

Identification of small molecule binding pocket for inhibition of Crimean–Congo hemorrhagic fever virus OTU protease

Fatih KOCABAŞ^{1,*}, Enes Kemal ERGİN²

¹Department of Genetics and Bioengineering, Faculty of Engineering, Yeditepe University, İstanbul, Turkey

²Department of Computer Science, North American University, Houston, TX, USA

Received: 16.01.2015 • Accepted/Published Online: 29.07.2015 • Final Version: 05.01.2016

Abstract: Crimean–Congo hemorrhagic fever virus (CCHFV) is a deadly tick-borne virus with high mortality rates. Current antivirals lack specificity, making them susceptible to off-target effects and cytotoxicity. There is an utmost need for the identification of active compounds for anti-CCHFV therapies. Inhibition of CCHFV ovarian tumor (OTU) protease by small molecules is an exciting potential antiviral therapy. In this study, computational approaches based on residue homology, the binding coordinates of ligands, and correlation analysis with in vitro data identified the pocket of Y89–W99 as the inhibition site of CCHFV OTU protease. In silico screening of more than 600,000 compounds against this newly discovered pocket can identify potent inhibitors of CCHFV OTU protease. This novel set of compounds exhibits a common substructure and higher binding affinities. These findings distinguish the pocket of Y89–W99 as a pharmaceutical target for the optimization and identification of CCHFV OTU protease inhibitors that could serve as lead structures for discovering therapies against CCHFV.

Key words: Viral OTU protease, docking, in silico screening, virtual screening, CCHFV, CCHF

1. Introduction

Crimean–Congo hemorrhagic fever virus (CCHFV) is a deadly tick-borne virus belonging to the family Bunyaviridae, genus Nairovirus. It constitutes a public health threat due to its high mortality rate, up to 6% in hospitalized patients and up to 30%–50% in severe forms (Whitehouse, 2004; Ozkaya et al., 2010; Ergunay et al., 2011; Bente et al., 2013). Since its first report in Crimea in 1944, there have been outbreaks in Africa, Asia, and Eastern Europe (Gargili et al., 2011; Ergonul, 2012; Dokuzoguz et al., 2013). Because infected ticks have been shown to travel through migratory birds to different parts of the world, this raises a concern regarding general healthcare in various parts of the world (Gale et al., 2012; Lindeborg et al., 2012; Palomar et al., 2013). There is currently no therapeutic treatment or commercial vaccine available for CCHFV. Due to its short incubation time (average 3–5 days), high viremia that lasts around 10 days, and associated severe hemorrhages, it can lead to shock and death in 1–2 weeks. This high viral activity is partly due to an invasion mechanism that shuts down antiviral response pathways through a unique CCHFV ovarian tumor (OTU) protease, which has broad deconjugation activities including deubiquitination (Arguello and Hiscott, 2007; Frias-Staheli et al., 2007; Weber and Mirazimi, 2008).

The CCHFV OTU domain is expressed from the L segment of the CCHFV. CCHFV OTU protease takes an important role in viral invasion through antagonizing NF- κ B signaling (Frias-Staheli et al., 2007; van Kasteren et al., 2012). The inhibition of antiviral activity by CCHFV OTU occurs by the removal of ubiquitin (UB) and interferon-stimulated gene products 15 (ISG15), modifications occurring in a broad manner in the infected cells in contrast to target-specific OTU containing mammalian proteases (Frias-Staheli et al., 2007; Malakhova and Zhang, 2008; Akutsu et al., 2011; Capodagli et al., 2011, 2013). UB and ISG15 share a conserved C terminus motif 'LRLRGG' that is recognized by CCHFV OTU protease (Akutsu et al., 2011; Capodagli et al., 2011).

The crystal structures of the CCHFV OTU protease bound with human UB and with ISG15 have been resolved and deposited in the protein data bank (3PRP and 3PHX, respectively) (Capodagli et al., 2011; James et al., 2011). A key similarity between the two structures is the binding of UB and ISG15 towards a similar pocket of amino acid residues. Mutation studies in several amino acids in close proximity to this pocket suggest that it could be a pharmaceutical target for CCHFV OTU inhibitors (Frias-Staheli et al., 2007). There are currently no CCHFV OTU inhibitors available for use. The PubChem database reports

* Correspondence: fatih.kocabas@yeditepe.edu.tr

a couple of preliminary studies with compounds having a potentially inhibitory effect on OTU protease. Previous studies on the alignment of a limited number of related and unrelated protein datasets suggested the conservation of amino acids D37, G38, N39, C40, W71, and H151, with the prediction of C40 and H151 as catalytic residues in putative OTU protease (Snijder et al., 1994, 1995; Makarova et al., 2000; Nanao et al., 2004; Frias-Staheli et al., 2007). However, a robust approach to determine the conserved residues that are important for the inhibition of CCHFV OTU protease and to utilize in silico screening to identify CCHFV OTU inhibitors and correlate their affinities with these proposed conserved residues was needed.

Computational chemistry for in silico screening is now an established platform for identification of lead compounds in the drug discovery process (Yang et al., 2007; Sudha et al., 2008; Chang et al., 2010; Wassman et al., 2013). This approach utilizes the prediction of binding poses and affinities of each docked compound to the crystal structure of a target protein. This requires knowledge of highly conserved residues of the protein of interest. In this study, we used computational approaches based on residue homology, the binding coordinates of ligands, and correlation analysis with in vitro data to determine the conserved residues of CCHFV OTU protease. We identified a previously unknown pocket of conserved residues of CCHFV OTU, which provides a tool for lead compound optimization and discovery. We performed an in silico screening with compounds up to CID_1000000 from the PubChem database (Han et al., 2008; Wang et al., 2011) on the newly revealed pocket of conserved residues using the advanced docking tool AutoDock Vina (Trott and Olson, 2009). We identified a set of compounds with 75% similar substructures. In addition, we identified forty-nine CCHFV OTU-related proteins that share a similar pocket of conserved amino acids that could be the target of future drug discovery to combat viral and nonviral OTU proteases and corresponding pathogens.

