

Three dimensional model of severe acute respiratory syndrome coronavirus helicase ATPase catalytic domain and molecular design of severe acute respiratory syndrome coronavirus helicase inhibitors

Marcin Hoffmann · Krystian Eitner · Marcin von Grotthuss ·
Leszek Rychlewski · Ewa Banachowicz · Tomasz Grabarkiewicz ·
Tomasz Szkoda · Andrzej Kolinski

Received: 22 December 2005 / Accepted: 17 July 2006 / Published online: 14 September 2006
© Springer Science+Business Media B.V. 2006

Abstract The modeling of the severe acute respiratory syndrome coronavirus helicase ATPase catalytic domain was performed using the protein structure prediction Meta Server and the 3D Jury method for model selection, which resulted in the identification of 1JPR, 1UAA and 1W36 PDB structures as suitable templates for creating a full atom 3D model. This model was further utilized to design small molecules that are expected to block an ATPase catalytic pocket thus inhibit the enzymatic activity. Binding sites for various functional groups were identified in a series of molecular dynamics calculation. Their positions in the catalytic pocket were used as constraints in the Cam-

bridge structural database search for molecules having the pharmacophores that interacted most strongly with the enzyme in a desired position. The subsequent MD simulations followed by calculations of binding energies of the designed molecules were compared to ATP identifying the most successful candidates, for likely inhibitors—molecules possessing two phosphonic acid moieties at distal ends of the molecule.

Keywords SARS · Inhibitors · Drug design · Computational chemistry · Bioinformatics

Introduction

Over the last 20 years many new viruses were identified, and many mutations resulting in increased virulence of the known viruses were discovered. HIV, HCV, influenza, and SARS are examples of diversity of the biohazard on the Earth. Such threads constitute a challenge to develop fast and productive methods to predict new target proteins, their functions, and active sites, as well as create the need for improved drug design procedures. The information gathered by molecular biologists is often stored in publicly available protein databases, like Swiss-Prot—with non-redundant protein sequences with accurate functional annotations [1]—or PDB—with macromolecular structural data [2]. Therefore such databases can serve as mines of information and speed up the process of drug design. In this study we utilized the information gathered in these databases to address the recent threat caused by severe acute respiratory syndrome (SARS). SARS is a life-threatening form of pneumonia

M. Hoffmann (✉) · K. Eitner · M. von Grotthuss ·
L. Rychlewski · T. Grabarkiewicz
BioInfoBank Institute, ul. Limanowskiego 24A, 60-744
Poznan, Poland
e-mail: mmh@bioinfo.pl

M. Hoffmann
Quantum Chemistry Group, Faculty of Chemistry, Adam
Mickiewicz University, ul. Grunwaldzka 6, 60-780 Poznan,
Poland

E. Banachowicz
Bioinformatics Unit, Department of Physics, Adam
Mickiewicz University, ul. Umultowska 85, 61-614 Poznan,
Poland

T. Szkoda
Respiratory Virus Laboratory, Department of Virology,
National Institute of Hygiene, ul. Chocimska 24, 00-791
Warsaw, Poland

A. Kolinski
Faculty of Chemistry, Warsaw University, ul. Pasteura 1,
02-093 Warsaw, Poland

characterized by high fever, nonproductive cough, chills, myalgia, lymphopenia, and progressing infiltrates on chest radiography [3]. Between 2002 and 2003, an epidemic emerged that, facilitated by international air travel, spread within a few weeks from its origin in Guangdong Province, China, to many other countries. WHO reported over 8,000 SARS cases and nearly 800 deaths resulting from the infection with the SARS-associated coronavirus (SARS CoV) [4]. Studies on SARS CoV resulted in identification of protein targets for potential drugs, which included SARS CoV protease, polymerase and helicase [5]. In this study we focused on SARS CoV Hel (helicase)—the enzyme which couples energy from nucleotide triphosphate (NTP) hydrolysis with the unwinding of duplex viral nucleic acid obtained after replication.

We used SARS CoV Hel polypeptide sequence from SARS CoV replicase polyprotein 1ab (UniProt accession number P59641) as an input for protein structure prediction Meta Server [6]. The Meta Server collects output from diverse structure prediction methods and generates a consensus model using the 3D-Jury approach. This approach has proven efficient and effective in successful fold prediction for many proteins [7–9]. In brief, the 3D-Jury utilizes groups of models generated by a set of servers predicting protein structure. All models are compared with each other and a similarity score is assigned to each pair, which equals to the number of C_α atom pairs that are within 3.5 Å after optimal superposition. 3D-Jury selects the model that has the highest average similarity to other models in the collected set. The average similarity expressed as the average number of superimposed C_α atoms is also a reliable measure of the accuracy of the model. If this value is above 50 the selected model has approximately a 90% chance to belong to the same fold class as the native structure of the target protein [9]. The 3D-Jury system is used to select a correct initial template for full atom modeling conducted usually with Modeller [10] as was also done in this case. The main goal for Modeller is the correct loop closure and side chain placement. Advanced minimization options feasible in Modeller are usually disabled to prevent a distortion of the correctly aligned core and functional residues.

The application of computer-based models using analytical potential energy functions within the framework of classical mechanics has proven effective and powerful in studying large-scale molecules like proteins or nucleic acids [11]. The binding of small molecule ligands to giant protein targets is central to numerous biological functions. Thus, docking ligands to the binding site of a receptor is often performed

using points of complementarities between host and guest molecules. The accurate prediction of the binding modes between the ligand and protein is of utmost importance in modern structure-based drug design. Moreover, drug design requires advanced force field based on non-covalent interactions between molecules, covalent bonding, charges, and atomic volume. Only then such a force field is powerful enough for analysis and prediction of molecular interactions. Despite various limitations and crude approximations molecular mechanics (MM) and molecular dynamics (MD) methods have been used to study a wide variety of phenomena, including structure and dynamics of simple and complex structures, thermodynamics of ligands binding to proteins, conformational transitions in nucleic acids, and many others [11–21]. Recent developments in hybrid quantum mechanics (QM) and MM method created new opportunities in more accurate assessment of interaction energies [22–25]. In particular ONIOM method nicely resolved problems with QM and MM boundary and was successfully utilized in studies on mechanism of enzymatic reactions or calculations of interaction energies between proteins and their ligands [26–31].

Methods

Fully atomic 3D structure of SARS CoV Hel ATPase domain

The 3D-Jury approach was used for the initial fold assignment. 3D-Jury, takes as input groups of models generated by a set of fold assignment servers (ORFeus [32], SamT02 [33], FFAS03 [34], mGenTHREADER [35], INBGU [36], FUGUE-2 [37], 3D-PSSM [38]), neglecting the confidence scores assigned by the servers to the models. All models are compared with each other and a similarity score is assigned to each pair, which equals to the number of C-alpha atom pairs that are within 3.5 Å after optimal superposition. The final 3D-Jury score of a model equals to the average similarity scores of considered model pairs. It can be expected that highly reliable models produced by fold recognition methods have less ambiguities in the alignments to their template structures, which would result in higher similarity between models and consequently in higher 3D-Jury scores. The final score of the model can be also calculated on a per residue basis enabling the detection of well and less well modeled regions. The application of the 3D-Jury approach on the sequence of the SARS CoV Hel ATPase catalytic domain resulted in the selection of models based on

three template structure: 1PJR [39], 1UAA [40], and 1W36 [41]. In order to obtain the full atom three dimensional model the side chains and missing loops were rebuilt with Modeller program [10, 42, 43]. The final model was energy minimized using Amber force field [44–47].

Protein ligand interactions

In order to study interactions between SARS CoV Hel ATPase domain and small molecules we carried out a series of MD simulations followed by energy minimizations with Amber force field [44–47] as implemented in Tinker software [48–51]. Figure 1 presents a schematic representation of computational procedure steps. Required parameters for ATP were taken from the work of Meagher et al. [52]. To derive the necessary force-field parameters for the functional groups, and the designed molecules that were unavailable in the standard force field database, we followed the procedure suggested by Fox and Kollman [44] to be consistent with the other Amber force field parameters.

