based docking: From hit discovery to metabolism and toxicity predictions. *Bioorg. Med. Chem.* 2016, 24, 4928–4935. [CrossRef] [PubMed]
- 43. Fayaz, S.M.; Rajanikant, G.K. Ensemble pharmacophore meets ensemble docking: A novel screening strategy for the identification of RIPK1 inhibitors. *J. Comput. Aided Mol. Des.* **2014**, *28*, 779–794. [CrossRef] [PubMed]
- 44. Geng, C.; Narasimhan, S.; Rodrigues, J.P.; Bonvin, A.M. Information-Driven, Ensemble Flexible Peptide Docking Using HADDOCK. *Methods Mol. Biol.* **2017**, *1561*, 109–138. [PubMed]
- Henriksen, S.T.; Liu, J.; Estiu, G.; Oltvai, Z.N.; Wiest, O. Identification of novel bacterial histidine biosynthesis inhibitors using docking, ensemble rescoring, and whole-cell assays. *Bioorg. Med. Chem.* 2010, *18*, 5148–5156. [CrossRef] [PubMed]
- 46. Huang, S.Y.; Zou, X. Ensemble docking of multiple protein structures: Considering protein structural variations in molecular docking. *Proteins* **2007**, *66*, 399–421. [CrossRef]
- 47. Okamoto, Y.; Kokubo, H.; Tanaka, T. Ligand docking simulations by generalized-ensemble algorithms. *Adv. Protein Chem. Struct. Biol.* **2013**, *92*, 63–91.
- 48. Zhang, Z.; Ehmann, U.; Zacharias, M. Monte Carlo replica-exchange based ensemble docking of protein conformations. *Proteins* **2017**, *85*, 924–937. [CrossRef]
- Xu, M.; Lill, M.A. Utilizing experimental data for reducing ensemble size in flexible-protein docking. J. Chem. Inf. Model. 2012, 52, 187–198. [CrossRef]
- 50. Lam, P.C.; Abagyan, R.; Totrov, M. Ligand-biased ensemble receptor docking (LigBEnD): A hybrid ligand/receptor structure-based approach. *J. Comput. Aided Mol. Des.* **2018**, *32*, 187–198. [CrossRef]
- Lape, M.; Elam, C.; Paula, S. Comparison of current docking tools for the simulation of inhibitor binding by the transmembrane domain of the sarco/endoplasmic reticulum calcium ATPase. *Biophys. Chem.* 2010, 150, 88–97. [CrossRef] [PubMed]
- 52. Westhead, D.R.; Clark, D.E.; Murray, C.W. A comparison of heuristic search algorithms for molecular docking. *J. Comput. Aided Mol. Des.* **1997**, *11*, 209–228. [CrossRef] [PubMed]
- 53. Meng, X.Y.; Zhang, H.X.; Mezei, M.; Cui, M. Molecular docking: A powerful approach for structure-based drug discovery. *Curr. Comput. Aided Drug Des.* **2011**, *7*, 146–157. [CrossRef] [PubMed]
- 54. Ewing, T.J.; Makino, S.; Skillman, A.G.; Kuntz, I.D. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput. Aided Mol. Des.* **2001**, *15*, 411–428. [CrossRef] [PubMed]
- 55. Forli, S.; Huey, R.; Pique, M.E.; Sanner, M.F.; Goodsell, D.S.; Olson, A.J. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **2016**, *11*, 905–919. [CrossRef]
- 56. Pagadala, N.S.; Syed, K.; Tuszynski, J. Software for molecular docking: A review. *Biophys. Rev.* 2017, 9, 91–102. [CrossRef]
- 57. Jones, G.; Willett, P.; Glen, R.C.; Leach, A.R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, 267, 727–748. [CrossRef]
- 58. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489. [CrossRef]