2. Materials and methods

2.1. CCHFV OTU protein alignments

An amino acid sequence of the CCHFV OTU was blasted to determine the OTU-related proteins in nonredundant protein sequences (nr) (all nonredundant GenBank CDS translations + PDB + SwissProt + PIR + PRF) using a blastp (protein-protein BLAST) algorithm (NCBI). A FASTA sequence of the related protein hits was collected to perform an amino acid alignment and to determine the conserved amino acid residues in CCHFV OTU protease. Multiple alignments of CCHFV OTU related protein amino acid sequences were performed using a constraint-based multiple alignment tool (COBALT) (NCBI). The accession numbers of the CCHFV OTU protease-related

protein sequences are provided in Figure S1 (see links at the end of the manuscript).

2.2. Conserved binding pocket analysis of CCHFV OTU protease by molecular docking

The 3D coordinates of the crystal structure of OTU at a resolution of 1.7 Å in complex with UB (PDB code: 3PRP) were selected as the receptor model. Using AutoDockTools (Trott and Olson, 2009), pockets of D37, C40, and H151 residues and a pocket of Y89–W99 residues were highlighted to determine the grid box locations with a 22–20–20 Å search space. The 3D SDF files of the CCHFV OTU inhibitors identified in vitro and reported in a primary screen were downloaded from PubChem (Bioassay AID: 686976 and partially AID: 651958). Docking studies were performed using AutoDock Vina 1.1.2 (Trott and Olson, 2009) and automated using PaDEL-ADV (Sanner et al., 1996; Sanner, 1999). In addition, a conserved binding pocket correlation analysis was repeated using the in silico-identified set of compounds as described.

2.3. In silico screening

The 3D SDF files for all of the small molecules in this paper were downloaded from PubChem (Wang et al., 2011). The OTU protein was prepared using AutoDockTools for docking studies, which is a module of the MGLTools. The OTU protein (PDB code: 3PRP) was downloaded from pdb.org in PDB format (text). Human UB and water molecules were removed from the 3PRP and then converted to PDBQT format using AutoDockTools. Only compounds with 3D SDF structures provided by PubChem were used for in silico screening. Small molecules (ligands) up to CID_1000000 were downloaded from PubChem (in 3D SDF format) and were ready for docking studies. The docking studies were performed using AutoDock Vina 1.1.2 with a 22–20–20 Å search space centered on the pocket of Y89–W99 residues. In silico screening was automated using PaDEL-ADV with 16 CPUs. The maximum predicted binding affinity (lowest binding energy of –9.2 kcal/mol) of the putative inhibitors identified in vitro was used as the cut-off value. The lower the binding affinity, the stronger the ligand binds to the protein. The compounds with a binding energy of at least –9.2 kcal/mol were rescreened for the whole OTU surface with a 50–50–50 Å search space, and a pocket of the D37, C40, and H151 residues with a 22–20–20 Å search space. Compounds with a binding energy of at least –9.2 kcal/mol for both the whole OTU surface and the pocket of Y89–W99 were determined. OTU–ligand diagrams were prepared using AutoDockTools (MGL Tools, Scripps Research Institute).

2.4. Cluster analysis

The top compounds were clustered according to structural similarity (2D and 3D). Compounds with PubChem IDs

were submitted to the Chemical Structure Clustering Tool (PubChem, NCBI) and compounds with scores of >0.7 similarities were grouped into clusters.

2.5. Statistical analysis

Conserved binding pocket correlation analyses were performed using Student's t test, Pearson's R, and the Spearman correlation. All statistical calculations were done in Microsoft Excel and confirmed in MATLAB equipped with a statistics toolbox. Statistical calculations are provided in Supplementary Tables 1, 2, and 3. Values with $P < 0.05$ were considered statistically significant.

3. Results

3.1. Multiple alignment of CCHFV OTU-related proteins identifies highly conserved residues of OTU protease

In the last decade, the pool of nucleotide sequences available from the NCBI has rapidly increased, which has made it possible to determine the conserved residues of CCHFV OTU protease in related viral and nonviral proteins with better resolution. In this study, we performed a basic protein blast of a CCHFV OTU amino acid sequence using an NCBI blastp algorithm and found that there are fifty CCHFV OTU-related proteins. Doing a multiple alignment of these fifty OTU related viral and nonviral proteins using a COBALT algorithm, we determined the previously undiscovered conserved residues Y89 and W99 with 100% identity and G100 with 96% identity in addition to the previously reported D37, C40, and H151 amino acid residues determined with at least 98% identity (Figures 1 and S1). However, we did not see any conservation of the previously suggested G38, N38, or F152 residues in the fifty CCHFV OTU-related viral and nonviral protein alignments (Snijder et al., 1994, 1995; Makarova et al., 2000; Nanao et al., 2004; Frias-Staheli et al., 2007).

Furthermore, we performed multiple alignments of the CCHFV OTU-related viral proteins. A blast analysis at NCBI brought only six different viral OTU-related proteins (see Figure S2a for the phylogenetic tree of OTU-related viral proteins in related viruses). This alignment clearly showed that there are three highly conserved areas. The highest conservation is in the D96–L108 residues, including the newly discovered conserved W99 and G100 residues (Figure S2b). In addition, this newly revealed pocket of Y89–W99 is located in the interaction site of UB and ISG15 with CCHFV OTU protease (Figures S3a and S3b). These findings suggest that a putative inhibition pocket of CCHFV OTU protease could be located in the pocket of the Y89, W99, and G100 residues or in the pocket of the previously suggested D37–C40 residues, or the pocket of the H151–L155 residues. Intriguingly, C40 was located in close proximity to the Y89–W99 pocket. Moreover, when we located these residues in the crystal structure of the

CCHFV OTU protease, we found that all of these three conserved areas were located in close proximity to each other; D37, C40, and H151 are facing towards the outer surface of the OTU protease, which could be considered a single pocket of D37, C40, and H151 residues (Figure 2).

