ATP dependent NS3 helicase interaction with RNA: insights from molecular simulations

Andrea Pérez-Villa, Maria Darvas and Giovanni Bussi[*]

Scuola Internazionale Superiore di Studi Avanzati, International School for Advanced Studies, 265, Via Bonomea, I-34136 Trieste, Italy

ABSTRACT

Non-structural protein 3 (NS3) helicase from hepatitis C virus is an enzyme that unwinds and translocates along nucleic acids with an ATP-dependent mechanism and has a key role in the replication of the viral RNA. An inchworm-like mechanism for translocation has been proposed based on crystal structures and single molecule experiments. We here perform atomistic molecular dynamics in explicit solvent on the microsecond time scale of the available experimental structures. We also construct and simulate putative intermediates for the translocation process, and we perform non-equilibrium targeted simulations to estimate their relative stability. For each of the simulated structures we carefully characterize the available conformational space, the ligand binding pocket, and the RNA binding cleft. The analysis of the hydrogen bond network and of the non-equilibrium trajectories indicates an ATP-dependent stabilization of one of the protein conformers. Additionally, enthalpy calculations suggest that entropic effects might be crucial for the stabilization of the experimentally observed structures.

INTRODUCTION

Non-structural protein 3 (NS3) is a molecular motor encoded by Hepatitis C virus (HCV) consisting of a N-terminal region with a serine protease domain and a C-terminal superfamily 2 (SF2) helicase that unwinds and translocates on nucleic acids. This protein translocates on single-stranded ribonucleic acid (ssRNA) in the $3' \rightarrow 5'$ direction through a periodic stepwise mechanism (1–4). Translocation is ATP dependent, so that this enzyme has been also classified as a DExH box ATPase (5). Understanding the mechanism of action of this molecular motor at the atomistic level is fundamental since the NS3 helicase domain has been proposed as a target for the development of antiviral agents (6).

NS3 helicase has been characterized by single-molecule experiments (7–11) and biochemical essays (12–14). Although sometime acting as a dimer or as an oligomer (15,16), NS3 also functions as a monomer, similarly to other SF1 and SF2 helicases (17,18). The N-terminal protease domain affects the binding of NS3 to RNA and plays an important role for the reaction kinetics (19). However, the protease domain is not essential for the helicase activity (20,21), thus the helicase domain (NS3h) can be characterized in isolation. Interestingly, optical tweezers experiments have provided estimates of the number of substeps per cycle, up to a resolution of single base pair (8,9).

Fluorescence resonance energy transfer (FRET) (7) on the NS3-DNA complex suggested a step of 3 bp with 3 hidden substeps where 1 bp is unwound per 1 ATP molecule consumed following an inchworm mechanism. However, although single molecule experiments allow the kinetics of the mechanism to be captured, they cannot provide detailed structural information. Additionally, the force applied during mechanical manipulation is often much larger than the actual force felt by biopolymers *in vivo* (22,23). On the other hand, X-ray crystallography can provide detailed snapshots at atomistic resolution. Only a few intermediate snapshots have been reported so far related to NS3h translocation on RNA (20,24,25), and conformational differences between these snapshots have been interpreted using elastic network models (26,27). In this context, molecular dynamics simulations (28) with accurate force fields could add dynamical information to the available crystal structures providing a new perspective on the mechanism of action of this important molecular motor.

In this paper, we describe atomistic molecular dynamics (MD) simulations in explicit solvent of NS3-ssRNA complex in the absence (apo) and presence of ATP/ADP. In order to understand the stability of the intermediates along the translocation cycle we constructed putative intermediate structures. We used a recent version of the AMBER force field and performed microsecond time scale simulations so as to provide statistically meaningful results. Results are complemented with non-equilibrium targeted molecular dynamics so as to assess the relative stability of the apo, ADP and ATP structures. Experimen-