- 59. Neves, M.A.; Totrov, M.; Abagyan, R. Docking and scoring with ICM: The benchmarking results and strategies for improvement. *J. Comput. Aided Mol. Des.* **2012**, *26*, 675–686. [CrossRef]

- 60. Schnecke, V.; Swanson, C.A.; Getzoff, E.D.; Tainer, J.A.; Kuhn, L.A. Screening a peptidyl database for potential ligands to proteins with side-chain flexibility. *Proteins* **1998**, *33*, 74–87. [CrossRef]
- 61. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K.; et al. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, 47, 1739–1749. [CrossRef] [PubMed]
- Halgren, T.A.; Murphy, R.B.; Friesner, R.A.; Beard, H.S.; Frye, L.L.; Pollard, W.T.; Banks, J.L. Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* 2004, 47, 1750–1759. [CrossRef] [PubMed]
- Abdel-Rahman, N.; Martinez-Arias, A.; Blundell, T.L. Probing the druggability of protein-protein interactions: Targeting the Notch1 receptor ankyrin domain using a fragment-based approach. *Biochem. Soc. Trans.* 2011, 39, 1327–1333. [CrossRef] [PubMed]
- 64. Huang, Z.; Zhu, L.; Cao, Y.; Wu, G.; Liu, X.; Chen, Y.; Wang, Q.; Shi, T.; Zhao, Y.; Wang, Y.; et al. ASD: A comprehensive database of allosteric proteins and modulators. *Nucleic Acids Res.* **2011**, *39*, D663–D669. [CrossRef] [PubMed]
- 65. Amor, B.R.; Schaub, M.T.; Yaliraki, S.N.; Barahona, M. Prediction of allosteric sites and mediating interactions through bond-to-bond propensities. *Nat. Commun.* **2016**, *7*, 12477. [CrossRef]
- Cosconati, S.; Marinelli, L.; Di Leva, F.S.; La Pietra, V.; De Simone, A.; Mancini, F.; Andrisano, V.; Novellino, E.; Goodsell, D.S.; Olson, A.J. Protein flexibility in virtual screening: The BACE-1 case study. *J. Chem. Inf. Model.* 2012, 52, 2697–2704. [CrossRef]
- 67. Fan, H.; Irwin, J.J.; Sali, A. Virtual ligand screening against comparative protein structure models. *Methods Mol. Biol.* **2011**, *819*, 105–126.
- 68. Xie, L.; Bourne, P.E. Structure-based systems biology for analyzing off-target binding. *Curr. Opin. Struct. Biol.* **2011**, *21*, 189–199. [CrossRef]
- 69. Chartier, M.; Morency, L.P.; Zylber, M.I.; Najmanovich, R.J. Large-scale detection of drug off-targets: Hypotheses for drug repurposing and understanding side-effects. *BMC Pharm. Toxicol.* **2017**, *18*, 18. [CrossRef]
- 70. Trofimov, V.A.; Varentsova, S.A. Essential Limitations of the Standard THz TDS Method for Substance Detection and Identification and a Way of Overcoming Them. *Sensors* **2016**, *16*, 502. [CrossRef]
- Law, S.; Panwar, P.; Li, J.; Aguda, A.H.; Jamroz, A.; Guido, R.V.C.; Bromme, D. A composite docking approach for the identification and characterization of ectosteric inhibitors of cathepsin K. *PLoS ONE* 2017, *12*, e0186869. [CrossRef] [PubMed]
- 72. Pabon, N.A.; Xia, Y.; Estabrooks, S.K.; Ye, Z.; Herbrand, A.K.; Suss, E.; Biondi, R.M.; Assimon, V.A.; Gestwicki, J.E.; Brodsky, J.L.; et al. Predicting protein targets for drug-like compounds using transcriptomics. *PLoS Comput. Biol.* **2018**, *14*, e1006651. [CrossRef]