3.2. Active compounds show selective affinity towards the proposed inhibition pocket

Molecular docking has been used to test which pocket of conserved residues are part of the putative protease inhibition pocket. To this end, we got the 3D protein structure of CCHFV OTU protease at a resolution of 1.7 Å from the Protein Database (pdb.org, 3PRP). Using AutoDockTools, we located the pocket of D37, C40, and H151 residues and the pocket of Y89 and W99 residues and determined the grid box locations with a 22–20–20 Å search space (Figure 2). In addition, we outlined a grid box with a 50–50–50 Å search space that allowed us to scan the whole OTU surface for the affinity and binding locations of the previously reported putative inhibitors of OTU protease from the PubChem Database. Thus, we collected the 3D SDF files of the CCHFV OTU inhibitors identified in vitro, reported in a preliminary screen from PubChem (Bioassay AID: 686976 and partially AID: 651958). Then we determined the predicted binding affinities toward 'whole surface', 'pocket of D37, C40, and H151', and 'pocket of Y89, W99, G100, and S101 residues' using the fast, and exhaustive docking software AutoDock Vina. The analysis of 346 compounds was automated using PaDEL-ADV. Using the coordinates and configurations provided in Figure 2, we performed affinity calculations. We report that the affinities toward the pocket of Y89, W99, G100, and S101 were very similar to the OTU surface maximum affinity calculations ($P = 0.13$) (Figure 3a). On the other hand, the affinities toward the pocket of D37, C40, and H151 were significantly different than the OTU surface maximum affinity calculations ($P = 3.7 \times 10^{-85}$) (Figure 3a). In addition, affinities toward the pocket of Y89, W99, G100, and S101 showed a high degree of correlation with the OTU surface maximum affinities as measured by Pearson's R ($R = 0.93$) (Figure 3a). Moreover, the binding pocket correlation analysis with the in vitro reported compounds showed that those compounds show higher affinities towards the proposed inhibition pocket (pocket of Y89, W99, G100, and S101) instead of the previously suggested catalytic residues, namely D37, C40, and H151. For instance, CID_15944915, CID_300375, and CID_17749659 demonstrate the same or highly similar affinities toward the pocket of Y89–S101 and the whole surface maximum affinity. This is also evident when we looked at where these compounds are predicted to bind in the whole OTU surface affinity analysis (Figure 3b, left column). Their predicted maximum affinity is toward the pocket of Y89, W99, G100, and S101.

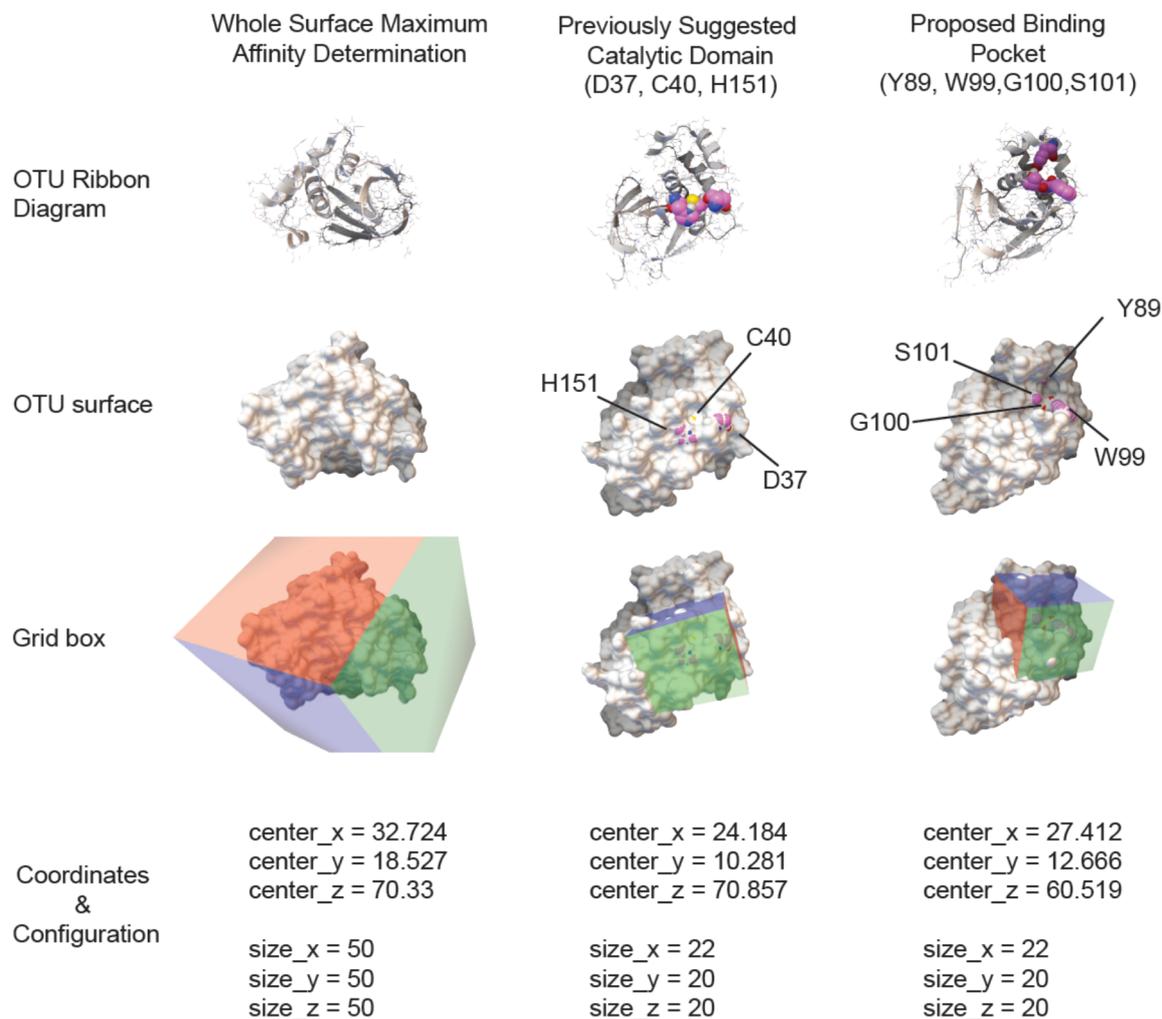


Figure 2. Analysis of the CCHFV OTU protease inhibition pocket by molecular docking. Top row = OTU ribbon diagram, where the proposed residues as catalytic residues are highlighted. Middle row = OTU surface with the positions of the conserved residues marked. Bottom rows = grid box showing where molecular docking took place in a cube, with its coordinates and configuration.

