ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2021, *66*(10-11), 845–853
DOI: 10.33224/rrch.2021.66.10-11.08

Dedicated to the memory of
Prof. Petre T. FRANGOPOL (1933-2020)

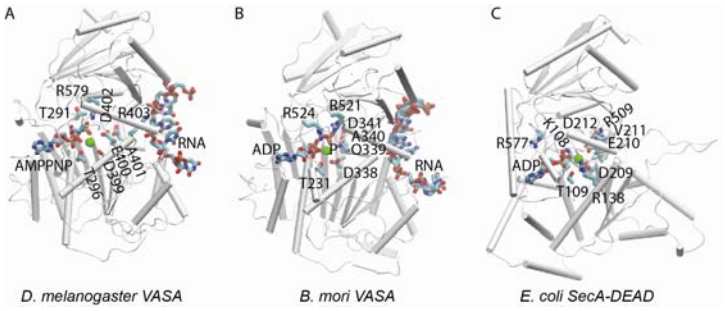
CONFORMATIONAL COUPLING VIA HYDROGEN-BONDING IN THE DEAD-BOX PROTEIN VASA

Konstantina KARATHANOU[a] and Ana-Nicoleta BONDAR[a,b,c]

[a]Freie Universität Berlin, Department of Physics, Theoretical Molecular Biophysics Group, Arnimallee 14, D-14195 Berlin, Germany
[b]University of Bucharest, Faculty of Physics, Atomiştilor 405, Măgurele 077125, Roumania
[c]Forschungszentrum Jülich, Institute of Computational Biomedicine, IAS-5/INM-9, Wilhelm-Johnen Straße, 5428 Jülich, Germany

Received September 30, 2021

The VASA DEAD-box helicase couples binding and hydrolysis of adenosine triphosphate with unwinding of RNA. We use atomistic simulations and graph-based analyses to explore the role of hydrogen-bond networks in long-distance conformational coupling between the ATP- and RNA-binding sites. We identify local H-bond networks at the ATP-binding site, and find that these networks can transiently connect via short, water-mediated bridges, to the RNA-binding site. H-bonds and H-bond networks identified here could help ensure long-distance conformational couplings in VASA.



INTRODUCTION

Asp-Glu-Ala-Asp (DEAD)-box motor RNA helicases are enzymes that couple binding and hydrolysis of adenosine triphosphate (ATP) with binding and alteration of RNA such as unwinding.[1] The family of DEAD-box helicases includes the cytoplasmic eukaryotic translation factor 4A (elF4A) that unwinds mRNA to facilitate binding to the ribosome,[2] and the cell-nucleus protein p68 –an RNA-unwinding enzyme[3] overexpressed and poly-ubiquitylated in colon cancer[4] and thus of potential interest as biomarker and drug target.[5] Host DEAD-box helicases might also be used by coronaviruses, including SARS-CoV-2, for viral replication,[6] and zotatifin, an mRNA-translation inhibitor of elF4A, reduces infectivity of SARS-CoV-2.[7] Another DEAD-box helicase, SecA, functions as pre-protein motor essential for protein bio-synthesis in bacteria.[8, 9] Description of the reaction mechanism of DEAD-box enzymes is thus valuable, as it could guide the development of therapeutics.

The reaction mechanism of an enzyme is given by the sequence of events and associated free energy profile along the reaction path from the reactant to the product state via the transition state(s).[10] In the case of DEAD-box helicases, the reaction coordinate involves electronic structure changes at the nucleotide-binding site, binding and unwinding of RNA, and conformational coupling between the nucleotide- and RNA-binding sites. Here we rely on molecular dynamics (MD) simulations to identify

* Corresponding author: nbondar@fizica.unibuc.ro; ana.bondar@unibuc.ro

Hydrogen(H) bond networks that could ensure conformational coupling between the ATP- and RNA-binding sites the *Drosophila melanogaster* VASA, a DEAD-box enzyme that might be implicated in the cell cycle.[11]

The three-dimensional structure of the *D. melanogaster* VASA was solved bound to an ATP analogue[12] and an RNA fragment.[12] The non-hydrolyzable ATP analogue AMPPNP and the magnesium ion bind deep into the binding cleft, close to the groups of the DEAD motif; adjacent in the sequence, R403 connects to the RNA fragment (Figure 1A). An overall similar arrangement was observed in the structure of the helicase core domain of the *Bombix mori* VASA with the DEAD-box Glu sidechain mutated to Gln (E339Q mutant), solved in the presence RNA, ADP, and inorganic phosphate (Figure 1B).[13]

Protein and water interactions at the nucleotide-binding site of the two VASA structures are similar to those seen in the ADP-bound structure of the DEAD motor domain of the *Escherichia coli* SecA (Figure 1C). The DEAD motor domain of SecA, which consists of the two nucleotide-binding domains, has ATPase activity, i.e., it can be used as a model system to probe protein-nucleotide interactions.[14] We have shown, using extensive atomistic MD simulations of the *E. coli* DEAD motor, that a dynamic H-bond network guides the binding of the nucleotide, and stabilizes it at the binding site.[15] Importantly, recent experiments and computations on SecA suggest that dynamic H-bond networks ensure long-distance conformational couplings associated with biological function.[16]

Analyses of dynamic H-bond networks for protein conformational couplings bring about the challenge of identifying transient H-bond paths that may connect remote regions of the protein. To facilitate analyses of dynamic H-bond networks in complex protein environments, we implemented efficient graph-based algorithms.[18-20] A graph representation of the protein H-bond network consists of nodes, which are H-bonding protein groups, and edges, which are the H-bonds between these groups. Once the H-bond graph has been computed, it can be filtered according to, e.g., occupancy criteria that indicate the percentage of time during which H-bonds are present. Or, for proteins such as DEAD-box enzymes, for which long-distance conformational coupling is essential, the protein H-bond graph may be queried to identify H-bond networks that may assemble transiently and connect different regions of the protein.