The ATP binding site in the SARS CoV Hel can be identified, based on the homology with proteins used as templates, because one of these helicases, namely PcrA one, has been observed in the crystal structure in complex with ATP (3PJR). Therefore, we could verify if simulations with TINKER package and Amber force field parameters are capable to identify ATP binding site correctly. In this test we derived ATP conformations observed in complexes with macromolecules with the help of PDB-Ligand database, where 435 models of ATP molecule interacting with proteins were collected, based on 205 entries in PDB. These ATP structures clustered at 1 Å RMS deviation led to identification of 37 unique ATP conformations. Then, we randomly rotated ATP molecule in a given conformation and placed it in a random position on the surface of the SARS CoV helicase. One hundred repetitions for each conformation generated starting points (3,700) for the initial energy pre-minimization, where protein coordinates were kept frozen, while flexible ATP molecules were docked to the enzyme. After initial MM minimization with Amber force field, the resulting ATP molecules were clustered at 1 Å RMS deviation identifying unique positions of ATP in complex with SARS CoV Hel. From this set one hundred structures of the lowest energy were selected for further simulations. These structures served as starting points for MD simulations with temperature varying from 1,000 to

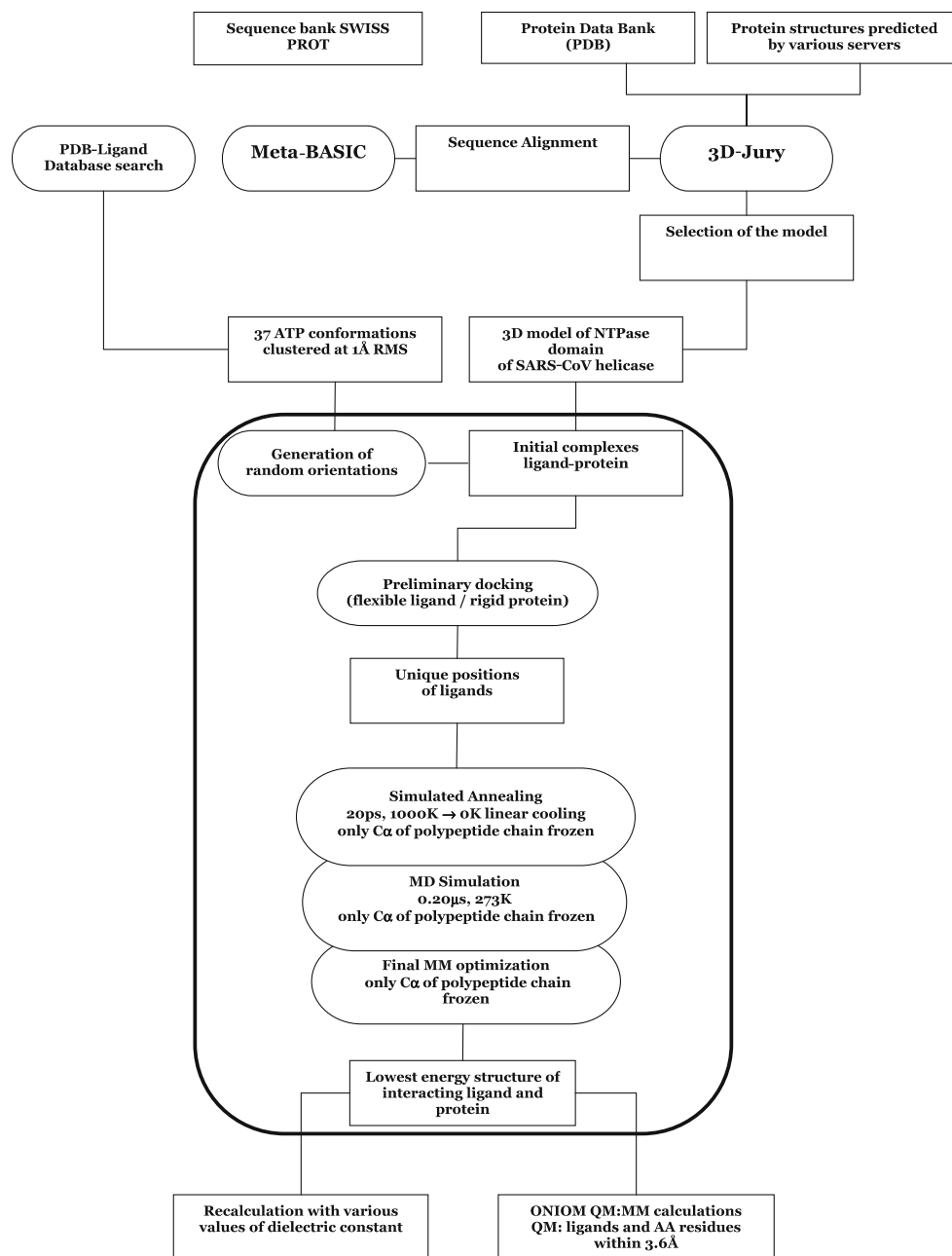
0 K (simulated annealing) over 0.1 ps, followed by MD simulations over 0.2 μs at 273 K. In all these MD simulations only the positions of C α carbon atoms of polypeptide chain were frozen, while the positions of other atoms of the protein, and all atoms of ATP, or putative inhibitors were optimized. Having had the MD results, the subsequent energy minimization was performed. The comparison of ATP binding energies led to identification of the most favorable mode of interaction between the protein and ATP.

Three runs of such a procedure were performed and in every case the same lowest energy structure of the complex of ATP with the protein was identified. As expected the conformation of ATP when interacting with protein was different from the conformation being the global minimum of ATP in vacuo [53]. The position of ATP molecule with the most favorable protein ligand binding energy was very similar to the one occupied by ATP in the crystal structure of one of the template proteins—PcrA helicase in complex with ATP (PDB code 3PJR) [54] (see supporting information Figure S1).

In the next step we identified attractors for various simple molecules on the surface of the enzyme using MD simulations. The employed computational procedure is analogous to multiple copy simultaneous search (MCSS) methods [55–57]. The simple molecules utilized in our search were: PO $_4^{3-}$ (see Fig. 3), CH $_3$ COO $^-$, CH $_3$ CONHCH $_3$ (peptide moiety), CH $_3$ OH, CH $_3$ C(=O)CH $_3$, C $_6$ H $_6$, NH $_4^+$, CH $_4$, C(NH $_2$) $_3^+$ (protonated guanidine), and imidazole. For each functional group we created a sphere of 10 Å radius filled with one thousand copies of randomly oriented molecules, placed in random positions in the sphere. The sphere was centered over the NTPase catalytic pocket of SARS CoV Hel (the average of Cartesian coordinates of ATP atoms in the lowest energy complex with the enzyme identified previously). Then simulated annealing procedure and MD simulations, as the one used for ATP, were carried out, followed by energy minimization optimizing the positions of the functional groups (see Fig. 2).

The individual copies of the molecules were not interacting with each other, so many copies collapsed into the same positions. We selected the attractors for the functional groups that were within 10 Å distance from the mass center of ATP bound to the enzyme. The attractors' positions were used to impose geometrical constraints in the Cambridge structural database (CSD) [58–60] search for molecules having desired functional groups in a given distance from each other. Because the two phosphate groups possessed the

Fig. 1 Summary of computational procedure steps, carried out in the search for ATP binding side via MD/MM calculations. In the case of the other ligands similar approach was adopted. However, they were not observed in PDB database [2] so the structures from CSD [58] were taken instead



highest binding energies, at least twice as favorable as any other small molecule, we focused our search on molecules possessing P atom in their structure [61]. The obtained hits were investigated and modified manually to fit into SARS CoV Hel ATP binding site so that 17 molecules were selected for further studies. These molecules were placed in the active side of the protein, so that two desired functional groups were in the positions identified by the attractors for the corresponding small molecules. In the next step MD simulations with temperature varying from 1,000 to 0 K (simulated annealing) over 0.1 ps were carried out,

followed by MD simulations over 0.2 μ s at 273 K, and energy minimization. In these simulations only the coordinates of C_{α} carbon atoms of the protein were frozen.

In order to assess binding energies in aqueous solution, the complexes of the protein and the designed molecules that maximized interaction energies were ‘immersed’ into a box of 8,000 water molecules and annealed in a fashion described above. Figure 3 presents thermodynamic cycle employed to calculate protein–ligand binding energy in aqueous solution according to the equation:

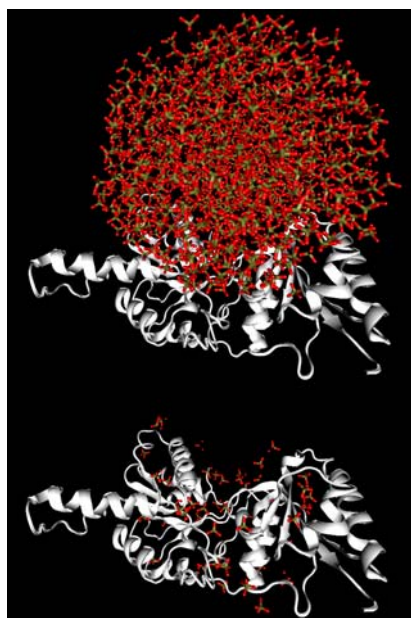


Fig. 2 Example of localizing attractors for small molecules. Initial position of one thousand of randomly oriented PO_4^{3-} ions centered over ATPase catalytic pocket of the enzyme (*top*). After calculations, many copies of PO_4^{3-} ions collapsed into the same positions on the surface of the enzyme (*bottom*), localizing attractors for PO_4^{3-} ions. (Figure prepared with VMD 1.8.3 [83])