3.3. In silico screening toward the pocket of Y89, W99, G100, and S101

The precise discovery of an inhibition pocket of CCHFV OTU protease using ensemble-based docking allowed us to design an in silico screen for inhibitors of CCHFV OTU with higher affinities. In vitro screening is often costly and the failure rate is still high. Although there are hits that show promise, they have either high IC₅₀ values or cytotoxicity. The utilization of bioinformatic tools that foresee and provide lead compounds with predicted higher affinities could yield lower IC₅₀ and thus lower toxicity. Therefore, we performed an in silico screening toward the newly revealed pocket of Y89, W99, G100, and S101. The screening included compounds with 3D structures available at PubChem

up to CID_1000000. Docking studies were performed using AutoDock Vina and PaDEL-ADV with a 22–20–20 Å search space centered on the pocket of Y89 and W99 residues. We identified 313 compounds with higher affinities (lower binding energy) than previously reported (CID_15944915, binding affinity = -9.2 kcal/mol) (Supplementary Tables 3 and 4).

3.4. In silico screening identifies a novel set of compounds with higher affinity toward the pocket of Y89, W99, G100, and S101

A binding pocket correlation analysis was repeated using a set of compounds identified in silico with an affinity of at least -9.2 kcal/mol. We performed docking for the whole OTU surface with a 50–50–50 Å search space, and a pocket of D37, C40, and H151 residues with a 22–20–

A Binding Pocket Correlation Analysis with in vitro Identified Compounds

PubChem CID	Compound Name	Whole OTU Surface	Previously Suggested Active Site (D37,C40,H151)		Proposed Active Site (Y89,W99,G100,S101)	
		Max Affinity (kcal/mol)	Affinity (kcal/mol)	Affinity Difference compared to whole	Affinity (kcal/mol)	Affinity Difference compared to whole
15944915	MLS000590115	-9.2	-5.4	-3.8	-9.2	0
4608646	AC1NDXU5	-8.7	-6	-2.7	-8.8	0.1
9586236	Ambcb5944659	-8.6	-6.8	-1.8	-8.7	0.1
1935947	AC1LYX5X	-7.5	-5.4	-2.1	-8.6	1.1
300375	AC1L6W6U	-8.4	-6	-2.4	-8.4	0
660152	AC1LD2EH	-8.3	-5.8	-2.5	-8.4	0.1
6163965	AC1O22IP	-8.3	-6.2	-2.1	-8.3	0
17749659	MLS000923759	-8.2	-6.9	-1.3	-8.2	0
664517	AC1LDBSM	-7.7	-6.5	-1.2	-8.1	0.4
5341375	AC1NSXRG	-7.8	-5.8	-2	-8.1	0.3

B Representative Images of in vitro Identified Compounds Docked into whole OTU, Pocket of D37-H151 and Y89-W99

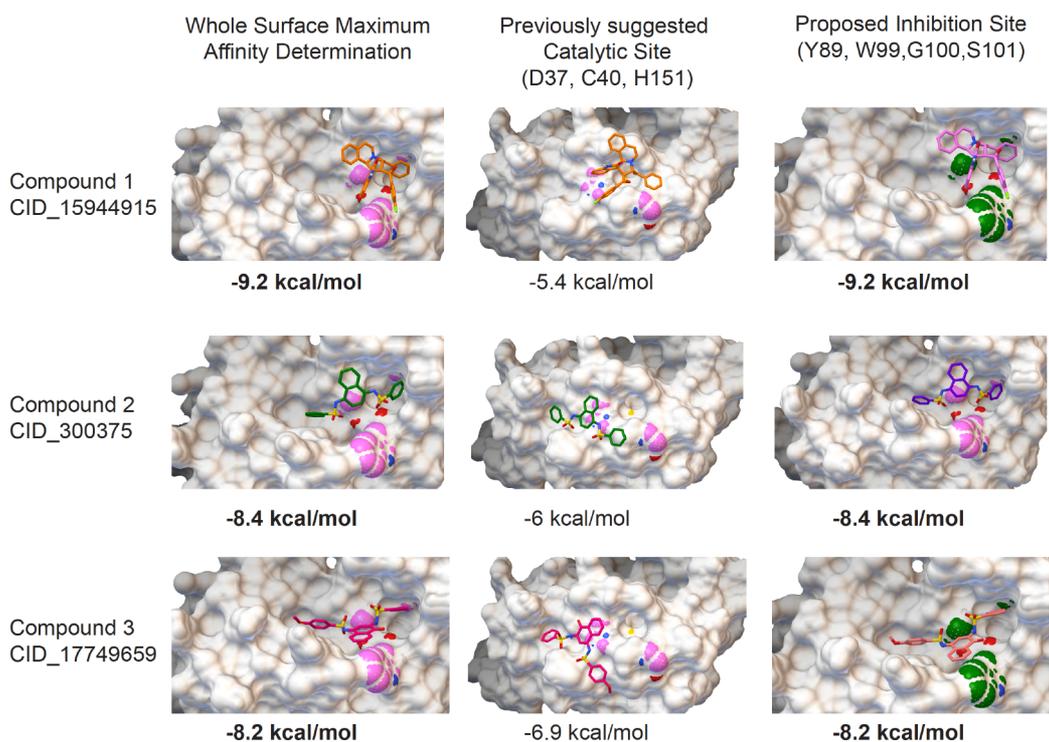


Figure 3. Binding pocket correlation analysis with compounds identified in vitro. A) Table of compounds showing affinity estimates towards the whole OTU surface, the previously suggested catalytic residues (D37, C40, and H151), and the proposed binding pocket (Y89, W99, G100, and S101). Note that the Y89, W99, G100, and S101 site affinities of the compounds are significantly similar to the whole surface maximum affinities and show a high degree of correlation ($R = 0.93$). However, affinity towards the catalytic residue (D37, C40, and H151) is significantly different ($P < 0.001$) and is lower (average difference = -2.2) compared to the whole surface maximum affinity scores. B) Representative images of the compounds identified in vitro docked into the whole OTU or the pockets of D37-H151 and Y89-W99 show that the putative inhibitors identified in vitro reported in PubChem against CCHFV OTU protease show a preference to bind to the pocket of Y89-W99, both in whole surface molecular docking and Y89-W99-G100-S100 molecular docking. See Supplementary Table 1 (see links at the end of the manuscript) for larger datasets and Supplementary Table 2 for Swissdock verification.