- 73. Bull, S.C.; Doig, A.J. Properties of protein drug target classes. *PLoS ONE* **2015**, *10*, e0117955. [CrossRef] [PubMed]
- 74. Bakker, A.; Caricasole, A.; Gaviraghi, G.; Pollio, G.; Robertson, G.; Terstappen, G.C.; Salerno, M.; Tunici, P. How to achieve confidence in drug discovery and development: Managing risk (from a reductionist to a holistic approach). *ChemMedChem* **2009**, *4*, 923–933. [CrossRef] [PubMed]
- 75. Lee, J.Y.; Krieger, J.M.; Li, H.; Bahar, I. Pharmmaker: Pharmacophore modeling and hit identification based on druggability simulations. *Protein Sci.* **2019**, *29*, 76–86. [CrossRef] [PubMed]
- 76. Zhou, Y.; Hou, Y.; Shen, J.; Huang, Y.; Martin, W.; Cheng, F. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov.* **2020**, *6*, 14. [CrossRef]
- 77. Jain, S.S.; Luiken, J.J.; Snook, L.A.; Han, X.X.; Holloway, G.P.; Glatz, J.F.; Bonen, A. Fatty acid transport and transporters in muscle are critically regulated by Akt2. *FEBS Lett.* **2015**, *589*, 2769–2775. [CrossRef]
- 78. Zhu, J.; Wang, M.; He, J.; Zhu, M.; Wang, J.C.; Jin, L.; Wang, X.F.; Yang, Y.J.; Xiang, J.Q.; Wei, Q. Polymorphisms in the AKT1 and AKT2 genes and oesophageal squamous cell carcinoma risk in an Eastern Chinese population. *J. Cell. Mol. Med.* **2016**, *20*, 666–677. [CrossRef]
- 79. Liao, P.; Wang, W.; Li, Y.; Wang, R.; Jin, J.; Pang, W.; Chen, Y.; Shen, M.; Wang, X.; Jiang, D.; et al. Palmitoylated SCP1 is targeted to the plasma membrane and negatively regulates angiogenesis. *Elife* **2017**, *6*, e22058. [CrossRef]
- 80. Hellesoy, M.; Lorens, J.B. Cellular context-mediated Akt dynamics regulates MAP kinase signaling thresholds during angiogenesis. *Mol. Biol. Cell* **2015**, *26*, 2698–2711. [CrossRef]

- Huang, J.J.; Shi, Y.Q.; Li, R.L.; Hu, A.; Lu, Z.Y.; Weng, L.; Wang, S.Q.; Han, Y.P.; Zhang, L.; Li, B.; et al. Angiogenesis effect of therapeutic ultrasound on HUVECs through activation of the PI3K-Akt-eNOS signal pathway. *Am. J. Transl. Res.* 2015, *7*, 1106–1115. [PubMed]
- 82. Tandon, M.; Chen, Z.; Othman, A.H.; Pratap, J. Role of Runx2 in IGF-1Rbeta/Akt- and AMPK/Erk-dependent growth, survival and sensitivity towards metformin in breast cancer bone metastasis. *Oncogene* **2016**, *35*, 4730–4740. [CrossRef]
- Tandon, M.; Chen, Z.; Pratap, J. Role of Runx2 in crosstalk between Mek/Erk and PI3K/Akt signaling in MCF-10A cells. J. Cell. Biochem. 2014, 115, 2208–2217. [CrossRef] [PubMed]
- 84. Tandon, M.; Chen, Z.; Pratap, J. Runx2 activates PI3K/Akt signaling via mTORC2 regulation in invasive breast cancer cells. *Breast Cancer Res.* **2014**, *16*, R16. [CrossRef] [PubMed]
- 85. Tandon, M.; Othman, A.H.; Ashok, V.; Stein, G.S.; Pratap, J. The role of Runx2 in facilitating autophagy in metastatic breast cancer cells. *J. Cell. Physiol.* **2017**, 233, 559–571. [CrossRef] [PubMed]
- 86. Lee, I.I.; Maniar, K.; Lydon, J.P.; Kim, J.J. Akt regulates progesterone receptor B-dependent transcription and angiogenesis in endometrial cancer cells. *Oncogene* **2016**, *35*, 5191–5201. [CrossRef]