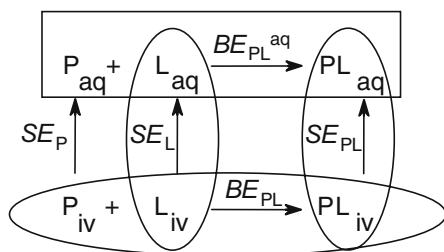


Fig. 3 Thermodynamic cycle used to calculate protein–ligand binding energy in aqueous solution (*aq*). The values of protein–ligand binding energy in vacuo (*iv*) were calculated after MD/MM calculations as well as the values of aqueous solvation energies for the ligand and protein ligand complexes. The aqueous solvation energy of the protein is constant for all examined ligands. Energies of the transitions that were calculated in MD/MM simulations are indicated by *ellipses*, while the protein ligand binding energy in aqueous solution calculated from the presented thermodynamic cycle indicated by the *rectangle*

$$\text{BE}_{\text{PL}}^{\text{aq}} = \text{BE}_{\text{PL}} + \text{SE}_{\text{PL}} - \text{SE}_{\text{P}} - \text{SE}_{\text{L}} \quad (1)$$

where $\text{BE}_{\text{PL}}^{\text{aq}}$ is protein ligand binding energy in aqueous solution, BE_{PL} is protein ligand binding energy for the isolated system, while SE_{P} , SE_{L} , and SE_{PL} stand for aqueous solvation energy of protein, ligand, and protein–ligand complex, respectively. The aqueous solvation energies of ligands and protein–ligand complexes

were calculated as usual in calculations with explicit water molecules [62]. The aqueous solvation energy of the protein was not calculated explicitly as this term has the same value for all studied ligands. Because we focused on the design of molecules that bind to SARS CoV Hel stronger than ATP we compared the calculated protein–ligand binding energies in aqueous solution with the same energy calculated with ATP. Such a comparison leads to binding energies relative to ATP (see supporting information Table S2). If the value of this relative energy is negative it means that a given ligand binds to SARS CoV Hel stronger than ATP.

We also tested the influence of various dielectric constant values on the protein–ligand interaction energies (see supporting information Table S1). The pronounced dependence of interaction energies calculated with MM methods on dielectric constant suggests that electrostatic interactions are strongly overestimated. Thus, we decided to employ quantum mechanical methods to calculate interaction energies between ligands and the amino acids residues surrounding them. Because of the size of the systems we performed ONIOM two layer QM: MM calculations (see Fig. 4) using Gaussian 03 program [63]. Amino acid residues within 3.6 Å from any atom of an inhibitor together with this inhibitor were treated quantum mechanically with AM1 [64–66] and PM3 [67, 68] semiempirical methods, while the remaining part of the protein was kept frozen and described with the use of Amber force field.

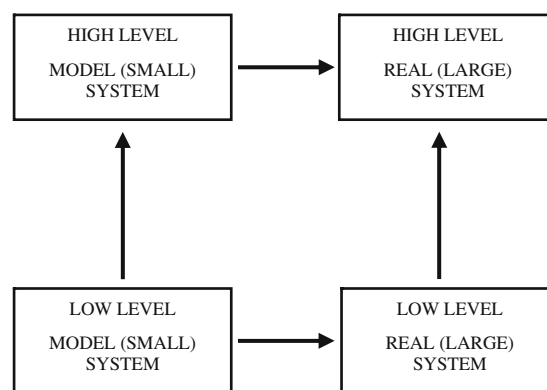


Fig. 4 In ONIOM extrapolation scheme energy of large system calculated at high level of theory is approximated as energy of large system calculated at low level + energy of small system calculated at high level—energy of small system calculated at low level. Thus, computationally very expensive calculations for large system at high level of theory are avoided. Instead, cheaper calculations are carried out for large system at low level, small system at high level, and small system at low level

Results and discussion

The SARS CoV belongs to a diverse group of enveloped, positive-strand RNA viruses. The genomic organization of SARS CoV has a gene order characteristic for coronaviruses: 5'-replicase [rep], spike, envelope, membrane, and nucleocapsid-3'. The SARS CoV rep gene, which comprises approximately two-thirds of the genome, is predicted to encode two polyproteins (encoded by orf1a and orf1b) that undergo co-translational proteolytic processing [69]. In this study we applied 3D-Jury [70] to predict the structure of the catalytic domain of the SARS CoV helicase (SARS CoV Hel) [54] encoded by orf1b.

Homology modeling

The initial model of SARS CoV Hel catalytic domain was obtained as described in the [Methods](#) section. Based on the obtained alignment (see Fig. 5) we chose structures of PcrA, Rep, and RecB DNA helicases as templates (PDB accession codes: 1PJR [39], 1UAA [40], and 1W36 [41], respectively) that obtained the

highest (> 150) 3D-Jury score. It should be noted that such high scores suggest very high probability for the correct prediction of the protein structure.

The PcrA/Rep/RecB helicases contain four structural domains: two parallel α - β domains which encompass the canonical helicase sequence motifs, and two additional domains encoded as a single insertion within the polypeptide sequence of the main domains [71]. NTP-binding site is situated in a cleft between α - β domains [72] which are both conserved in the SARS CoV Hel. The 3D alignment [17] of the 1PJR, 1UAA, and 1W36 NTPase domains and the corresponding superposition of the protein backbones are presented in Figs. 5 and 6, respectively. The sequence of SARS CoV Hel catalytic domain was added to the alignment based on the result obtained from 3D-Jury system. A percentage of sequence identity and similarity between each pair of the proteins is shown in Table 1.

Conserved motifs

SARS CoV Hel has six motifs (see Fig. 6 and Table 2) that are characteristic for helicases from the superfamily-1 (SF1). [73]. The motif I (Walker A) and motif

Fig. 5 Sequence-structure multiple alignment of four NTPase domains of SF1 helicases. Structures of NTP-binding domains of 1PJR [39], 1UAA [40], and 1W36 [41] proteins were aligned using three-dimensional similarity [17]. The sequence of SARS CoV Hel was added based on alignment obtained from 3D-Jury. RMSD of superposition of C_α atoms and percentages of sequence identity and sequence similarity are presented in Table 1.

Residues from six conserved motifs are marked by *cross* or *hash* symbols. *Hash character* highlights amino acids placed in the active pocket, which are shown in Fig. 6. Identical or similar (BLOSUM62) residues among four proteins are presented by *white symbols* with background colored by amino acid types

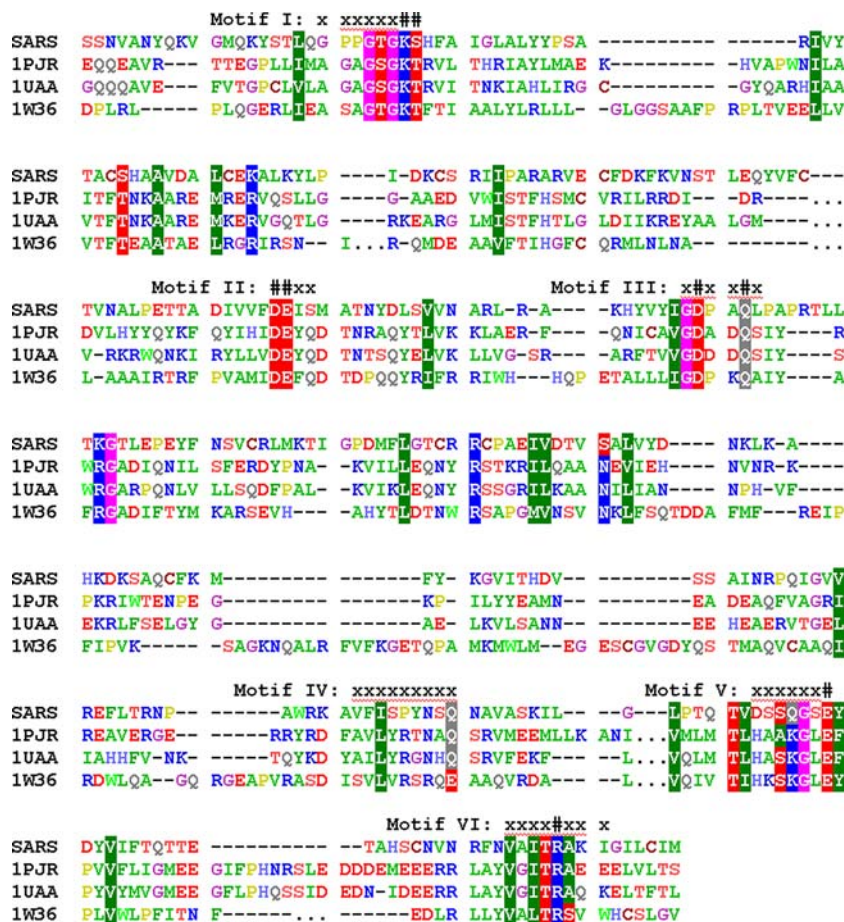


Fig. 6 Superposition of three chain backbones of the NTPase domain, which come from PcrA, Rep and RecB helicases (PDB accession symbols: 1PJR [39], 1UAA [40], and 1W36 [41]). Secondary structure elements of each of the proteins are colored by succession from *blue* (N-end) to *red* (C-end). Side-chains of eight conserved amino acids from the NTP-binding pocket are presented by sticks and colored by atom types. Values of root mean square deviation (RMSD) of aligned C_{α} atoms and RMSD of all atoms of the shown side-chains are presented in Table 1