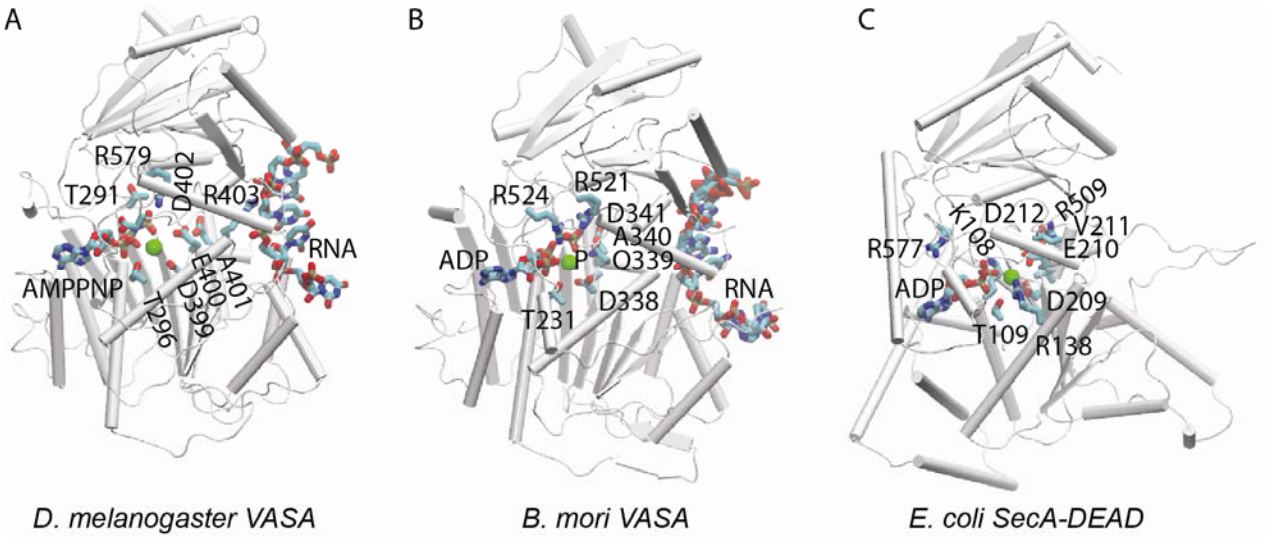


Fig. 1 – ATP and RNA binding sites of VASA. For clarity, the protein is shown as white cartoons. Nucleotide, magnesium ion, RNA fragment, and selected protein sidechains are shown as bonds with carbon atoms colored cyan, nitrogen, blue, oxygen, red, and phosphorous, brown; the magnesium ion is shown as a green sphere. (A) *D. melanogaster* VASA bound to AMPPNP, magnesium ion, and RNA fragment. The molecular graphics is based on structure PDB ID:2DB3 chain A, solved at a resolution of 2.2Å.[12] (B) E339Q mutant of *B. mori* VASA bound to ADP, inorganic phosphate (labeled P), magnesium ion and RNA fragment, structure PDB ID:4D26, solved at a resolution of 2.1Å.[13] (C) ADP and magnesium ion interactions in the nucleotide binding pocket of the DEAD-motor domain of the *E. coli* SecA, structure PDB ID:3BXZ, solved at a resolution of 3Å.[14] All molecular graphics were prepared with Visual Molecular Dynamics, VMD.[17]

To find out whether VASA might use dynamic H-bond networks for conformational couplings, here we use atomistic MD simulations to probe protein motions, and subject the MD simulation trajectories to H-bond analyses. We find that the protein hosts an extensive H-bond network that includes H-bonds that are rather short, which is indicative of strong interactions. Computations of H-bond graphs indicate that DEAD-box amino acid residues are part of networks of dynamic protein-water H-bond bridges and H-bond clusters that could relay to the protein changes in local interactions upon ATP binding and hydrolysis.

METHODS

ATP-bound VASA

Starting coordinates for ATP-bound VASA were taken from the crystal structure PDB ID:2DB3, chain A.[12] This chain includes coordinates for 420 amino acid residues (Y202 to R621), ANP, a magnesium ion, water molecules, and a poly-uracil RNA fragment. We included in the simulation system all protein atoms and 256 crystallographic water molecules. The first and the last of the seven monomers from the poly-uracil RNA fragment are incomplete, and we thus modeled a poly-uracil chain consisting of 5 monomers. To model ATP, the Nβ atom of ANP was replaced by oxygen; we included the coordinates of the magnesium ion. Hydrogen atoms were constructed using CHARMM (Chemistry at Harvard Molecular Mechanics).[21] The system consisting of the VASA protein bound to ATP, poly-uracil, ATP, magnesium, and crystallographic waters, was placed in the center of a box of water molecules, and bulk waters overlapping with the crystal-structure system were deleted. Sodium ions were added for charge neutrality. The complete simulation system for VASA includes 209.676 atoms.