20 Å search space. We show that compounds with an affinity of at least -9.2 kcal/mol on the whole OTU surface docking were not significantly different than the affinities towards the proposed binding pocket of Y89, W99, G100,

and S101 ($P = 0.30$) (Figure 4a; Supplementary Tables 3 and 4). Similar to the binding pocket correlation analysis with compounds identified in vitro towards the pocket of D37, C40 and H151, compounds identified in silico show

A Binding Pocket Correlation Analysis with in silico Identified Compounds

PubChem CID	Compound Name	Whole OTU Surface	Previously Suggested Active Site (D37,C40,H151)		Proposed Active Site (Y89,W99,G100,S101)	
		Max Affinity (kcal/mol)	Affinity (kcal/mol)	Affinity Difference compared to whole	Affinity (kcal/mol)	Affinity Difference compared to whole
636081	AC1LCRNJ	-12	-8.2	3.8	-12	0
62448	Vat brown 1	-11.2	-10	1.2	-11.2	0
77186	AC1L2SIX	-10	-7.8	2.2	-11.2	-1.2
113522	AC1L3DEZ	-10.7	-7.4	3.3	-11.1	-0.4
377321*	NSC658721	-11.1	-8.6	2.5	-11.1	0
80929	EINECS 229-251-6	-11.9	-9.4	2.5	-11	0.9
388441*	NSC683367	-11	-6.8	4.2	-11	0
174177	EINECS 279-223-2	-10.8	-7.8	3	-11	-0.2
74783	EINECS 217-812-8	-10.9	-7.5	3.4	-10.9	0
77767	4118-16-5	-10.9	-6.4	4.5	-10.8	0.1

*Reported as non-cytotoxic in many bioassays

B Representative Images of in silico Identified Compounds Docked into whole OTU surface, Pocket of D37-H151 and Y89-W99

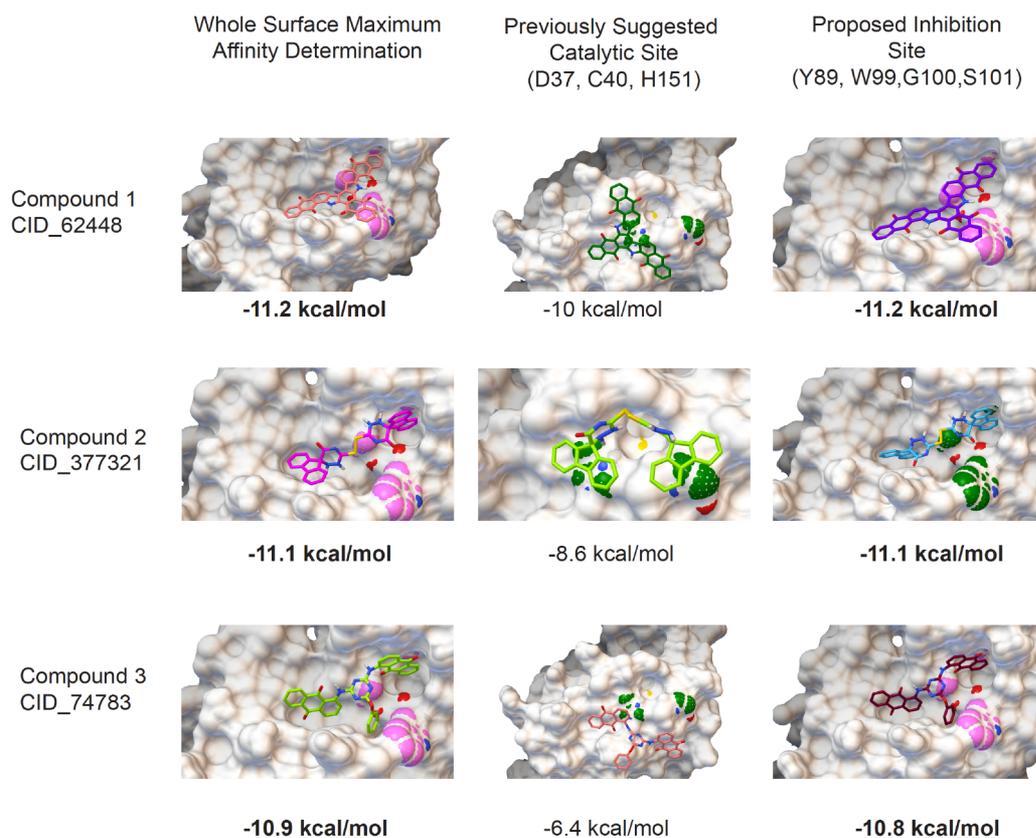


Figure 4. Binding pocket correlation analysis with compounds identified in silico. A) The set of compounds identified in silico from a virtual screen of compounds up to CID_1000000 that demonstrate preferential binding to the pocket of Y89–W99–G100–S101 residues instead of the previously proposed catalytic domains (D37, C40, and H151). See Supplementary Table 3 for larger datasets. B) Representative images of compounds identified in silico docked into the whole OTU surface, the D37–H151 pocket, and the Y89–W99 pocket. See Supplementary Table 4 for the binding affinity predictions of the compounds up to CID_1000000 tested for the Y89–W99 pocket.

significantly different affinities compared to the whole surface maximum affinities ($P = 1.5 \times 10^{-240}$). In addition, hits identified in silico demonstrated a correlated affinity between the whole surface maximum affinities and the

proposed binding pocket affinities ($R = 0.80$), while there were no correlations between the whole OTU surface maximum affinities and the previously suggested catalytic site affinities ($R = 0.46$). Moreover, hits identified in silico

demonstrated a preferential binding to the pocket of Y89, W99, G100, and S101 on the whole OTU surface docking (Figure 4b). For instance, CID_62448, CID_377321, and CID_74783 demonstrated the same affinity predictions and were found to bind to the same location (the pocket of Y89–S101), as determined by both the whole surface and the proposed pocket of Y89–S101 molecular docking. This is in contrast to the D37, C40, and H151 pocket, where affinities are largely different than the whole surface maximum affinities of the compounds CID_62448, CID_377321, and CID_74783.