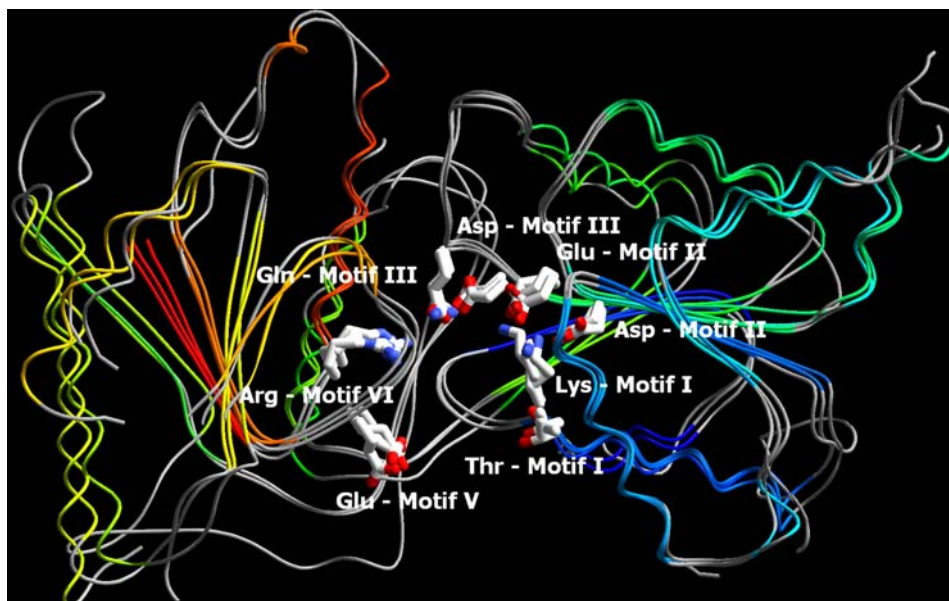


Table 1 Values of similarity between four NTPase domains based on multiple alignment presented in Fig. 5: percentage of sequence identity (first value), percentage of sequence similarity

	SARS	1PJR (PcrA)	1UAA (Rep)	1W36 (RecB)
SARS	–	13% 28%	12% 26%	11% 21%
1PJR (1PcrA)	13% 28%	–	48% 65% 1.13 1.34	20% 38% 1.46 1.35
1UAA (Rep)	12% 26%	48% 65% 1.13 1.34	–	17% 34% 1.48 1.79
1W36 (RecB)	11% 21%	20% 38% 1.46 1.35	17% 34% 1.48 1.79	–

using the BLOSUM62 matrix (second value), root mean square deviation (RMSD) of aligned C_{α} atoms (third value), RMSD of all side-chain atoms highlighted by symbol (fourth value)

II (Walker B) have residues that form the pocket and interact with MgNTP/MgNDP. The Walker A motif, initially defined as a GxxxxGKT [74] sequence and later as a (G/A)x(A/P)GxGK(S/T) consensus [75], requires the three final residues GK(T/S) in order to be functional [71]. However, in the SARS CoV Hel the motif I is completely conserved (GPPGTGKS). The side chain of the key lysine (Lys288) occupies the position that will be occupied by the bound Mg^{2+} ion when the NTP- Mg^{2+} complex binds the SARS CoV helicase. Upon binding NTP- Mg^{2+} , the lysine side chain contacts the β phosphate of the bound NTP and can stabilize the transition state during catalysis [72]. Replacement of this residue in mutant helicases resulted in large decrease of the rate of NTP hydrolysis. Mutated helicases were unable to catalyze duplex nucleic acid unwinding [76].

Motif II (known as a Walker B) plays a key role in the NTP hydrolysis reaction. The Walker B motif originally defined as a single aspartic acid residue [74], took later the general form of DE, across the helicase superfamily 1 and 2 [73]. Both of these key amino acids

are conserved in the SARS CoV Hel. The carboxyl group of the aspartic acid (Asp374) coordinates the Mg^{2+} ion of MgNTP/MgADP through outer sphere interactions, while the glutamic acid (Glu375) might be a catalytic base in NTP hydrolysis. Clearly, these amino acids are in position to co-ordinate the NTP-associated Mg^{2+} ion and activate the attacking water molecule, respectively, as proposed previously for related NTPases [77].

Motif V (DSSQGSE) and the first part of motif III (GDPAQ) participate in a complex network of interaction including ligation of MgNTP/MgNDP and formation of a specific salt bridge between α - β domains. The glutamic acid (Glu540) from motif V ligates the ribose while an amide group of the glutamine (Gln404) from motif III binds the γ phosphate of NTP. Hence, motif III is referred to as the ‘sensor I’ motif [71]. Asp401 of motif III modulates interaction between domains by forming the salt bridge to lysine (Lys309) of domain 2.

Motif IV, which is not strongly conserved in the SARS CoV helicase, probably interacts with ssRNA.

Table 2 Conserved motifs of the SARS CoV nsp13 and alignment between the SARS CoV helicase sequence and the PcrA (1PJR [39]) helicase sequence

Residue # alignment	Motif # function
SARS 282–289 PcrA 31–38 GPPGTGKS AGAGSGKT	Motif I Lys288 binds β phosphate of NTP Ser289 ligates the Mg^{2+} ion of NTP
SARS 374–378 PcrA 223–227	Motif II Asp374 coordinates the Mg^{2+} ion of NTP, Glu375 is a catalytic base in NTP hydrolysis, Met378 makes hydrogen bond with Gln537
DEISM DEYQD SARS 400–405 PcrA 250–255 GDPAQL GDADQS	Motif III Asp401 forms salt bridge with Lys309 Gln404 binds γ phosphate of NTP
SARS 510–520 PcrA 355–365 VFISPYNSQNA AVLYRTNAQSR	Motif IV Gln516 probably interacts with ssRNA
SARS 534–540 PcrA 565–571	Motif V Gln537 makes hydrogen bond with Met378, Glu540 ligates ribose of NTP
DSSQGSE HAAKGLE SARS 563–570 PcrA 606–613 VAITRAKI VGITRAEE	Motif VI R568 forms salt bridge with phosphate of NTP

Contrary to this, most of the residues from the motif VI (VAITRAKI) are conserved. The guanidinium group of the middle arginine (Arg567) forms a salt bridge with the γ phosphate of NTP. The key role of this residue was confirmed by mutagenesis experiments. The results carried out on eIF4A helicase support a model in which the arginine interacts with NTP [78].

Not conserved motifs

Other motifs (Ia, the second part of III, TxGx and QxxR) are not conserved in SARS CoV Hel. Key residues (e.g. F64, Y257, W259, H587, and F626 in 1PJR) from these motifs in the PcrA/Rep/RecB helicases bind ssDNA during unwinding duplex oligonucleotide. These residues are also important during unidirectional translocation, which has 3'–5' polarity [74]. The lack of their conservation confirms 5'–3' polarity of the SARS CoV Hel.

Catalytic pocket identification

The position of ATP in the NTPase catalytic pocket of SARS CoV Hel should be very similar to the ATP position in PcrA helicase 3PJR [54]. Indeed, the complex of ATP with SARS CoV Hel obtained as the lowest energy structure from our calculations closely

resembles the one observed in the crystal structure of PcrA helicase (see Fig. 7) Root mean square deviation between the corresponding ATP atoms from these



Fig. 7 Energetically favored position of ATP in the ATPase catalytic pocket of SARS CoV Hel (cyan protein and yellow ATP molecule) superimposed with PcrA DNA helicase (lime) in complex with ATP (orange) as observed in crystal structure 3PJR. RMSD between the positions of the corresponding atoms of both ATP molecules was as low as 0.571 Å

two structures after optimal superposition of the two proteins is as low as 0.571 Å. Such an excellent agreement indicates that the choice of Amber force field in our calculations results in the correct identification of ATPase of SARS CoV Hel catalytic pocket.