- Gao, X.; Lowry, P.R.; Zhou, X.; Depry, C.; Wei, Z.; Wong, G.W.; Zhang, J. PI3K/Akt signaling requires spatial compartmentalization in plasma membrane microdomains. *Proc. Natl. Acad. Sci. USA* 2011, 108, 14509–14514. [CrossRef]
- Krasnov, G.S.; Puzanov, G.A.; Afanasyeva, M.A.; Dashinimaev, E.B.; Vishnyakova, K.S.; Beniaminov, A.D.; Adzhubei, A.A.; Kondratieva, T.T.; Yegorov, Y.E.; Senchenko, V.N. Tumor suppressor properties of the small C-terminal domain phosphatases in non-small cell lung cancer. *Biosci. Rep.* 2019, 39, BSR20193094. [CrossRef]
- 89. Dai, M.; Al-Odaini, A.A.; Arakelian, A.; Rabbani, S.A.; Ali, S.; Lebrun, J.J. A novel function for p21Cip1 and acetyltransferase p/CAF as critical transcriptional regulators of TGFbeta-mediated breast cancer cell migration and invasion. *Breast Cancer Res.* **2012**, *14*, R127. [CrossRef]
- 90. Lin, Y.C.; Lu, L.T.; Chen, H.Y.; Duan, X.; Lin, X.; Feng, X.H.; Tang, M.J.; Chen, R.H. SCP phosphatases suppress renal cell carcinoma by stabilizing PML and inhibiting mTOR/HIF signaling. *Cancer Res.* **2014**, *74*, 6935–6946. [CrossRef]
- Khan, M.A.; Tania, M.; Wei, C.; Mei, Z.; Fu, S.; Cheng, J.; Xu, J.; Fu, J. Thymoquinone inhibits cancer metastasis by downregulating TWIST1 expression to reduce epithelial to mesenchymal transition. *Oncotarget* 2015, 6, 19580–19591. [CrossRef] [PubMed]
- 92. Kim, Y. Emerging Roles of CTD Phosphatases. J. Life Sci. 2017, 27, 370–381. [CrossRef]
- Yeo, M.; Lin, P.S. Functional characterization of small CTD phosphatases. *Methods Mol. Biol.* 2007, 365, 335–346. [PubMed]
- 94. Thompson, J.; Lepikhova, T.; Teixido-Travesa, N.; Whitehead, M.A.; Palvimo, J.J.; Janne, O.A. Small carboxyl-terminal domain phosphatase 2 attenuates androgen-dependent transcription. *EMBO J.* **2006**, *25*, 2757–2767. [CrossRef] [PubMed]
- 95. Kloet, D.E.; Polderman, P.E.; Eijkelenboom, A.; Smits, L.M.; van Triest, M.H.; van den Berg, M.C.; Groot Koerkamp, M.J.; van Leenen, D.; Lijnzaad, P.; Holstege, F.C.; et al. FOXO target gene CTDSP2 regulates cell cycle progression through Ras and p21(Cip1/Waf1). *Biochem. J.* **2015**, *469*, 289–298. [CrossRef]
- 96. Kim, Y.; Gentry, M.S.; Harris, T.E.; Wiley, S.E.; Lawrence, J.C., Jr.; Dixon, J.E. A conserved phosphatase cascade that regulates nuclear membrane biogenesis. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 6596–6601. [CrossRef]
- Ghosh, A.; Shuman, S.; Lima, C.D. The structure of Fcp1, an essential RNA polymerase II CTD phosphatase. *Mol. Cell* 2008, 32, 478–490. [CrossRef]
- Zhang, M.; Liu, J.; Kim, Y.; Dixon, J.E.; Pfaff, S.L.; Gill, G.N.; Noel, J.P.; Zhang, Y. Structural and functional analysis of the phosphoryl transfer reaction mediated by the human small C-terminal domain phosphatase, Scp1. *Protein Sci.* 2010, *19*, 974–986. [CrossRef]
- Almo, S.C.; Bonanno, J.B.; Sauder, J.M.; Emtage, S.; Dilorenzo, T.P.; Malashkevich, V.; Wasserman, S.R.; Swaminathan, S.; Eswaramoorthy, S.; Agarwal, R.; et al. Structural genomics of protein phosphatases. *J. Struct. Funct. Genom.* 2007, *8*, 121–140. [CrossRef]