Force field and MD simulation protocol

We used the CHARMM36 force-field parameters with CMAP corrections[21-27] for the protein and nucleotide atoms, and the TIP3P water model.[28] All MD simulations were performed with NAMD[29, 30] with fixed lengths of covalent bonds to H atoms.[31] Heating was performed in the *NVT* ensemble (constant number of atoms *N*, constant volume *V*, and constant temperature *T*), and the production run, in the *NPT* ensemble (constant pressure P = 1bar) with isotropic pressure coupling using a Langevin dynamics scheme and a Nosé-Hoover piston.[32, 33] We used a switch function between 8Å and 12Å for the short-range real-space interactions, and smooth particle mesh Ewald for Coulomb interactions.[34,35] During heating to 300K and the first 1ns of the simulations we used an integration step of 1fs; for the remaining of the production run we used a multiple time-step integration scheme with 1fs for the bonded forces, 2fs for short-range nonbonded, and 4fs for long-range electrostatics.[36] Coordinates were saved each 10ps.

H-bond criteria and H-bond graph computations

As geometric criteria for H-bonding we use the distance d_{HA} between the H atom and the acceptor heavy atom; alternatively, we use the distance d_{AD} between the H-bond donor and acceptor atoms, and the H-bond angle θ. Our previous test computations on coordinate sets from a SecA simulation have indicated that H-bond graphs computed with the combined criteria $d_{AD} \leq 3.5$Å and $\theta \leq 60^{\circ}$ are largely the same as obtained with the single distance criterion $d_{HA} \leq 2.5$Å[37]

As a simple measure of the strength of H-bonding, here we proceed as in our previous work on SecA[18] and consider H-bonds as weak, medium, or strong when, respectively, $d_{HA} \leq 1.7$Å, 1.7Å < $d_{HA} \leq$ 1.9Å, and 1.9Å < $d_{HA} \leq$ 2.5Å. For each H-bond pair we monitor the H-bond distance along the simulation trajectory and count how many times the H-bond distance is in one of these three intervals. An H-bond is considered as strong when the number of times $d_{HA} \leq 1.7$Å is larger than 25% of the sum of the number of times d_{HA} met the criteria for weak or medium strength. A similar condition is verified for medium-strength H-bonds: For each H-bond pair during the simulation, when the number of times 1.7Å < $d_{HA} \leq$ 1.9Å was larger than 25% of the sum of strong and weak H bonds, the H bond was considered of medium strength. When neither the medium not the strong H-bond criteria are met, and d_{HA} is within the interval noted above for weak H-bonds, the H-bond is considered as weak.[18] These analyses of H-bond strength were performed using the last 50ns of the simulation. The H-bond graphs were computed with Bridge[19, 20] using 6215 equally spaced coordinate snapshots from the last segment of the simulations.

To visualize H-bond graphs we used the Bridge2 graphical user interface[19, 20] and the combined distance and angle criteria $d_{DA} \leq 3.5$Å and $\theta < 60°$. We computed H-bonds between protein sidechains and one-water bridges between protein sidechains.

RESULTS AND DISCUSSION

The overall structure of VASA is stable throughout the simulations, with Cα root-mean-square distance within 2Å for both the structured (Figures 2A, 3B) and loops/termini regions of the protein (Figure 2A, 3C). The ATP and the magnesium ion remain bound at the nucleotide-binding cleft between the N-terminal (NTD) and C-terminal (CTD) domains (Figures 2A, 3A); likewise, the RNA fragment remains bound to VASA (Figure 2A). During the simulations, both the ATP molecule and the RNA fragment interact with water molecules (Figure 2A).

We suggest that H-bonds between amino acid residue sidechains largely contribute to the protein structural stability indicated by the RMSD profiles. There are numerous H bonds between protein sidechains, such that the entire protein hosts H-bonds (Figures 3D-F). Most of these H bonds are dynamic, with relatively small occupancies (Figure 3E) and d_{HA} values in the intervals for weak- or medium-strength H-bonding (Figure 3F). H-bonds with high occupancies and relatively short distances, indicative of stronger interactions, are also observed throughout the protein (Figure 3F).

Persistent H-bonding at the DEAD-box and the RNA-binding regions

We used Bridge2[20] to compute separately the H-bond network for protein-sidechains (Figure 4) and for one-water bridges between sidechains (Figure 5). The graph of direct H-bonds between protein sidechains contains numerous persistent interactions, i.e., with high occupancies (Figure 4). The DEAD-box carboxylates E400 and D402 are part of a highly stable local H-bond cluster with S343 and R551 (Figures 4A,B). At the end of the simulation, E400 is within ~4Å from the γ-phosphate group of ATP; within <3Å the γ-phosphate group, R582 connects to D554 and K556 (Figures 4A,B). Near the DEAD-box, D571 interacts transiently with K569 – which very infrequently connects to D406 (Figure 4D); D571 is adjacent in the sequence with D572, which in turn is part of a local cluster with E497 and R576 (Figures 4 A, B).