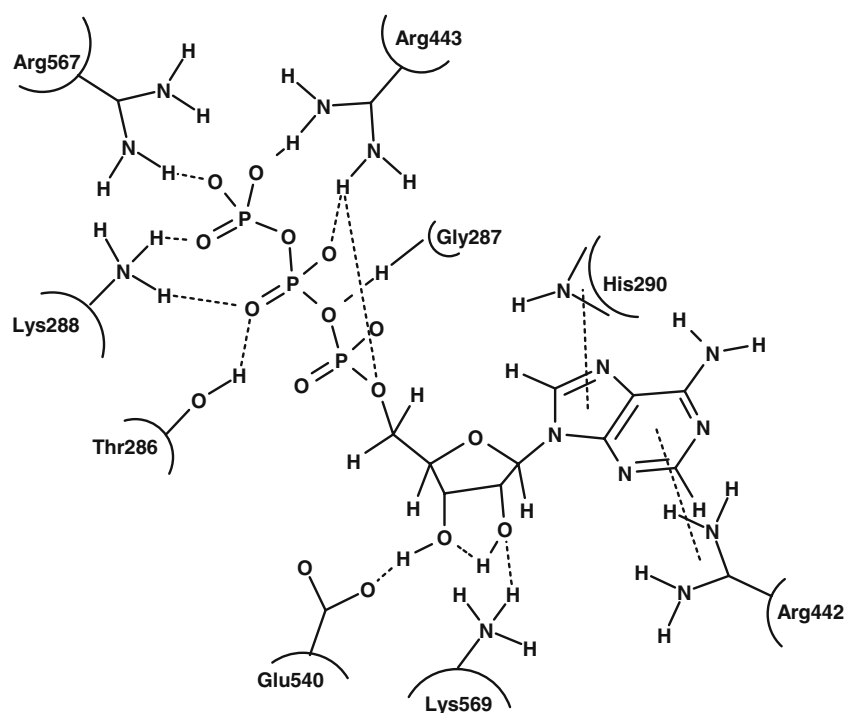
ATP binding mode inside catalytic pocket

The analysis of the lowest energy complex of ATP with SARS CoV Hel indicates that the energetically most favorable interactions are the ones that involve: (1) γ -phosphate and Arg567, Arg443, and Lys288 via charge assisted hydrogen bonds, (2) the oxygen atom bridging γ - and β -phosphates and hydrogen bonds donors: peptide NH of Gly287, and positively charged NH_2^+ of Arg443, (3) β -phosphate and peptide NH groups of Gly287 and Lys288, and OH of Thr286, (4) 3'OH of ribose and COO^- of Glu540 being an acceptor of a hydrogen bond, (5) 2'OH of ribose being an acceptor of a hydrogen bond and Lys569 serving as a donor. Moreover, adenine ring is placed in the vicinity of Arg442 and His290 (see Fig. 8). The observed pattern of interactions agrees very well with the analysis of the conserved motifs (vide supra). To our best knowledge there are two published results on SARS CoV Hel NTPase activity [54, 79]. In both of these studies authors examined the kinetics of hydrolysis of various NTP's (ATP, dATP, GTP,

dGTP, CTP, dCTP, UTP, dTTP) by SARS CoV Hel. The results obtained in these two studies resulted in totally different values. For example Michaelis constant K_m measured for ATP was 0.33 mM in one study [54] while 1.23 μM in the other [79]. The correlation coefficient between the k_{cat}/K_m recorded in both of these studies is as low as 0.18. Moreover, one study pointed at dATP closely followed by dCTP as the best substrates [54]. In the other ATP was the best substrate while dCTP second worst [79].

Our results provide information on interaction modes between SARS CoV Hel and ATP (vide supra) and indicate that one of the interactions between the enzyme and ATP is a charge assisted hydrogen bond where 3'OH of ribose serves as a donor, while negatively charged carboxylate moiety of Glu540 as an acceptor. Moreover, 2'OH of ribose is a hydrogen bond donor to 3'OH. Due to cooperativity of these hydrogen bonds ribose should interact with SARS CoV Hel stronger than deoxyribose. What is more 2'OH of ribose is an acceptor of a hydrogen bond, whose donor is the amine group of Lys569. Therefore, the involvement of two OH groups in the interaction with the protein indicates that ribose is more strongly bound than deoxyribose. This in turn suggests that ribonucleotide triphosphates should be better substrates than the corresponding deoxyribonucleotide triphosphates and seems to favor the results obtained by Ivanov et al. [79].

Fig. 8 Schematic representation of main interactions between SARS CoV Hel amino acid residues and ATP molecule. The residues within 3.6 Å distance from the ATP molecule are shown



Putative SARS CoV helicase inhibitors

As it was mentioned earlier, mutations in motives, which are responsible for the binding of the terminal phosphate group, eliminate the NTP hydrolytic activity of the helicase. On the basis of these results one could expect that a reduction of accessibility of the NTP-binding site for NTP should lead to inhibition of the NTPase and, consequently, of the helicase activity of the SARS CoV NTPase/helicase [80]. Therefore, the compounds that reduce the activity of the enzyme could act as inhibitors of virus replication [81]. In this study we attempted the design of molecules that are expected to bind in the catalytic pocket. After, we had identified attractors for various functional groups and selected the positions where these groups are most strongly bound, we searched CSD [58] for molecules having the desired functional groups in a given position (see Fig. 9). We searched for molecules having three or two functional groups at a given distance from one another. Because phosphate groups were bound at least twice as strongly as any other small molecule (no surprise owing the role in ATP hydrolysis) we focused on searching for molecules possessing at least one P atom. By modifying the obtained hits, we designed a set of molecules that are suspected to bind to the SARS CoV Hel (see Fig. 10). The subsequent calculations for these compounds revealed the binding

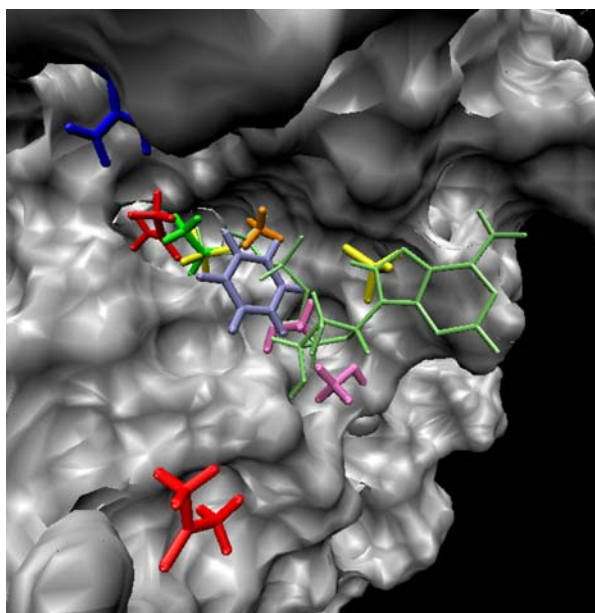


Fig. 9 Attractors for various functional groups around ATP molecule in the catalytic pocket of SARS CoV helicase. ATP: lime, CH_3COO^- : green, ketone: red, PO_4^{3-} : yellow, $\text{C}(\text{NH}_2)_3^+$: blue, CH_4 : orange, CH_3OH : mauve, C_6H_6 : iceblue

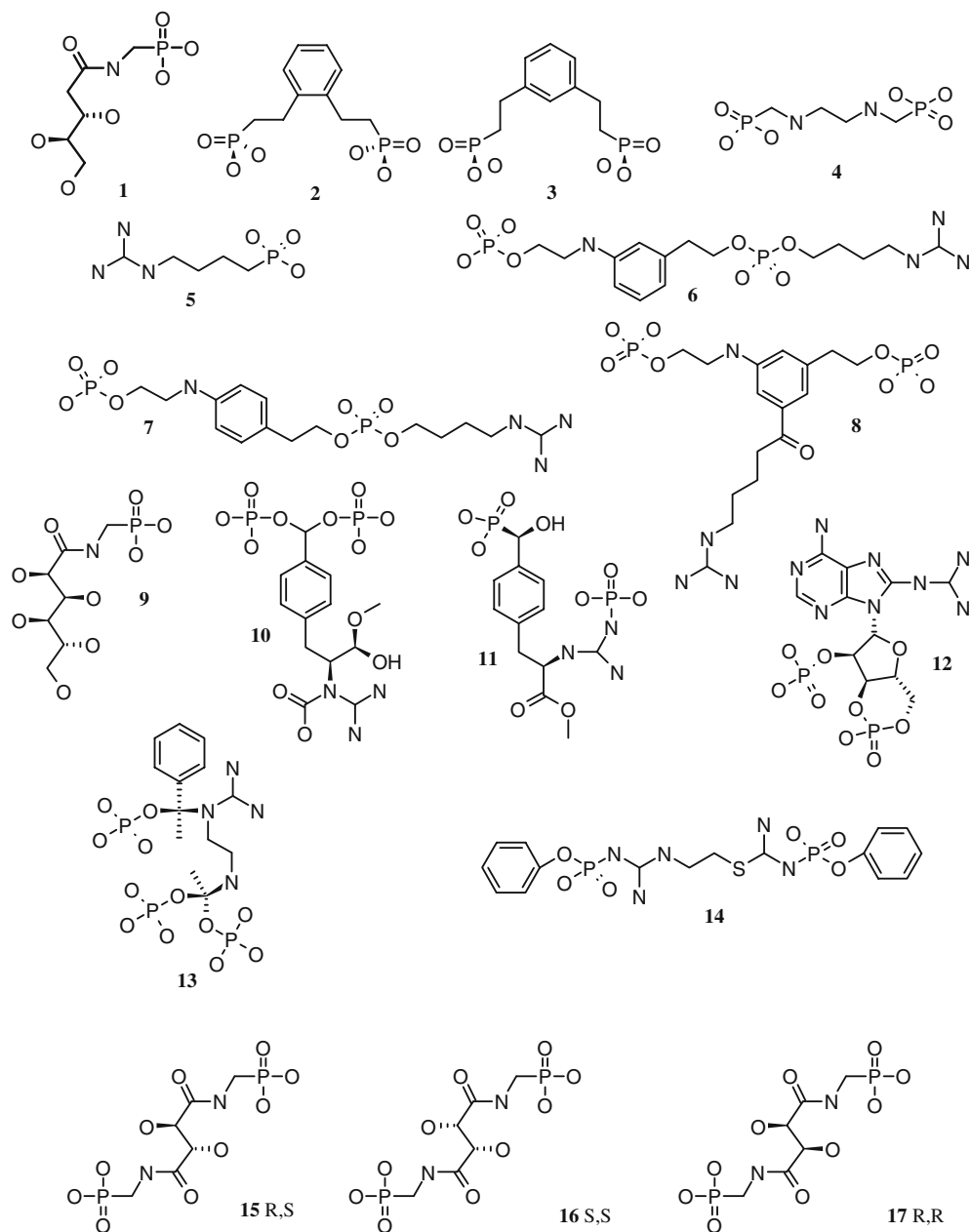
modes of these molecules in the catalytic pocket and provided data for assessment of binding energies.