- 100. Schwer, B.; Ghosh, A.; Sanchez, A.M.; Lima, C.D.; Shuman, S. Genetic and structural analysis of the essential fission yeast RNA polymerase II CTD phosphatase Fcp1. *RNA* **2015**, *21*, 1135–1146. [CrossRef]

- 101. Yun, J.H.; Ko, S.; Lee, C.K.; Cheong, H.K.; Cheong, C.; Yoon, J.B.; Lee, W. Solution structure and Rpn1 interaction of the UBL domain of human RNA polymerase II C-terminal domain phosphatase. *PLoS ONE* 2013, *8*, e62981. [CrossRef] [PubMed]
- 102. Guo, X.; Engel, J.L.; Xiao, J.; Tagliabracci, V.S.; Wang, X.; Huang, L.; Dixon, J.E. UBLCP1 is a 26S proteasome phosphatase that regulates nuclear proteasome activity. *Proc. Natl. Acad. Sci. USA* 2011, 108, 18649–18654. [CrossRef] [PubMed]
- McConnell, J.L.; Wadzinski, B.E. Targeting protein serine/threonine phosphatases for drug development. *Mol. Pharm.* 2009, 75, 1249–1261. [CrossRef] [PubMed]
- 104. Prabhu, S.V.; Singh, S.K. Identification of Potential Dual Negative Allosteric Modulators of Group I mGluR Family: A Shape Based Screening, ADME Prediction, Induced Fit Docking and Molecular Dynamics Approach Against Neurodegenerative Diseases. *Curr. Top. Med. Chem.* 2019, 19, 2687–2707. [CrossRef] [PubMed]
- 105. SarathKumar, B.; Lakshmi, B.S. In silico investigations on the binding efficacy and allosteric mechanism of six different natural product compounds towards PTP1B inhibition through docking and molecular dynamics simulations. *J. Mol. Model.* **2019**, *25*, 272. [CrossRef]
- 106. Ivasiv, V.; Albertini, C.; Goncalves, A.E.; Rossi, M.; Bolognesi, M.L. Molecular Hybridization as a Tool for Designing Multitarget Drug Candidates for Complex Diseases. *Curr. Top. Med. Chem.* 2019, 19, 1694–1711. [CrossRef]
- 107. Wu, Y.; Jiang, S.; Ying, T. From therapeutic antibodies to chimeric antigen receptors (CARs): Making better CARs based on antigen-binding domain. *Expert Opin. Biol.* **2016**, *16*, 1469–1478. [CrossRef]
- 108. Fraczek, J.; Vanhaecke, T.; Rogiers, V. Toxicological and metabolic considerations for histone deacetylase inhibitors. *Expert Opin. Drug Metab. Toxicol.* **2013**, *9*, 441–457. [CrossRef]
- 109. Kumar, S.; Tiwari, M. Variable selection based QSAR modeling on Bisphenylbenzimidazole as Inhibitor of HIV-1 reverse transcriptase. *Med. Chem.* **2012**, *9*, 955–967. [CrossRef]
- 110. Faivre, S.; Sablin, M.P.; Dreyer, C.; Raymond, E. Novel anticancer agents in clinical trials for well-differentiated neuroendocrine tumors. *Endocrinol. Metab. Clin. N. Am.* **2010**, *39*, 811–826. [CrossRef]
- 111. Selz, K.A.; Samoylova, T.I.; Samoylov, A.M.; Vodyanoy, V.J.; Mandell, A.J. Designing allosteric peptide ligands targeting a globular protein. *Biopolymers* **2007**, *85*, 38–59. [CrossRef] [PubMed]
- Stains, C.I.; Mondal, K.; Ghosh, I. Molecules that target beta-amyloid. *ChemMedChem* 2007, 2, 1674–1692. [CrossRef] [PubMed]
- Urban, J.D.; Clarke, W.P.; von Zastrow, M.; Nichols, D.E.; Kobilka, B.; Weinstein, H.; Javitch, J.A.; Roth, B.L.; Christopoulos, A.; Sexton, P.M.; et al. Functional selectivity and classical concepts of quantitative pharmacology. *J. Pharm. Exp.* 2007, 320, 1–13. [CrossRef] [PubMed]