3.5. Structural clustering of in silico hits identifies common substructures

We performed a structural classification of hits identified in silico using 2D and 3D Tanimoto similarity (Figures 5a and S4a and S4b). We found that there are common substructures shared in 75% of the compounds described in the set (Figures 4a, 5a, and 5b). Unique compounds with unique structures were put into group 1 while others with at least 0.7 similarity scores were put into group 2 (Figure 5c). The investigation of the group 2 compounds having at least a -10.8 kcal/mol predicted affinity towards CCHFV

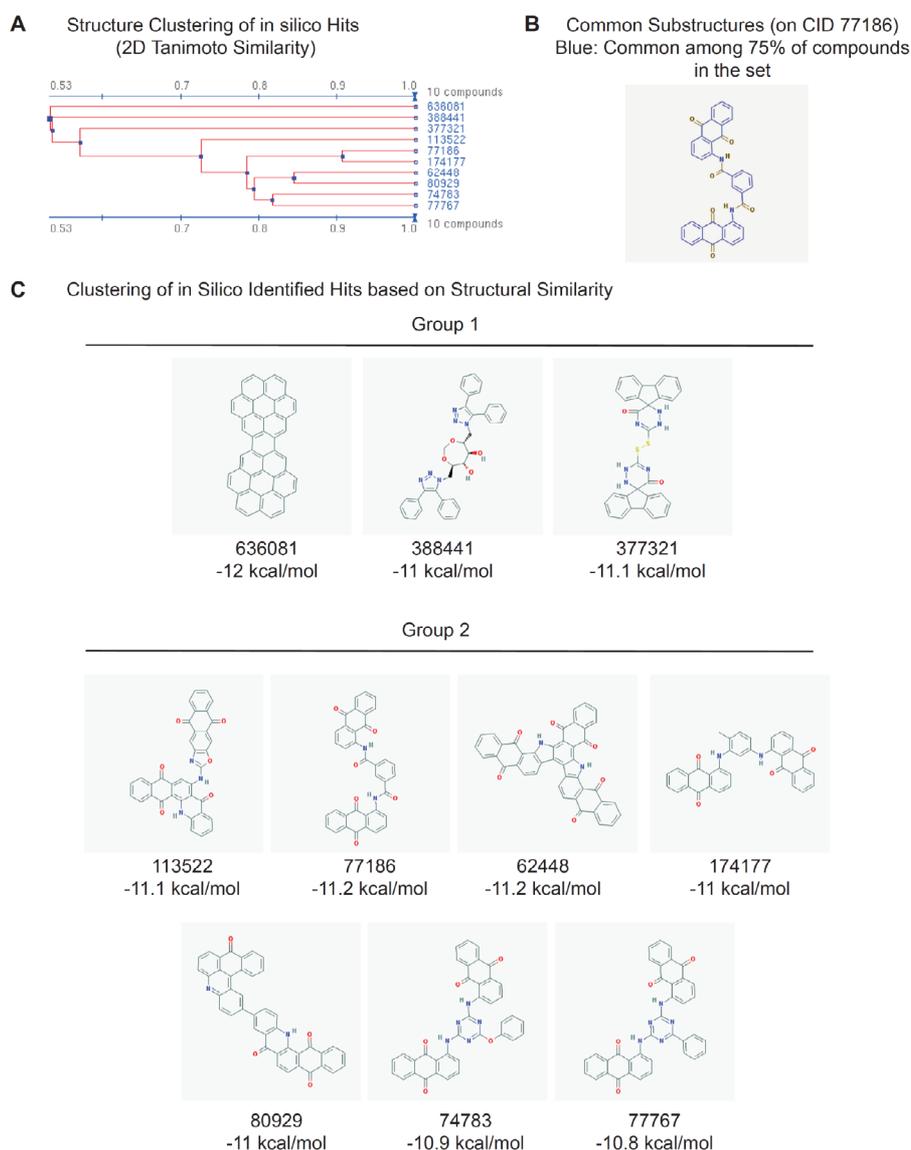


Figure 5. Clustering of hits identified in silico pinpoints common substructures. A) Structure clustering of in silico hits based on 2D Tanimoto similarity. Note that the majority have a similarity score greater than 0.7. (See Figure S4 for 3D structure similarities). B) Substructure analysis identifies common substructures among 75% of the set as shown in blue on CID 77186. C) Group 1 is composed of nonclustered compounds CID 113522, CID 77186, and CID 62448 with unique molecular structures. Group 2 is composed of compounds with high structure similarity scores up to 0.9. Note that all of the compounds listed here have predicted affinity scores of at least -10.8 .

OTU protease provided a novel group of compounds for generation of CCHFV antivirals.

3.6. Cytotoxicity analysis of putative inhibitors of CCHFV OTU protease

It is possible that the identified hits could have cytotoxicity. Thus, we analyzed the hits identified *in silico* in the PubChem bioassays database to determine if they have any reported cytotoxicity (Han et al., 2008). None of the hits in Figure 4a were reported to have any cytotoxicity. In addition, compounds CID_377321 and CID_388441 were tested on various cell lines and showed no cytotoxicity as reported by PubChem bioassays. On the other hand, 70% of putative OTU protease inhibitors identified *in vitro* showed in Figure 3a are predicted to be potentially cytotoxic as reported in PubChem (see Supplementary Table 1 for more info). We have determined some potentially cytotoxic and noncytotoxic CCHFV OTU protease inhibitors that could be used in further evaluation of treatments of CCHFV infections.

4. Discussion

It is a great challenge to identify compounds that can inhibit the virulence of deadly viruses. Large library screenings are costly and require additional validation and toxicity analysis. On the other hand, virtual screening could provide the means to filter out compounds with low affinity, thus possibly providing compounds with high IC₅₀ values that could overcome the toxicities related to high doses. Computational identification of such compounds requires accurate knowledge of the conserved residues in the inhibition pocket of the target protein. In this study, we used a computational method based on the correlation analysis of putative inhibitors with the predicted pockets of highly conserved residues to identify the inhibition pocket of the CCHFV OTU protease. Robust homology studies among CCHFV OTU protease-related proteins revealed a new pocket of conserved residues in close proximity to the UB and ISG15 binding sites. We have shown that the putative OTU protease inhibitors identified *in vitro* preferentially bind to the newly revealed pocket of Y89–W99 rather than the previously suggested D37, C40, and H151 residues.