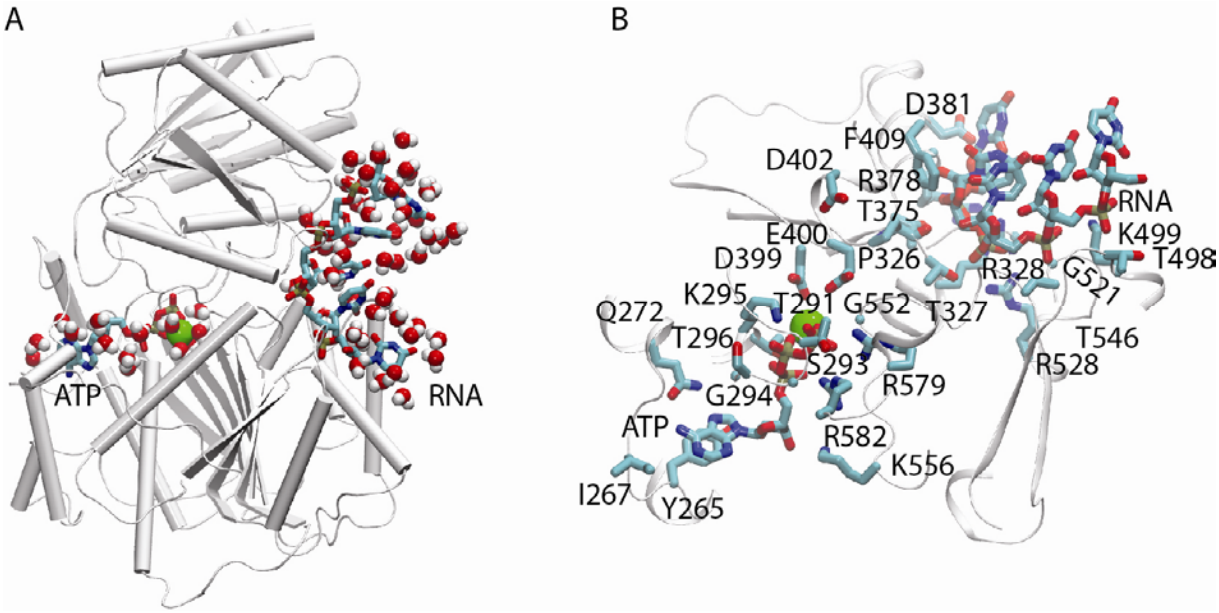


Fig. 2 – ATP and RNA remain bound to *D. melanogaster* VASA during simulations. (A) Coordinate snapshot from the end of the MD simulations of D. melanogaster VASA in water box. The ATP molecule and the RNA fragment, shown as bonds, remain bound to the protein and interact with water molecules, which are shown as small spheres; the magnesium ion, shown as van der Waals sphere colored green, remains at its binding site. (B) Close view of ATP and RNA with nearby protein groups. In this coordinate snapshot, the closest distance between the gamma-phosphate atom of ATP and an RNA phosphate atom is ~17Å. For clarity, H atoms are not shown.