- Paul, M.; Leibovici, L. Combination antimicrobial treatment versus monotherapy: The contribution of meta-analyses. *Infect. Dis. Clin. N. Am.* 2009, 23, 277–293. [CrossRef]
- 115. Wang, Y.; Qi, W.; Zhang, L.; Ying, Z.; Sha, O.; Li, C.; Lu, L.; Chen, X.; Li, Z.; Niu, F.; et al. The novel targets of DL-3-n-butylphthalide predicted by similarity ensemble approach in combination with molecular docking study. *Quant. Imaging Med. Surg.* 2017, 7, 532–536. [CrossRef]
- 116. Ramirez, D.; Caballero, J. Is It Reliable to Use Common Molecular Docking Methods for Comparing the Binding Affinities of Enantiomer Pairs for Their Protein Target? *Int. J. Mol. Sci.* **2016**, *17*, 525. [CrossRef]
- Goldhirsch, A.; Coates, A.S.; Gelber, R.D.; Glick, J.H.; Thurlimann, B.; Senn, H.J. First—Select the target: Better choice of adjuvant treatments for breast cancer patients. *Ann. Oncol.* 2006, 17, 1772–1776. [CrossRef]
- 118. Evers, A.; Klabunde, T. Structure-based drug discovery using GPCR homology modeling: Successful virtual screening for antagonists of the alpha1A adrenergic receptor. J. Med. Chem. 2005, 48, 1088–1097. [CrossRef]
- 119. Arrondeau, J.; Huillard, O.; Tlemsani, C.; Cessot, A.; Boudou-Rouquette, P.; Blanchet, B.; Thomas-Schoemann, A.; Vidal, M.; Tigaud, J.M.; Durand, J.P.; et al. Investigational therapies up to Phase II which target PDGF receptors: Potential anti-cancer therapeutics. *Expert Opin. Investig. Drugs* 2015, 24, 673–687. [CrossRef]
- Chaput, L.; Martinez-Sanz, J.; Quiniou, E.; Rigolet, P.; Saettel, N.; Mouawad, L. vSDC: A method to improve early recognition in virtual screening when limited experimental resources are available. *J. Cheminform.* 2016, *8*, 1. [CrossRef]

- 121. Manning, A.; Highland, H.M.; Gasser, J.; Sim, X.; Tukiainen, T.; Fontanillas, P.; Grarup, N.; Rivas, M.A.; Mahajan, A.; Locke, A.E.; et al. A Low-Frequency Inactivating AKT2 Variant Enriched in the Finnish Population Is Associated With Fasting Insulin Levels and Type 2 Diabetes Risk. *Diabetes* 2017, 66, 2019–2032. [CrossRef] [PubMed]
- 122. Uehara, S.; Tanaka, S. Cosolvent-Based Molecular Dynamics for Ensemble Docking: Practical Method for Generating Druggable Protein Conformations. *J. Chem. Inf. Model.* **2017**, *57*, 742–756. [CrossRef] [PubMed]
- 123. Dube, M.P.; de Denus, S.; Tardif, J.C. Pharmacogenomics to Revive Drug Development in Cardiovascular Disease. *Cardiovasc. Drugs* **2016**, *30*, 59–64. [CrossRef] [PubMed]
- Lill, M.A.; Danielson, M.L. Computer-aided drug design platform using PyMOL. J. Comput. Aided Mol. Des. 2010, 25, 13–19. [CrossRef]
- 125. Mbonye, A.K.; Magnussen, P.; Chandler, C.I.; Hansen, K.S.; Lal, S.; Cundill, B.; Lynch, C.A.; Clarke, S.E. Introducing rapid diagnostic tests for malaria into drug shops in Uganda: Design and implementation of a cluster randomized trial. *Trials* 2014, *15*, 303. [CrossRef]