Mutation studies in viral OTU protease are in alignment with our findings. Mutation studies with a closely related OTU protease at H151 resulted in the destabilization of OTU in a human cell line (HEK293 cells) but not in a monkey cell line (Vero cells) (Bakshi et al., 2013). This points to a major difference between human and other vertebrates in the invasion of CCHFV and related viruses and the requirement of H151 in the stability of OTU protein in human tissues, rather than involving OTU protease activity. Moreover, it has been recently shown that a mutated OTU protein at C40 still has a significant

effect on the induction of the JAK/STAT and TNF- α /NF- κ B pathways, although it loses cleavage ability (Bergeron et al., 2010; Bakshi et al., 2013). This suggests that C40, which is partially located in the pocket of Y89–W99, is required for CCHFV OTU protease activity.

We have identified the Y89–W99 pocket of CCHFV OTU protease as a pharmaceutical target for optimization and identification of CCHFV OTU protease inhibitors. The discovery of the Y89–W99 pocket allowed us to design *in silico* screening to identify compounds with higher affinities toward the OTU protease inhibition pocket. We automated the *in silico* screening system using PaDEL-ADV and successfully identified compounds that show preferential binding to the Y89–W99 pocket with higher binding affinities (lower binding energies). Our screening provided a cost-effective and timesaving screening process for identification of the inhibition pocket of CCHFV OTU protease and highly potent inhibitors. This led us to identify over 300 candidate compounds. Among these compounds, we found that NSC658721 (CID_377321) and NSC683337 (CID_388441) have at least a -11.0 kcal/mol calculated binding affinity towards the pocket of Y89–W89 and the whole OTU protease. In addition, we found a common backbone of the chemical structure that could shed light on future inhibitor development studies targeting the pocket of viral OTU protease. Further investigation of these compounds is required to determine whether they could be used for the inactivation of invasion mechanisms through the inhibition of viral OTU protease activity in the fight against CCHFV infections.

In addition, robust characterization of CCHFV OTU-related proteins led us to discover that CCHFV OTU is related to intracellular parasites that use OTU-like proteases. They are likely to use a similar system to invade cells and overcome innate cellular immunity through the broad deconjugation activities of OTU protease. This study also highlights that lead compounds identified *in vitro* could be used successfully to determine the inhibition pocket of proteins through a correlation analysis following a computational calculation of docking affinities of compounds with the target protein. This study could be extended experimentally to evaluate the biological activity of the identified compounds, which would help in designing compounds with higher potency and lower toxicity.

Acknowledgments

We thank Dr Savaş Kaya from the Immunology Department of the School of Medicine at Dicle University in Diyarbakır, Turkey, for his valuable input regarding MGLTools and Autodock Vina. We thank Dr Emrah Nikerel from the Department of Genetics and Bioengineering at Yeditepe University for his valuable

input regarding statistical analysis. We are thankful for the support of the cofunded Brain Circulation Scheme of the Scientific and Technological Research Council of Turkey (TÜBİTAK) and the Marie Curie Action COFUND of the 7th Framework Programme (FP7) of the European

Commission grant number 115C039, the Science Academy Young Scientist Award Program (BAGEP-2015, Turkey), and funds provided by both North American University in Houston, TX, USA, and Yeditepe University in İstanbul, Turkey.

Supplementary material

Supplementary Figure 1: https://drive.google.com/open?id=0B6Il4DEh_MdMYktQcDA0RzVoUUU

Supplementary Figure 2: https://drive.google.com/open?id=0B6Il4DEh_MdMSmdVdINkRlhObW8

Supplementary Figure 3: https://drive.google.com/open?id=0B6Il4DEh_MdMQWhRZEs2R3NaTVU

Supplementary Figure 4: https://drive.google.com/open?id=0B6Il4DEh_MdMc2poRUG0c1BsYIE

Supplementary Table 1: https://drive.google.com/open?id=0B6Il4DEh_MdMMkVFMkhzQU5qSTg

Supplementary Table 2: https://drive.google.com/open?id=0B6Il4DEh_MdMMTNrVWtReHJBbFU

Supplementary Table 3: https://drive.google.com/open?id=0B6Il4DEh_MdMbHRTQkxJYTFrRWs

Supplementary Table 4: https://drive.google.com/open?id=0B6Il4DEh_MdMVkRLNDZjTjZaZlk