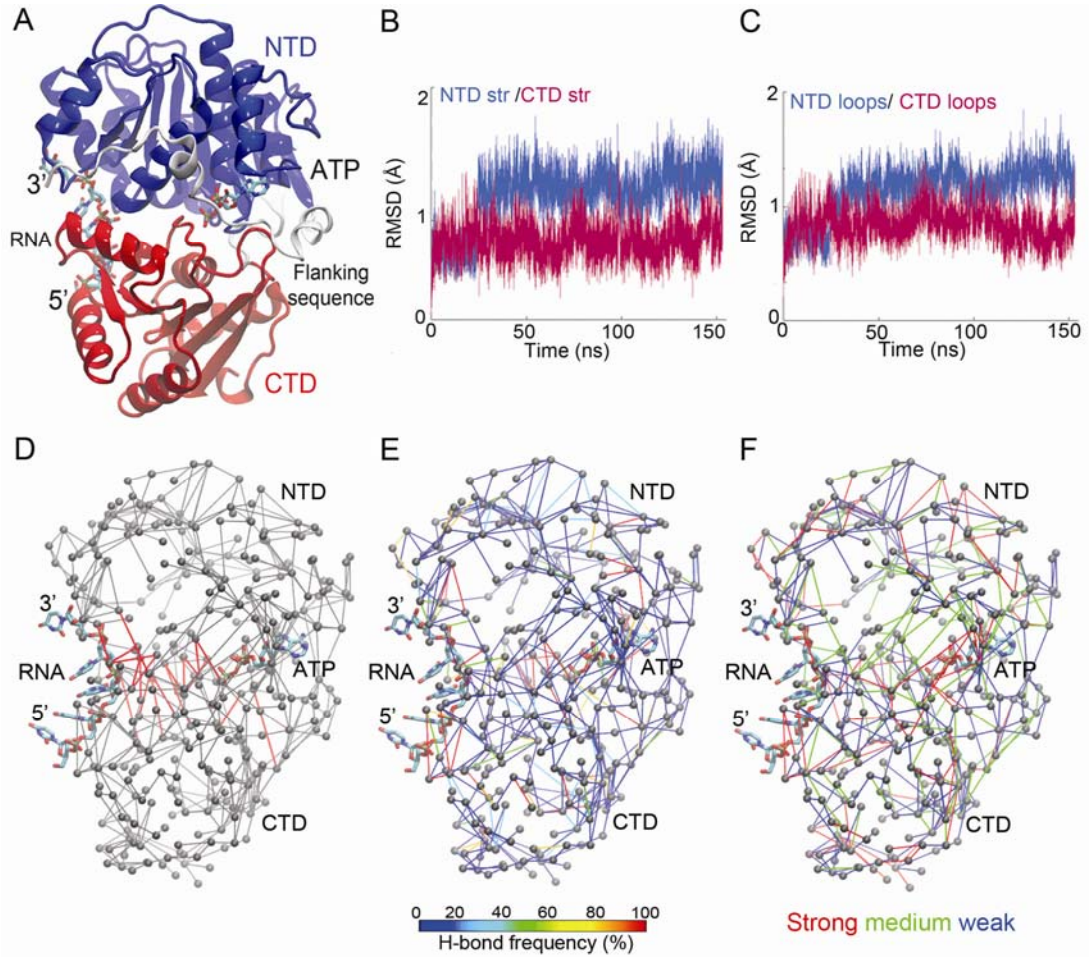


Fig. 3 – Architecture and conformational dynamics of *D. melanogaster* VASA. (A) Molecular graphics of VASA indicating its functional domains, based on the starting crystal structure.[12] The flanking sequence extends from amino acid residue 202 to 232; the N-terminal domain (NTD) and the C-terminal domain (CTD) extend from V233 to A454 and, respectively, from S463 to R621.[12] ANP and RNA are shown as bonds. (B, C) Cα RMSD profiles for structured (panel A) vs. loops and termini (panel B) of VASA as computed from the simulations. (D-F) H-bonds computed for ATP-bound VASA during the simulations. Cα atoms of the amino acid residues involved in H bonding are shown as grey spheres using the last coordinate snapshot from the simulation. (D) H bonds between NTD and CTD domains are represented as red lines and H bonds between groups of the same domain as grey lines. On the average, at any moment of time during the simulations there are ~140 H bonds, and ~20 between NTD and CTD domains. (E) H bond interactions represented as lines color coded according to the occupancy of the H bond, from red (100% occupancy) to blue. (F) H-bond represented according to the strength of H bonds sampled during the simulation. Strong, medium, vs. weak interactions are depicted as red-green-blue lines in the graph, respectively.

R328 and K499 are close to the RNA fragment in the starting protein crystal structure (Figure 1B). Both of these sidechains are part of persistent local H-bond clusters: R328 with three other protein sidechains (Q325, E329, and R528), and K499, with five other protein sidechains (R500, D503, D522, H520, and R523, Figures 4A-C). The K499 H-bond cluster can further extend, via somewhat more infrequent H-bonding, to S518 and H520 (Figure 4C).

Taken together, the analysis here of the direct H-bonds between protein sidechains suggests that, in VASA, the protein region between the ATP- and RNA-binding interfaces hosts H-bonds and H-bond clusters that together could contribute to conformational couplings.

Transient protein-water H-bond clusters at the nucleotide-binding pocket

Water molecules can mediate interactions between protein groups. In the case of a soluble protein, water-mediated bridges between bulk-exposed side chains can be rather short lived – for example, in the case of the soluble protein PsbO, we found that the lifetime of bridges of 1-5 H-bonded waters between carboxylate pairs is <1ps, and it decays exponentially with the length of the bridge.[38] Here, for simplicity, we consider only short H-bond bridges mediated by one water molecule.

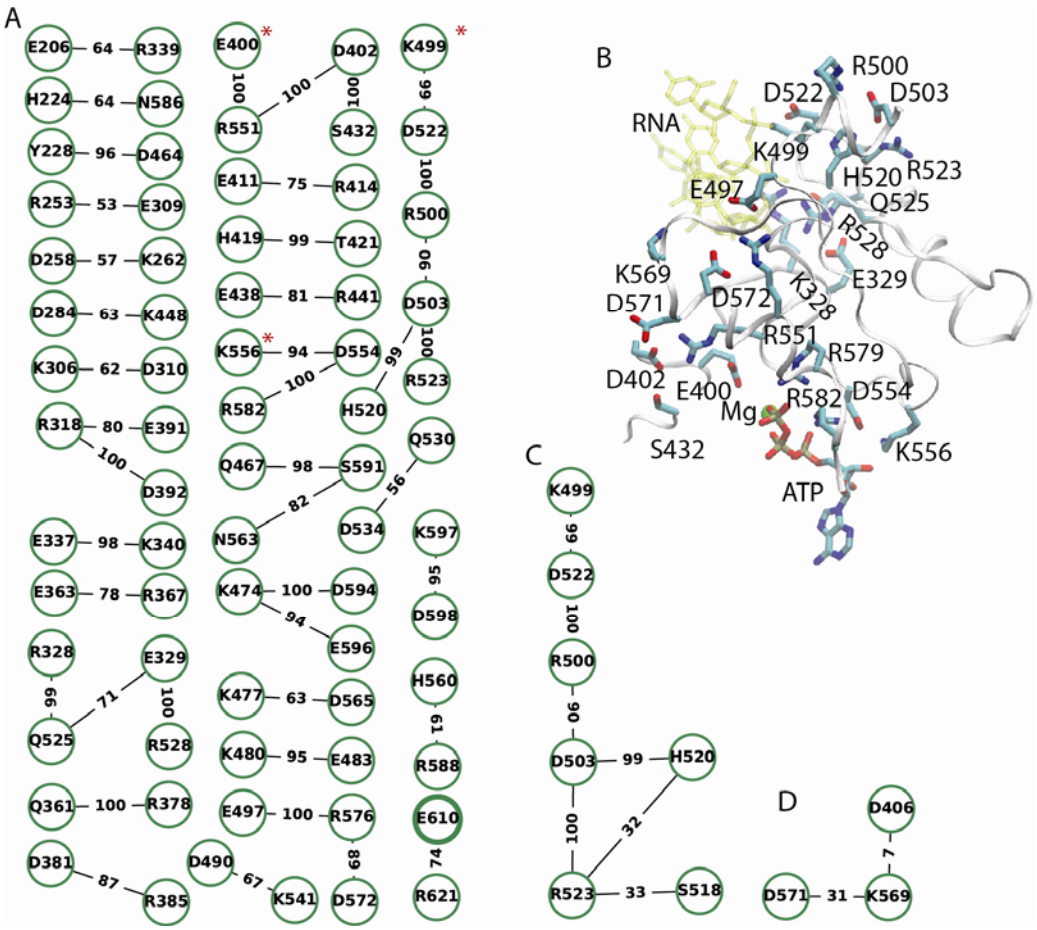


Fig. 4 – H-bonds between sidechains of VASA. (A) H-bond graph computed for protein sidechains. For clarity, we show only H-bonds with occupancies of at least 50% during the trajectory segment used for analyses. The red asterisks indicate local H-bonds/H-bond clusters discussed explicitly in the text. (B) Molecular graphics illustrating selected interactions shown in the H-bond graphs. The RNA molecule is shown as transparent yellow bonds, and the magnesium ion close to the ATP molecule, as a small green sphere; for clarity, H atoms and water molecules are not shown. (C) Local H-bond cluster of K499, with the occupancy criterion lowered to 30%; note that, compared to the K499 cluster from panel A, two additional H-bonds are present with occupancies of 32-33%. (D) Transient H-bonding of D571.