The binding energies obtained in Amber force field calculations seem to be non-sensually high, in particular when compared with AM1 and PM3 values. Nevertheless, all computational methods predict that in the set of the investigated molecules there are molecules that interact with SARS CoV Hel stronger than ATP. Comparison of the results obtained from different computational methods (see Table 3 for AM1 and PM3 results and Table 4 for Amber force field results) identifies molecules whose protein ligand interactions are more favorable than for ATP. Molecules **3**, **4**, **11**, **13**, **16**, **17** are likely to compete with ATP for the access to the active site, thus serve as inhibitors. Detailed interactions between these ligands and SARS CoV Hel are presented in Fig. 11.

In the case of molecule **3** the energy released due to interaction between the protein and **3** calculated with PM3 is 3.5 kcal/mol, while with AM1 it is 20.6 kcal/mol, as compared to 8.9 and 6.9 kcal/mol released due to interactions with ATP in PM3 and AM1 calculations, respectively. In Amber force field calculations the binding energy between the protein and **3** in aqueous solution, was by 86 kcal/mol more attractive than for ATP. However, as it was stated previously MM energies of interactions between the protein and the ligands were non-sensually large and showed strong dependence on dielectric constant values, which suggested that electrostatic component of the interaction was highly overestimated (see Supporting Informations Table S1). On the other hand, the inspection of specific interactions reveals that **3** is capable of forming strong interactions within catalytic pocket of SARS CoV Hel. In the energetically favored position the NH_2^+ group of Arg443 interacts with phosphonate group of **3** via a salt bridge. The hydrogen atom connected to C_α from Ser539 approaches the aromatic ring of **3** perpendicularly creating a $\text{CH}-\pi$ interaction. The NH groups of Lys508, and Arg507 as well as OH of Tyr541 bind the other phosphonate moiety due to salt bridges and charge assisted hydrogen bonds.

Relatively small molecule **4** forms four salt bridges with the enzyme. One PO_3^{2-} moiety interacts with NH_3^+ group of Lys288, and NH_2^+ group of Arg567, and additionally its position is stabilized due to hydrogen bond in which NH group from peptide bond of Gly285, which serves as a donor. At the other end of the molecule the second phosphonate group is involved in a salt bridge with His290 and the amine group of **4** interacts with carboxylate moiety of Glu540. The energy released as the result of the interactions between **4** and the protein is 8.4 and 15 kcal/mol, calculated with PM3 and AM1

Fig. 10 Designed set of molecules that are suspected to bind to the SARS CoV Hel



methods. PM3 predicts that **4** interacts with the protein almost as strongly as ATP (8.9 kcal/mol) while AM1 indicates that **4** interacts with the enzyme of about 8 kcal/mol more favorably than ATP (6.9 kcal/mol).

In the case of **11** energy gain due to formation of the protein ligand complex, relative to ATP, favors the complex formation and the interaction energy for **11** is more attractive than for ATP by 2.8 kcal/mol in PM3, and 6.4 kcal/mol in AM1. The binding energy (energy released upon binding of **11** with the enzyme) calculated with Amber force field in aqueous solution relative to ATP was non-sensually high: 666 kcal/mol (see Supporting Information Table 4). For this complex the

oxygen atoms from the phosphonate moiety containing the C–P bond interacts with the NH groups of Arg442 creating salt bridges. Lys569 uses its NH_3^+ as a hydrogen bond donor forming hydrogen bond with O atom connected to the same C atom as P atom of the first phosphate. The NH groups of Lys288 and Arg443 interact with the second phosphazane group which contains P–N bond. The carboxylate group of Glu540 interacts with the NH of guanidine unit of **11**, while OH group of Ser289 donates a hydrogen bond to the oxygen atoms of the ester part of **11**.

For **13** the energy released as the result of interaction between SARS CoV Hel and **13** is more favorable

Table 3 Energy released (in kcal/mol) due to interaction between the ligand and the amino acid residues that are within 3.6 Å from it, as calculated with ONIOM method. High level calculations were performed with AM1 and PM3, while low level with Amber force field

Molecule	AM1 -ΔE	PM3 -ΔE
ATP	6.9	8.9
1	9.8	5.0
2	9.6	6.5
3	20.6	3.5
4	15.0	8.4
5	1.1	9.6
6	15.6	6.0
7	15.0	4.4
8	1.4	8.8
9	9.4	2.3
10	24.0	2.3
11	13.3	11.8
12	0.8	14.7
13	43.8	23.7
14	0.8	2.2
15	21.9	5.4
16	30.1	5.9
17	31.9	10.8

$$\Delta E = E_{PL} - E_P - E_L$$

than for ATP by 14.8 kcal/mol in PM3 and 36.9 kcal/mol in AM1. Molecule **13** possesses three phosphate moieties each interacting with arginine side chains via salt bridges. It should be noted that **13** has geminal

Table 4 BE_{PL} is the energetic effect (in kcal/mol) of binding between a given ligand and SARS CoV Hel in vacuo calculated with Amber force field. The presented values $BE_{PL}^{aq,R}$ (protein ligand binding energy in aqueous solution) are relative to ATP—negative values indicate that binding between a given ligand and the protein in aqueous solution is energetically more favorable than for ATP. Non-sensually large energies were obtained, which depended strongly on dielectric constant value—see Supporting Information Table S1