- 126. Boucherit, H.; Chikhi, A.; Bensegueni, A.; Merzoug, A.; Bolla, J.M. The Research of New Inhibitors of Bacterial Methionine Aminopeptidase by Structure Based Virtual Screening Approache OF ZINC DATABASE and in vitro Validation. *Curr. Comput. Aided Drug Des.* **2019**, (Epub ahead of print). [CrossRef]
- 127. Verdonk, M.L.; Cole, J.C.; Hartshorn, M.J.; Murray, C.W.; Taylor, R.D. Improved protein-ligand docking using GOLD. *Proteins* 2003, *52*, 609–623. [CrossRef]
- Verdonk, M.L.; Chessari, G.; Cole, J.C.; Hartshorn, M.J.; Murray, C.W.; Nissink, J.W.; Taylor, R.D.; Taylor, R. Modeling water molecules in protein-ligand docking using GOLD. J. Med. Chem. 2005, 48, 6504–6515. [CrossRef]
- 129. Zhang, H.; Liao, L.; Saravanan, K.M.; Yin, P.; Wei, Y. DeepBindRG: A deep learning based method for estimating effective protein-ligand affinity. *Peer J.* **2019**, *7*, e7362. [CrossRef]
- Torres, P.H.M.; Sodero, A.C.R.; Jofily, P.; Silva-Jr, F.P. Key Topics in Molecular Docking for Drug Design. Int. J. Mol. Sci. 2019, 20, 4574. [CrossRef]
- Lindh, M.; Svensson, F.; Schaal, W.; Zhang, J.; Skold, C.; Brandt, P.; Karlen, A. Toward a Benchmarking Data Set Able to Evaluate Ligand- and Structure-based Virtual Screening Using Public HTS Data. *J. Chem. Inf. Model.* 2015, 55, 343–353. [CrossRef] [PubMed]
- De Smet, F.; Christopoulos, A.; Carmeliet, P. Allosteric targeting of receptor tyrosine kinases. *Nat. Biotechnol.* 2014, 32, 1113–1120. [CrossRef] [PubMed]
- 133. Chio, C.M.; Lim, C.S.; Bishop, A.C. Targeting a cryptic allosteric site for selective inhibition of the oncogenic protein tyrosine phosphatase Shp2. *Biochemistry* **2014**, *54*, 497–504. [CrossRef] [PubMed]
- 134. Hou, X.; Rooklin, D.; Yang, D.; Liang, X.; Li, K.; Lu, J.; Wang, C.; Xiao, P.; Zhang, Y.; Sun, J.P.; et al. Computational Strategy for Bound State Structure Prediction in Structure-Based Virtual Screening: A Case Study of Protein Tyrosine Phosphatase Receptor Type O Inhibitors. *J. Chem. Inf. Model.* 2018, *58*, 2331–2342. [CrossRef] [PubMed]
- Reddy, R.H.; Kim, H.; Cha, S.; Lee, B.; Kim, Y.J. Structure-Based Virtual Screening of Protein Tyrosine Phosphatase Inhibitors: Significance, Challenges, and Solutions. J. Microbiol. Biotechnol. 2017, 27, 878–895. [PubMed]
- 136. Lamba, V.; Ghosh, I. New directions in targeting protein kinases: Focusing upon true allosteric and bivalent inhibitors. *Curr. Pharm. Des.* **2012**, *18*, 2936–2945. [CrossRef]
- Cossins, B.P.; Lawson, A.D. Small Molecule Targeting of Protein-Protein Interactions through Allosteric Modulation of Dynamics. *Molecules* 2015, 20, 16435–16445. [CrossRef]
- 138. Grigoriadis, D.E.; Hoare, S.R.; Lechner, S.M.; Slee, D.H.; Williams, J.A. Drugability of extracellular targets: Discovery of small molecule drugs targeting allosteric, functional, and subunit-selective sites on GPCRs and ion channels. *Neuropsychopharmacology* **2009**, *34*, 106–125. [CrossRef]
- 139. Hogg, P.J. Targeting allosteric disulphide bonds in cancer. Nat. Rev. Cancer 2013, 13, 425–431. [CrossRef]