The VASA protein-water network is crowded, as numerous sidechains bridge via H-bonding water (Figure 5A). Most of these interactions are dynamic and thus absent from the graph of H-bonds with occupancies of at least 50% during the trajectory segment used for analyses (Figure 5B).

In general, the one-water bridges and local protein-water H-bond clusters from the ≥50% occupancy graph contain an Asp/Glu or Arg/Lys sidechain (Figure 5B). D399 and E400 of the DEAD-box motif (Figure 1) are part of a local, frequently visited cluster of water-mediated bridges with T296, Q333, and S325, whereas D402 has a somewhat more dynamic bridge with H575 (Figure 5B). Near H575, R576 connects to N563 (Figure 5B).

Adjacent in the sequence to D402 of the DEAD-box, R403 is part of a local H-bond cluster with two carboxylate (D406 and D572) and two hydroxyl groups (S547 and S550). As D572 is close in sequence to H575, it follows that the R403 and DEAD-box H-bond clusters are located close to each other.

To identify transient connections that could connect the nucleotide-binding pocket to other parts of the protein, we started from the H-bond graph computed for one-water bridges between protein sidechains, lowered the H-bond occupancy criterion to 20%, and used the Connected Components analysis tool of Bridge2[20] to examine local H-bond clusters of selected protein groups. We found that the local H-bond cluster of E400 extends by recruiting K295 and R551 (Figure 5C), and the H-bond between D402 and H575 extends to R579 (Figure 5D). The H-bond between R576 and N563 now connects, via relatively infrequent H-bonds with occupancies of 24-30%, to the R403-D406 H-bond cluster, which further extends to include K569, E497, and S568 (Figure 5E).