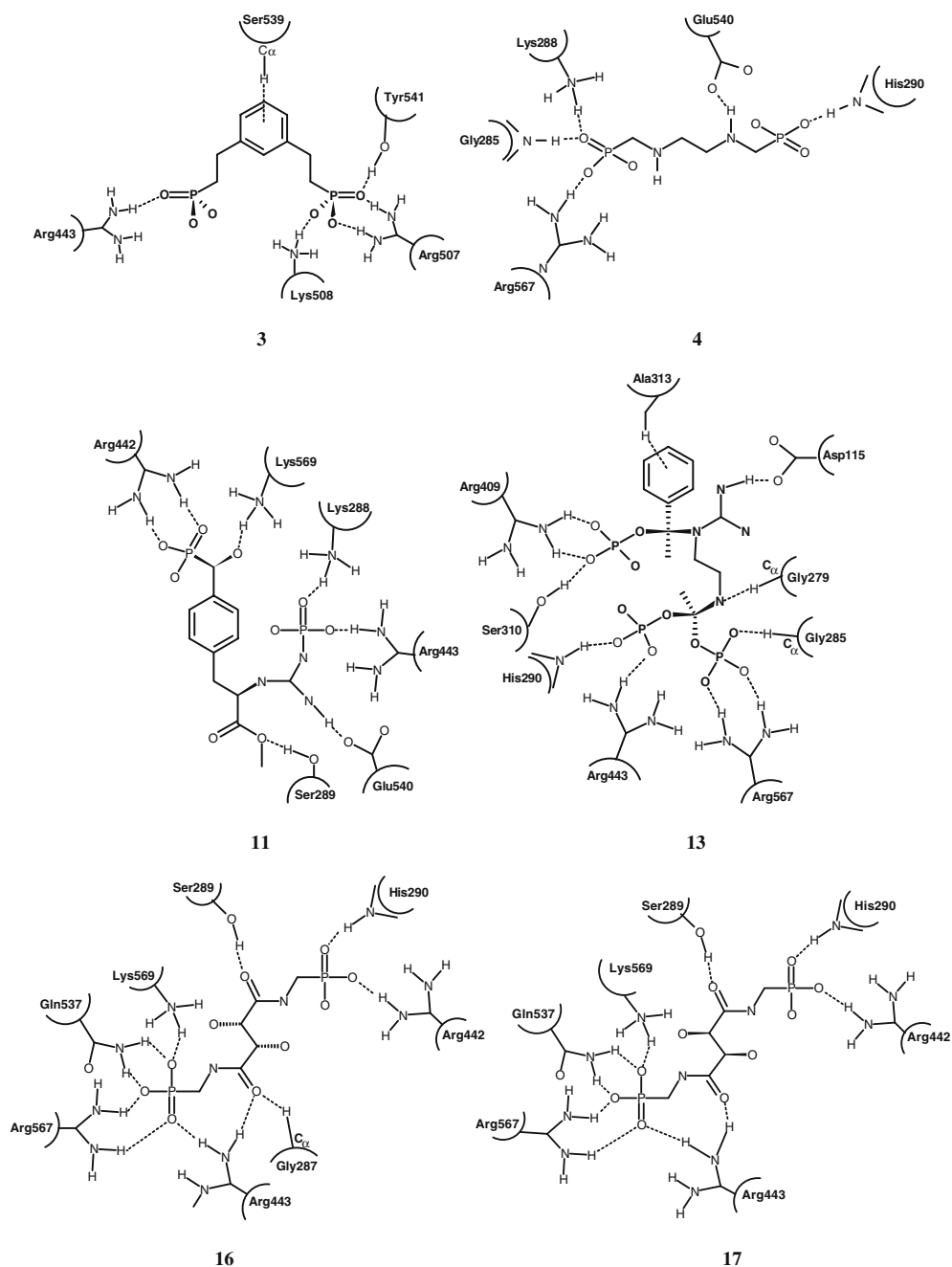
Molecule	$BE_{PL}^{aq,R}$	BE_{PL}
ATP	0.0	-1,891.8
1	454.0	-788.3
2	-45.8	-1,872.6
3	-85.9	-1,792.3
4	408.7	-895.8
5	565.3	-424.0
6	594.1	-626.9
7	666.0	-567.9
8	720.5	-1,060.6
9	702.4	-517.6
10	128.5	-1,693.9
11	116.5	-1,424.9
12	335.9	-847.8
13	-159.8	-2,075.6
14	563.3	-663.5
15	44.7	-1,715.8
16	47.8	-1,760.5
17	96.1	-1,674.5

phosphate moieties, which are not likely to be stable in aqueous solution. However, structures deposited in CSD are biased towards unusual molecules and the molecules with geminal phosphates were present in the set of molecules that we obtained in our search (HIBYIA, QEQNOP, QURYAD, XOBMOQ). Nevertheless, this molecule in all computational methods employed in our study binds to SARS CoV Hel more strongly than ATP. The phosphate closer to phenyl ring interacts with Arg409 and additionally with OH group of Ser310. Another phosphate forms a bonding with NH group of Arg443 and His290, while the last phosphate forms a salt bridge with NH group of Arg567 and additionally much weaker interaction with CH of Gly285. NH group of guanidine part of **13** forms a salt bridge with carboxylate group of Asp115. The position of the phenyl ring of **13** is stabilized due to CH- π interaction with CH of Ala313. In addition there is a close contact between ethyleneamine nitrogen atom and hydrogen atom attached to C_α of Gly279.

Molecules **16** and **17** are (*S,S*) and (*R,R*) enantiomers of tartaric acid derivative. For these molecules the energies released due to interaction with SARS CoV Hel, calculated with PM3, are 10.8 and 5.9 kcal/mol for **17** and **16**, respectively. In AM1 calculations these values are 31.9 and 30.1 kcal/mol. In the case of ATP the corresponding energies were 8.9 and 6.9 kcal/mol for PM3 and AM1, respectively. Therefore, both PM3 and AM1 calculations indicate that **17** interacts with the protein stronger than ATP. Molecules **16** and **17** have similar modes of interaction with the enzyme. Both for **16** and **17** one of the phosphonates is involved in salt bridges from NH groups of Arg442 and His290, while the other phosphonate moiety forms salt bridge with NH groups of Arg567, Lys569 and Arg443. The latter phosphate interacts also with the amide NH group of Gln537. One of the carbonyl groups of **16** or **17** serves as an acceptor of a hydrogen bond from OH group of Ser289, while the other accepts hydrogen bonding from NH group of Arg443. In the case of **16** this carbonyl oxygen atom forms also a close contact with CH of Gly287.

All in all, the compounds bearing two phosphonic acid moieties or phosphates located at the distal ends of a molecule seem to be promising candidates for experimental studies on SARS CoV helicase inhibitors. Moreover, the search for ATP-protein interactions in the PDB-Ligand database [82] indicates that among known 205 ATP-protein complexes, the amino acids environment surrounding ATP molecule in SARS CoV Hel, namely Arg, Arg, Lys, Thr, Gly, Glu, Lys, Arg, and His, is unique among all known ATP-binding proteins deposited in PDB. This may suggest that the

Fig. 11 Schematic representation of main interactions between SARS CoV Hel amino acid residues and selected inhibitors. The residues within 3.6 Å distance from a given molecule are shown



inhibitors designed specifically for SARS CoV helicase could be specific inhibitors of this helicase not general inhibitors of any ATP binding protein.

Conclusions

The designed molecules, bearing two phosphonic acid and/or phosphate moieties at distal ends of the molecule, are likely to bind stronger to the SARS CoV helicase in the NTPase catalytic pocket than ATP,

even in aqueous solution. Thus they are expected to inhibit its enzymatic activity. Therefore, they fully deserve further experimental studies.

Supporting information available

PDB file with cartesian coordinates of atoms of SARS CoV Hel. Figure presenting NTPase catalytic pocket with ATP and attractors for various functional groups in it (PO_4^{3-} yellow, CH_4 orange, ketone

red, CH_3COO^- green, $\text{C}(\text{NH}_2)_3^+$ blue, C_6H_6 ice blue, and CH_3OH pink). Figure comparing interactions between SARS CoV Hel and ATP with PcrA Hel and ATP. Table S1 presenting dependence of binding energies on various values of dielectric constant. Table S2 splitting Amber calculated aqueous binding energies into components.

Acknowledgment Financial support from European committee grant no. SP22-CT-2004-003831 and Polish Ministry for Science is gratefully acknowledged. K. E. and M. v. G. thank the Foundation for Polish Science for a fellowship. Calculations were performed in Poznan Supercomputing and Networking Center (PCSS).