- Jadaun, A.; Subbarao, N.; Dixit, A. Allosteric inhibition of topoisomerase I by pinostrobin: Molecular docking, spectroscopic and topoisomerase I activity studies. J. Photochem. Photobiol. B 2017, 167, 299–308. [CrossRef]
- 141. Yang, L.; Shi, P.; Zhao, G.; Xu, J.; Peng, W.; Zhang, J.; Zhang, G.; Wang, X.; Dong, Z.; Chen, F.; et al. Targeting cancer stem cell pathways for cancer therapy. *Signal Transduct. Target. Therapy* 2020, *5*, 8. [CrossRef] [PubMed]

- 142. Wu, Q.; Jiang, L.; Li, S.C.; He, Q.J.; Yang, B.; Cao, J. Small molecule inhibitors targeting the PD-1/PD-L1 signaling pathway. *Acta Pharm. Sin.* **2020**, 10. [CrossRef] [PubMed]
- 143. Gu, X.; Wang, Y.; Wang, M.; Wang, J.; Li, N. Computational investigation of imidazopyridine analogs as protein kinase B (Akt1) allosteric inhibitors by using 3D-QSAR, molecular docking and molecular dynamics simulations. *J. Biomol. Struct. Dyn.* **2019**, 1–16. [CrossRef] [PubMed]
- 144. Sheik Amamuddy, O.; Veldman, W.; Manyumwa, C.; Khairallah, A.; Agajanian, S.; Oluyemi, O.; Verkhivker, G.; Tastan Bishop, O. Integrated Computational Approaches and Tools forAllosteric Drug Discovery. *Int. J. Mol. Sci.* 2020, *21*, 847. [CrossRef]
- 145. Zhang, W.; Li, R.; Shin, R.; Wang, Y.; Padmalayam, I.; Zhai, L.; Krishna, N.R. Identification of the binding site of an allosteric ligand using STD-NMR, docking, and CORCEMA-ST calculations. *ChemMedChem* 2013, *8*, 1629–1633. [CrossRef]
- 146. Ramirez, U.D.; Myachina, F.; Stith, L.; Jaffe, E.K. Docking to large allosteric binding sites on protein surfaces. *Adv. Exp. Med. Biol.* **2010**, *680*, 481–488.
- 147. Li, S.; Zhang, J.; Lu, S.; Huang, W.; Geng, L.; Shen, Q. The mechanism of allosteric inhibition of protein tyrosine phosphatase 1B. *PLoS ONE* **2014**, *9*, e97668. [CrossRef]
- Craig, I.R.; Essex, J.W.; Spiegel, K. Ensemble docking into multiple crystallographically derived protein structures: An evaluation based on the statistical analysis of enrichments. *J. Chem. Inf. Model.* 2010, 50, 511–524. [CrossRef]
- 149. Rueda, M.; Bottegoni, G.; Abagyan, R. Recipes for the selection of experimental protein conformations for virtual screening. *J. Chem. Inf. Model.* 2009, *50*, 186–193. [CrossRef]
- 150. Rao, S.; Sanschagrin, P.C.; Greenwood, J.R.; Repasky, M.P.; Sherman, W.; Farid, R. Improving database enrichment through ensemble docking. *J. Comput. Aided Mol. Des.* **2008**, *22*, 621–627. [CrossRef]
- Campbell, A.J.; Lamb, M.L.; Joseph-McCarthy, D. Ensemble-based docking using biased molecular dynamics. J. Chem. Inf. Model. 2014, 54, 2127–2138. [CrossRef] [PubMed]
- 152. Polgar, T.; Keseru, G.M. Ensemble docking into flexible active sites. Critical evaluation of FlexE against JNK-3 and beta-secretase. *J. Chem. Inf. Model.* **2006**, *46*, 1795–1805. [CrossRef] [PubMed]
- 153. Rallabandi, H.R.; Lee, D.; Sung, J.; Kim, Y.J. Peripheral Inhibition of Small C-terminal Domain Phosphatase 1 with Napthoquinone Analogues. *Bull. Korean Chem. Soc.* **2020**, in press.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).