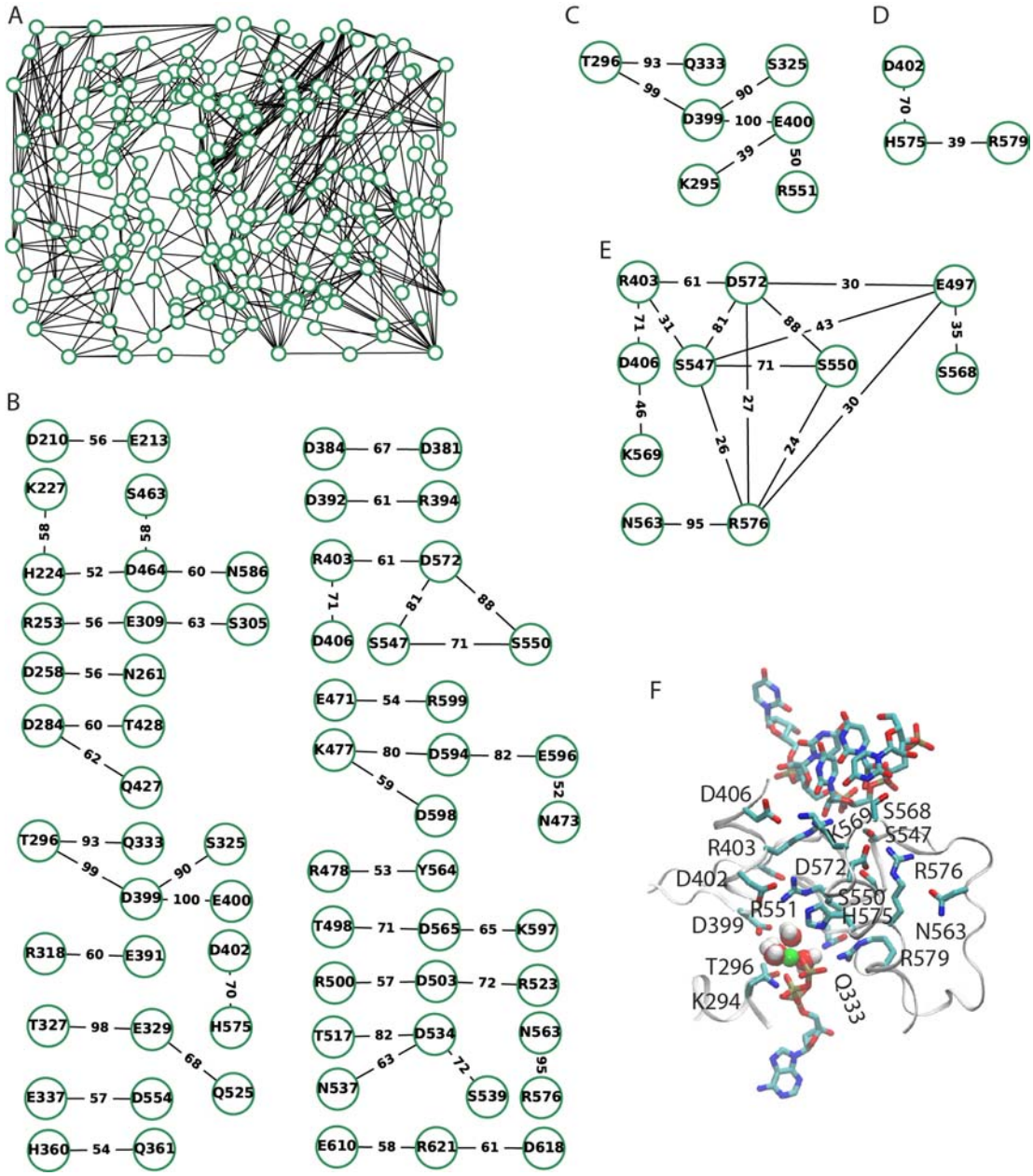


Fig. 5 – Water-mediated bridges between protein sidechains indicate connections between different regions of VASA. (A) Complete H-bond graph computed for one-water H-bond bridges between protein sidechains. The green empty circles are nodes of the graph (H-bonding sidechains), and the lines are one-water-mediated bridges. (B) H-bond graph obtained from the complete graph by selecting only bridges occupied during at least 50% of the trajectory fragment analyzed. Numbers along the edges indicate the occupancies (in integer percentages) of the one-water bridges. (C-E) Local H-bond clusters of E400 (panel C), D402 (panel D), and R576 (panel E) extracted from the graph with H-bond occupancies of at least 20%. (F) Molecular graphics illustrating interactions between protein groups included in the local H-bond networks presented in panels C-E.

CONCLUSIONS

We relied on graph-based analyses to identify dynamic H-bond networks between protein sidechains, and between protein sidechains and water, that could ensure conformational couplings in VASA. We found that the region comprising the DEAD-box hosts persistent H-bond and local H-bonds clusters. Other H-bond networks we identified here include amino acid residues for which site-directed mutagenesis on the VASA construct used to solve the crystal structure[12] suggest a potential functional role. For example, Ala mutation of Q333, which is part of the local protein-water H-bond network of E400 (Figure 5), reduced RNA crosslinking,[12] and Ala mutation of R551, to which the E400 protein-water H-bond network can connect transiently, and which is also

part of the stable protein H-bond cluster of E400 (Figures 4 A, B), reduces RNA unwinding.[12] Both Q333A and R551A are thought to have ATP hydrolysis uncoupled from RNA unwinding.[12] D554 is part of the stable local cluster with K556 and R582 (Figures 4 A, B); D554 is thought to be important for coupling the ATPase activity with RNA unwinding, as its Ala mutation in full-length VASA associated with *Drosophila* oogenesis defects. The H-bond networks identified (Figures 4, 5) here could provide a rationale for the uncoupling in mutant VASA proteins, as mutations might alter the long-distance connections established transiently via dynamic H-bonds and H-bond networks.

Acknowledgements. A-NB wishes to dedicate this work to the memory of Prof. Petre T. Frangopol and his wife, Dr. Maria Frangopol, who shared with her their passion for science. A-NB became interested in biophysics as an undergraduate student in the laboratory Prof. Petre T. Frangopol established at the University "Al. I. Cuza" Iaşi. Research reported here was initiated with financial support from the Marie Curie International Reintegration Grant IRG-276920 (to A-NB) and pursued with financial support in part from the Excellence Initiative of the German Research Foundation via the Freie Universität Berlin, and a computing time allocation from HLRN, The North-German Supercomputing Alliance. We thank Mr. Jens Dreger from the Physics Department of the FU Berlin, and Dr. Christian Tuma from the HLRN, for excellent technical assistance.