References

- Wu CH, Apweiler R, Bairoch A, Natale DA, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Mazumder R, O'Donovan C, Redaschi N, Suzek B (2006) *Nucleic Acids Res* 34(Database issue):D187–D191
- Berman H, Henrick K, Nakamura H (2003) *Nat Struct Biol* 10(12):980
- Neuman BW, Stein DA, Kroeker AD, Churchill MJ, Kim AM, Kuhn P, Dawson P, Moulton HM, Bestwick RK, Iversen PL, Buchmeier MJ (2005) *J Virol* 79(15):9665–9676
- http://www.who.int/csr/resources/publications/WHO_CDS_CSR_ARO_2004_1/en/index.html
- Kliger Y, Levanon EY, Gerber D (2005) *Drug Discov Today* 10(5):345–352
- Bujnicki JM, Elofsson A, Fischer D, Rychlewski L (2001) *Bioinformatics* 17(8):750–751
- <http://predictioncenter.org/>
- Ginalski K, Rychlewski L (2003) *Proteins* 53(Suppl 6):410–417
- Ginalski K, Grishin NV, Godzik A, Rychlewski L (2005) *Nucleic Acids Res* 33(6):1874–1891
- Sali A, Blundell TL (1993) *J Mol Biol* 234(3):779–815
- Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson JDM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) *J Am Chem Soc* 117:5179–5197
- Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M (1995) *Proteins* 23(3):318–326
- Spector FC, Liang L, Giordano H, Sivaraja M, Peterson MG (1998) *J Virol* 72(9):6979–6987
- Ravna AW, Schroder KE, Edvardsen O (1999) *Comput Chem* 23(5):435–437
- Ljungberg KB, Marelius J, Musil D, Svensson P, Norden B, Aqvist J (2001) *Eur J Pharm Sci* 12(4):441–446
- Egan WJ, Walters WP, Murcko MA (2002) *Curr Opin Drug Discov Dev* 5(4):540–549
- Plewczynski D, Pas J, von Grotthuss M, Rychlewski L (2002) *Appl Bioinf* 1(4):223–225
- Shen X, Xue JH, Yu CY, Luo HB, Qin L, Yu XJ, Chen J, Chen LL, Xiong B, Yue LD, Cai JH, Shen JH, Luo XM, Chen KX, Shi TL, Li YX, Hu GX, Jiang HL (2003) *Acta Pharmacol Sin* 24(6):505–511
- Oda A, Yamaotsu N, Hirono S (2005) *J Comput Chem* 26(8):818–826
- Joart B, Martinek TA, Marki A (2005) *J Comput Aided Mol Des* 19(5):341–356
- Park H, Lee S (2004) *J Comput Aided Mol Des* 18(6):375–388
- Vreven T, Morokuma K, Farkas O, Schlegel HB, Frisch MJ (2003) *J Comput Chem* 24(6):760–769
- Saen-oon S, Kuno M, Hannongbua S (2005) *Proteins* 61(4):859–869
- Wieczorek R, Dannenberg JJ (2005) *J Am Chem Soc* 127(42):14534–14535
- Hoffmann M, Khavrutskii IV, Musaev DG, Morokuma K (2004) *Int J Quant Chem* 99(6):972–980
- Burgoyne NJ, Jackson RM (2006) *Bioinformatics* 22(11):1335–1342
- Nemoto T, Fedorov DG, Uebayasi M, Kanazawa K, Kitaura K, Komeiji Y (2005) *Comput Biol Chem* 29(6):434–439
- Klingenberg M (2005) *Biochemistry* 44(24):8563–8570
- Waldron TT, Schrifft GL, Murphy KP (2005) *J Mol Biol* 346(3):895–905
- Nissink JW, Murray C, Hartshorn M, Verdonk ML, Cole JC, Taylor R (2002) *Proteins* 49(4):457–471
- Swegat W, Schlitter J, Kruger P, Wollmer A (2003) *Biophys J* 84(3):1493–1506
- Ginalski K, Pas J, Wyrwicz LS, von Grotthuss M, Bujnicki JM, Rychlewski L (2003) *Nucleic Acids Res* 31(13):3804–3807
- Karplus K, Karchin R, Barrett C, Tu S, Cline M, Diekhans M, Grate L, Casper J, Hughey R (2001) *Proteins* 5(Suppl): 86–91
- Jaroszewski L, Rychlewski L, Li Z, Li W, Godzik A (2005) *Nucleic Acids Res* 33(Web Server issue):W284–W288
- Jones DT (1999) *J Mol Biol* 287(4):797–815
- Fischer, D (2000) *Pac Symp Biocomput* 5:119–130
- Shi J, Blundell TL, Mizuguchi K (2001) *J Mol Biol* 310(1):243–257
- Kelley LA, MacCallum RM, Sternberg MJ (2000) *J Mol Biol* 299(2):499–520
- Hogbom M, Andersson ME, Nordlund P (2001) *J Biol Inorg Chem* 6(3):315–323
- Korolev S, Hsieh J, Gauss GH, Lohman TM, Waksman G (1997) *Cell* 90(4):635–647
- Singleton MR, Dillingham MS, Gaudier M, Kowalczykowski SC, Wigley DB (2004) *Nature* 432(7014):187–193
- Eswar N, John B, Mirkovic N, Fiser A, Ilyin VA, Pieper U, Stuart AC, Marti-Renom MA, Madhusudhan MS, Yerkovich B, Sali A (2003) *Nucleic Acids Res* 31(13):3375–3380
- Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A (2000) *Annu Rev Biophys Biomol Struct* 29:291–325
- Fox T, Kollman PA (1996) *Proteins* 25(3):315–334
- Cheatham TE, Cieplak P, Kollman PA (1999) *J Biomol Struct Dyn* 16(4):845–862
- Weiner SJ, Kollman PA, Nguyen DT, Case DA (1986) *J Comput Chem* 7(2):230–252
- Cornell WDCP, Bayly CI, Gould IR, Merz KM, Ferguson DM Jr, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) *J Am Chem Soc* 117:5179–5197
- Ren P, Ponder JW (2003) *J Phys Chem B* 107:5933–5947
- Ren P, Ponder JW (2002) *J Comput Chem* 23(16):1497–1506
- Pappu RV, Hart RK, Ponder JW (1998) *J Phys Chem B* 102:9725–9742
- Hodsdon ME, Ponder JW, Cistola DP (1996) *J Mol Biol* 264(3):585–602
- Meagher KL, Redman LT, Carlson HA (2003) *J Comput Chem* 24(9):1016–1025
- Taylor RD, Jewsbury PJ, Essex JW (2002) *J Comput Aided Mol Des* 16(3):151–166
- Tanner JA, Watt RM, Chai YB, Lu LY, Lin MC, Peiris JS, Poon LL, Kung HF, Huang JD (2003) *J Biol Chem* 278(41):39578–39582

55. Joseph-McCarthy D, Alvarez JC (2003) *Proteins* 51(2):189–202
56. Miranker A, Karplus M (1995) *Proteins* 23(4):472–490
57. Caffisch A, Miranker A, Karplus M (1993) *J Med Chem* 36(15):2142–2167
58. van de Streek J (2006) *Acta Crystallogr B* 62(Pt 4):567–579
59. Allen FH (2002) *Acta Crystallogr B* 58(Pt 3 Pt 1):380–388
60. Allen FH, Motherwell WDS (2002) *Acta Crystallogr B* 58(Pt 3):407–422
61. CDS refcodes of the structures: ABITOU, AEADMP, AMAFAP, BACVAD, BAYHAK, BEBCEQ, BEJRAJ01, BERDIL, BERDOR, BIBTIP, BOLHEP, BUFDEL, BUNYOY, CACLAC, CAGUCP10, CEPBUU, CIDSUD, CITWUX10, CIZCUJ, COWDOH, COYDUP, CYACPH10, DEFLAB, DINYIII10, DUMJEA10, DUVDON10, EHOWUT, EJETES, EVOYET, EWAHAL, FELSUK, FILCUY, FILDAF, GLCTSM, GLPCHO, GUPCYT20, HIBYIA, HUNBUN, ISOTEP, LAGXIA, LASDEO, LENCEM, LIHLET, LOTLEL, LUBLID, MAHNAK, NUKFII, PIDPHA10, QAZMIN, QEQNOP, QIXMEP, QURYAD, REBPET, SOYCUE, SPGLUC01, SUSBEN01, TAPFUL, THYTHY10, TICNIC, TOSDOU, TYRPX110, UKOWEW, XABSEZ, XOBMOQ, XUSBOC, YEJJUS, YITJOA
62. Cramer CJ, Truhlar DG (1999) *Chem Rev* 99(8):2161–2200
63. Frisch MJ, Schlegel GWTHB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMWB, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian, Inc, Wallingford
64. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) *J Am Chem Soc* 107(13):3902–3909
65. Dewar MJS, Reynolds CH (1986) *J Comput Chem* 7(2):140–143
66. Dewar MJS, Mckee ML, Rzepa HS (1978) *J Am Chem Soc* 100(11):3607–3607
67. Stewart JJP (1989) *J Comput Chem* 10(2):209–220
68. Stewart JJP (1989) *J Comput Chem* 10(2):221–264
69. Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, Tong S, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TC, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rasmussen M, Fouchier R, Gunther S, Osterhaus AD, Drosten C, Pallsansch MA, Anderson LJ, Bellini WJ (2003) *Science* 300(5624):1394–1399
70. Ginalski K, Elofsson A, Fischer D, Rychlewski L (2003) *Bioinformatics* 19(8):1015–1018
71. Caruthers JM, McKay DB (2002) *Curr Opin Struct Biol* 12(1):123–133
72. Velankar SS, Soutlanas P, Dillingham MS, Subramanya HS, Wigley DB (1999) *Cell* 97(1):75–84
73. Gorbalenya AE, Koonin EV (1993) *Curr Opin Struct Biol* 3:419–429
74. Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) *Embo J* 1(8):945–951
75. Hall MC, Matson SW (1999) *Mol Microbiol* 34(5):867–877
76. George JW, Brosh RM Jr, Matson SW (1994) *J Mol Biol* 235(2):424–435
77. Story RM, Steitz TA (1992) *Nature* 355(6358):374–376
78. Pause A, Methot N, Sonenberg N (1993) *Mol Cell Biol* 13(11):6789–6798
79. Ivanov KA, Thiel V, Dobbe JC, van der Meer Y, Snijder EJ, Ziebuhr J (2004) *J Virol* 78(11):5619–5632
80. Borowski P, Lang M, Niebuhr A, Haag A, Schmitz H, zur Wiesch JS, Choe J, Siwecka MA, Kulikowski T (2001) *Acta Biochim Pol* 48(3):739–744
81. Borowski P, Deinert J, Schalinski S, Bretner M, Ginalski K, Kulikowski T, Shugar D (2003) *Eur J Biochem* 270(8):1645–1653
82. Shin JM, Cho DH (2005) *Nucleic Acids Res* 33(Database issue):D238–241
83. Humphrey W, Dalke A, Schulten K (1996) *J Mol Graph* 14(1):33–38–27–38