References

- Akutsu M, Ye Y, Virdee S, Chin JW, Komander D (2011). Molecular basis for ubiquitin and ISG15 cross-reactivity in viral ovarian tumor domains. *P Natl Acad Sci* 108: 2228–2233.
- Arguello MD, Hiscott J (2007). Ub surprised: viral ovarian tumor domain proteases remove ubiquitin and ISG15 conjugates. *Cell Host Microbe* 2: 367–369.
- Bakshi S, Holzer B, Bridgen A, McMullan G, Quinn DG, Baron MD (2013). Dugbe virus ovarian tumour domain interferes with ubiquitin/ISG15-regulated innate immune cell signalling. *J Gen Virol* 94: 298–307.
- Bente DA, Forrester NL, Watts DM, McAuley AJ, Whitehouse CA, Bray M (2013). Crimean–Congo hemorrhagic fever: history epidemiology pathogenesis clinical syndrome and genetic diversity. *Antiviral Res* 100: 159–189.
- Bergeron E, Albariño CG, Khristova ML, Nichol ST (2010). Crimean–Congo hemorrhagic fever virus-encoded ovarian tumor protease activity is dispensable for virus RNA polymerase function. *J Virol* 84: 216–226.
- Capodagli GC, Deaton MK, Baker EA, Lumpkin RJ, Pegan SD (2013). Diversity of ubiquitin and ISG15 specificity among nairoviruses' viral ovarian tumor domain proteases. *J Virol* 87: 3815–3827.
- Capodagli GC, McKercher MA, Baker EA, Masters EM, Brunzelle JS, Pegan SD (2011). Structural analysis of a viral ovarian tumor domain protease from the Crimean–Congo hemorrhagic fever virus in complex with covalently bonded ubiquitin. *J Virol* 85: 3621–3630.
- Chang MW, Ayeni C, Breuer S, Torbett BE (2010). Virtual screening for HIV protease inhibitors: a comparison of AutoDock 4 and Vina. *PLoS ONE* 5: e11955.
- Dokuzoguz B, Celikbas AK, Gök SE, Baykam N, Eroglu MN, Ergonul O (2013). Severity scoring index for Crimean–Congo hemorrhagic fever and the impact of ribavirin and corticosteroids on fatality. *Clin Infect Dis* 57: 1270–1274.
- Ergonul O (2012). Crimean–Congo hemorrhagic fever virus: new outbreaks new discoveries. *Curr Opin Virol* 2: 215–220.
- Ergunay K, Whitehouse CA, Ozkul A (2011). Current status of human arboviral diseases in Turkey. *Vector Borne Zoonotic Dis* 11: 731–741.
- Frias-Staheli N, Giannakopoulos NV, Kikkert M, Taylor SL, Bridgen A, Paragas J, Richt JA, Rowland RR, Schmaljohn CS, Lenschow DJ et al. (2007). Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses. *Cell Host Microbe* 2: 406–416.
- Gale P, Stephenson B, Brouwer A, Martinez M, la Torre de A, Bosch J, Foley-Fisher M, Bonilauri P, Lindström A, Ulrich RG et al. (2012). Impact of climate change on risk of incursion of Crimean–Congo haemorrhagic fever virus in livestock in Europe through migratory birds. *J Appl Microbiol* 112: 246–257.
- Gargili A, Midilli K, Ergonul O, Ergin S, Alp HG, Vatansever Z, Iyisan S, Cerit C, Yilmaz G, Altas K et al. (2011). Crimean–Congo hemorrhagic fever in European part of Turkey: genetic analysis of the virus strains from ticks and a seroepidemiological study in humans. *Vector Borne Zoonotic Dis* 11: 747–752.
- Han L, Wang Y, Bryant SH (2008). Developing and validating predictive decision tree models from mining chemical structural fingerprints and high-throughput screening data in PubChem. *BMC Bioinformatics* 9: 401.

- James TW, Frias-Staheli N, Bacik JP, Levingston Macleod JM, Khajehpour M, García-Sastre A, Mark BL (2011). Structural basis for the removal of ubiquitin and interferon-stimulated gene 15 by a viral ovarian tumor domain-containing protease. *P Natl Acad Sci* 108: 2222–2227.
- Lindeborg M, Barboutis C, Ehrenborg C, Fransson T, Jaenson TG, Lindgren PE, Lundkvist A, Nyström F, Salaneck E, Waldenström J et al. (2012). Migratory birds, ticks, and Crimean–Congo hemorrhagic fever virus. *Emerg Infect Dis* 18: 2095–2097.
- Makarova KS, Aravind L, Koonin EV (2000). A novel superfamily of predicted cysteine proteases from eukaryotes viruses and *Chlamydia pneumoniae*. *Trends Biochem Sci* 25: 50–52.
- Malakhova OA, Zhang DE (2008). ISG15 inhibits Nedd4 ubiquitin E3 activity and enhances the innate antiviral response. *J Biol Chem* 283: 8783–8787.
- Nanao MH, Tcherniuk SO, Chroboczek J, Dideberg O, Dessen A, Balakirev MY (2004). Crystal structure of human otubain 2. *EMBO Rep* 5: 783–788.
- Ozkaya E, Dincer E, Carhan A, Uyar Y, Ertek M, Whitehouse CA, Ozkul A (2010). Molecular epidemiology of Crimean–Congo hemorrhagic fever virus in Turkey: occurrence of local topotype. *Virus Res* 149: 64–70.
- Palomar AM, Portillo A, Santibáñez P, Mazuelas D, Arizaga J, Crespo A, Gutiérrez Ó, Cuadrado JF, Oteo JA (2013). Crimean–Congo hemorrhagic fever virus in ticks from migratory birds, Morocco. *Emerg Infect Dis* 19: 260–263.
- Sanner MF (1999). Python: a programming language for software integration and development. *J Mol Graph Model* 17: 57–61.
- Sanner MF, Olson AJ, Spehner JC (1996). Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers* 38: 305–320.
- Snijder EJ, Wassenaar AL, Spaan WJ (1994). Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. *J Virol* 68: 5755–5764.
- Snijder EJ, Wassenaar AL, Spaan WJ, Gorbalenya AE (1995). The arterivirus Nsp2 protease, an unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases. *J Biol Chem* 270: 16671–16676.
- Sudha KN, Shakira M, Prasanthi P, Sarika N, Kumar CN, Babu PA (2008). Virtual screening for novel COX-2 inhibitors using the ZINC database. *Bioinformation* 2: 325–329.
- Trott O, Olson AJ (2009). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function efficient optimization and multithreading. *J Comput Chem* 31: 455–461.
- van Kasteren PB, Beugeling C, Ninaber DK (2012). Arterivirus and nairovirus ovarian tumor domain-containing deubiquitinases target activated RIG-I to control innate immune signaling. *J Virol* 86: 773–785.
- Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Zhou Z, Han L, Karapetyan K, Dracheva S, Shoemaker BA et al. (2011). PubChem's BioAssay database. *Nucleic Acids Res* 40: D400–D412.
- Wassman CD, Baronio R, Demir OZ, Wallentine BD, Chen CK, Hall LV, Salehi F, Lin DW, Chung BP, Hatfield GW et al. (2013). Computational identification of a transiently open L1/S3 pocket for reactivation of mutant p53. *Nat Commun* 4: 1407.
- Weber F, Mirazimi A (2008). Interferon and cytokine responses to Crimean Congo hemorrhagic fever virus; an emerging and neglected viral zoonosis. *Cytokine Growth Factor Rev* 19: 395–404.
- Whitehouse CA (2004). Crimean–Congo hemorrhagic fever. *Antiviral Res* 64: 145–160.
- Yang JM, Chen YF, Tu YY, Yen KR, Yang YL (2007). Combinatorial computational approaches to identify tetracycline derivatives as flavivirus inhibitors. *PLoS ONE* 2: e428.