REFERENCES

1. P. Linder, P. F. Lasko, M. Ashburner, P. Leroy, P. Nielsen, K. Nishi and P. P. Slonimski, *Nature*, **1989**, *337*, 121-122.
2. F. Rozen, I. Edery, K. Meerovitch, T. E. Dever, W. C. Merrick and N. Sonenberg, *Mol. Biol. Cell*, **1990**, *10*, 1134-1144.
3. H. Hirling, M. Scheffner, T. Restle and H. Stahl, *Nature*, **1989**, *339*, 562-564.
4. M. Causevic, R. G. Hislop, N. M. Kernohan, F. A. Carey, R. A.; Kay, R. J. C. Steele and F. V. Fuller-Pace, *Onocogene* **2001**, *20*, 7734-7743.
5. S. Shin, K. L. Rossow, J. P. Grande and R. Janknecht, *Cancer Res.*, **2007**, *67*, 7572-7578.
6. F. Squeglia, M. Romano, A. Ruggiero, G. Maga and R. Berisio, *Frontiers in Chem.*, **2020**, 8.
7. D. E. Gordon, G. M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K. M. White, M. J. OMeara, R. Hüttenhain, R. M. Kaake, A. L. Richards, B. Tutuncouglu, H. Foussard, J. Batra, K. Haas, M. Modak, M. Kim, P. Haas, B. J. Polacco, H. Branberg, J. M. Fabius, M. Eckhardt, M. Soucheray, M. J. Bennet, M. Cakir, M. J. McGregor, Q. Li, B. Meyer, F. Roesch, T. Vallet, A. Mac Kain, L. Miorin, E. Moreno, Z. Z. C. Naing, Y. Zhou, S. Peng, Y. Shi, Z. Zhang, W. Shen, I. T. Kirby, J. E. Melnyk, J. S. Chorba, K. Lou, S. A. Dai, I. Barrio-Hernandez, D. Memon, C. Hernandez-Armenta, J. Lyu, C. J. P. Mathy, T. Perica, K. B. Pilla, S. J. Ganesan, D. J. Saltzberg, R. Rakesh, X. Liu, S. B. Rosenthal, L. Calviello, S. Venkataraman, J. Liboy-Lugo, Y. Lin, X.-P. Huang, Y. Liu, S. A. Wankowicz, M. Bohn, M. Safari, F. S. Ugur, C. S. Koh, N. Sadat Savar, Q. D. Tran, D. Shengjuler, D. J. Broadhurst, S. Klippsten, P. P. Sharp, N. A. Wenzell, D. K. Kuzuoglu-Ozturk, H.-Y. Wang, R. Trenker, J. M. Young, D. A. Cavero, J. Hiatt, T. L. Roth, U. Rathore, A. Subsramaniam, J. Noack, M. Hubert, R. M. Stroud, A. D. Frankel, O. S. Rosenberg, K. A. Verba, D. A. Agard, M. Ott, M. Emerman, N. Jura, M. von Zastrow, E. Verdin, A. Ashworth, O. Schwartz, C. d'Enfret, S. Mukherjee, M. Javcobson, H. S. Malik, D. G. Fujimori, T. Ideker, C. S. Craik, S. N. Floor, J. S. Fraser, J. D. Grosss, A. Sali, B. L. Roth, D. Ruggero, J. Taunton, T. Kortemme, P. Beltrao, M. Vignuzzi, A. García-Sastre, K. M. Shokat, B. K. Shoichet, N. J. Krogan, *Nature*, **2020**, *583*, 459-468.
8. N. K. Tanner and P. Linder, *Mol. Cell*, **2001**, *8*, 251-262.
9. E. Vrontou and A. Economou, *Biochim. Biophys. Acta*, **2004**, *1694*, 67-80.
10. A.-N. Bondar and J. C. Smith, *Photochem. Photobiol.*, **2017**, *93*, 1336-1344.
11. M. Yajima and G. M. Wessel, *Molecular Reproduction & Development*, **2011**, *78*, 861-867.
12. T. Sengoku, O. Nureki, A. Nakamura, S. Kobayashi and S. Yokoyama, *Cell*, **2006**, *125*, 287-300.
13. J. Xiol, P. Spinelli, M. A. Laussmann, D. Homolka, Z. Yang, E. Cora, Y. Coute, S. Conn, J. Kadlec, R. Sachidanandam, M. Kaksonen, S. Cusack, A. Ephrussi and R. S. Pillai, *Cell*, **2014**, *157*, 1698.
14. S. Nithianantham and B. H. Shilton, *J. Mol. Biol.*, **2008**, *383*, 380-389.
15. A.-N. Bondar, H. Mishima and Y. Okamoto, *Biochim. Biophys. Acta*, **2020**, *1864*, 129654.
16. S. Krishnamurthy, N. Elefteriadis, K. Karathanou, J. H. Smit, A. G. Portaliou, K. E. Chatzi, S. Karamanou, A.-N. Bondar, G. Gouridis and A. Economou, *Structure*, **2021**, *29*, 1-13.
17. W. Humphrey, W. Dalke and K. Schulten, *J. Mol. Graph.*, **1996**, *14*, 33-38.
18. K. Karathanou and A.-N. Bondar, *J. Chem. Information and Modeling*, **2019**, *59*, 1882-1896.
19. M. Siemers, M. Lazaratos, K. Karathanou, F. Guerra, L. S. Brown and A.-N. Bondar, *J. Chem. Theory and Computation*, **2019**, *15*, 6781-6798.
20. M. Siemers and A.-N. Bondar, *J. Chem. Information and Modeling*, **2021**, *6*1, 2998-3014.
21. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, **1983**, *4*, 187-217.
22. N. Foloppe and A. D. MacKerell Jr., *J. Comput. Chem.*, **2000**, *21*, 86-104.
23. A. D. MacKerell Jr. and N. Banavali, *J. Comput. Chem.*, **2000**, *21*, 105-120.
24. A. D. MacKerell Jr., M. Feig and C. L. I. Brooks, *J. Comput. Chem.*, **2004**, *25*, 1400-1415.
25. A. D. MacKerell Jr., D. Bashford, M. Bellot, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. I. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, **1998**, *102*, 3586-3616.

26. B. R. Brooks, C. L. I. Brooks, A. D. MacKerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. D. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. York and M. Karplus, *J. Comput. Chem.*, **2009**, *30*, 1545-1614.
27. E. J. Denning, U. D. Priyakumar, L. Nilsson and A. D. MacKerell Jr., *J. Comput. Chem.*, **2011**, *32*, 1929-1943.
28. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, **1983**, *79*, 926-935.
29. L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan and K. Schulten, *J. Comput. Phys.*, **1999**, *151*, 283-312.
30. J. C. Phillips, B. Braun, W. Wang, J. Gumbart, E. Takjkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, **2005**, *26*, 1781-1802.
31. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, **1977**, *23*, 327-341.
32. G. J. Martyna, D. J. Tobias and M. L. Klein, *J. Chem. Phys.*, **1994**, *101*, 4177-4189.
33. S. E. Feller, Y. Zhang, R. W. Pastor and B. Brooks, *J. Chem. Phys.*, **1995**, *103*, 4613-4621.
34. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, **1993**, *98*, 10089-10092.
35. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577-8593.
36. M. Tuckermann, B. J. Berne and G. J. Martyna, *J. Chem. Phys.*, **1992**, *97*, 1990-2001.
37. K. Karathanou, M. Lazaratos, E. Bertalan, M. Siemers, K. Buzar, G. F. X. Schertler, C. del Val and A.-N. Bondar, *J. Struct. Biol.*, **2020**, *212*, 107617.
38. L. Kemmler, M. Ibrahim, H. Dobbek, A. Zouni and A.-N. Bondar, *Phys. Chem. Chem. Phys.*, **2019**, *21*, 25449-25466.