[*]To whom correspondence should be addressed. Tel: +39 040 3787 407; Fax: +39 040 3787 528; Email: bussi@sissa.it
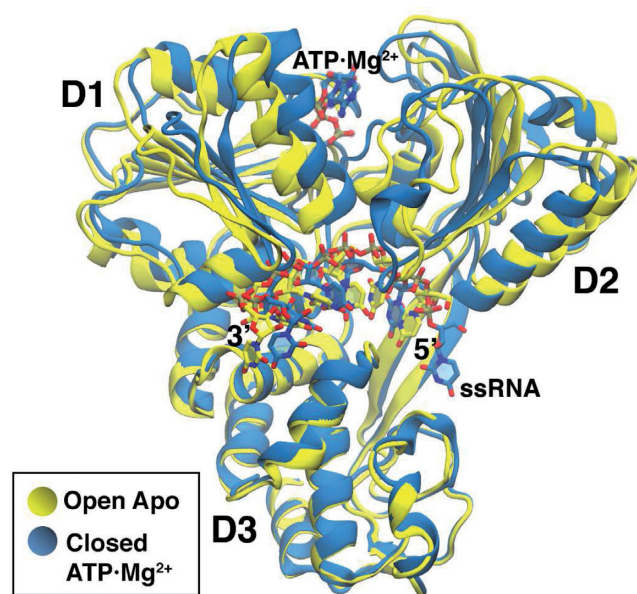
Figure 1. Comparison between open-apo and closed-ATP crystal structures. NS3h is represented with ribbons, while ATP and ssRNA are displayed with sticks. The gap between domain 1 (D1) and domain 2 (D2) is larger for the open conformation. The distance between centers of mass of D1 and D2 is ~29 Å in the open structure and ~26 Å in the closed one. Water and ions are not shown.

tally determined structures are shown to be stable within this timescale. Both the experimental structures and the putative intermediate ones are analyzed in details. Only a handful of MD simulations have been reported on nucleic-acid/helicase complexes so far (27,29), all of them on a much shorter time scale. To the best of our knowledge, only a few MD simulations have been performed on the microsecond timescale for RNA-protein complexes of comparable or larger size (30–32), and thus our results provide a valuable benchmark for state-of-the-art molecular dynamics of these systems.

MATERIALS AND METHODS

Atomistic molecular dynamics simulations in explicit solvent were performed for the NS3h-ssRNA complex for three different molecular systems (apo, with ADP and with ATP) starting from two different conformations (open and closed), for a total of six simulations. In the open form the distance between domains D1 and D2 is larger than in the closed form. Crystal structures are available in the protein data bank (PDB) for the apo open form (PDB: 3O8C) and for the holo (ATP) closed form (PDB:3O8R) (25). The four missing combinations, namely open-ATP, open-ADP, closed-ADP and closed-apo, were constructed as described below. A superimposed representation of open-apo and closed-ATP crystal structures can be seen in Figure 1. The molecular visualizations were generated using VMD (33,34).

Modeling the intermediate states

The missing intermediate states were built based on other experimental structures available from the PDB. The

closed-apo form was constructed by ATP removal from the available closed-ATP structure (PDB: 3O8R) (25). The open-ATP/ADP form was constructed by means of a structural alignment procedure, adding the ATP/ADP to the available open-apo structure (PDB: 3O8C) (25). To properly place the ATP, we used the ATP coordinates from the closed-ATP structure (PDB: 3O8R) after structural alignment of the binding pocket site, Motif I (Walker A, residues 204 to 211). Due to the lack of a crystal structure for the NS3h HCV-ssRNA-ADP complex, other NS3 helicases of flaviviridae virus were analyzed to characterize the ADP binding on the protein. The possible candidates that belong to this family of viruses are Dengue and Yellow Fever proteins. Crystal structures of complex protein-ssRNA-ATP/ADP are available for Dengue virus (PDB: 2JLV and 2JLZ, respectively) (35). Walker A is a well conserved motif across SF2 NS3 helicases, and a structural alignment showed that ATP and ADP placement is virtually identical in PDB: 2JLV and PDB: 2JLZ of Dengue virus. We thus used coordinates of ADP taken from PDB: 3O8R in the closed NS3h, replacing ATP, and in the open NS3h (PDB: 3O8C).

We observe that in reference (27) the open-ATP form was constructed by performing a targeted MD starting from the closed-apo structure. We preferred here to use the structural alignment procedure discussed above to avoid potential artifacts resulting from the non-equilibrium pulling in the generation of the starting points for our long MD simulations.

Molecular dynamics simulations

Each of the starting structures consisting of NS3 peptide (436 aminoacids), polyUracil ssRNA (6 nucleotides), and, when present, ATP/ADP·Mg^{2+}, was solvated in a box containing 31 058 water molecules, 70 Na^+ ions, corresponding to a 0.1 M concentration, and the missing Cl^- ions to neutralize the total charge, for a total of ~100 000 atoms (see Supplementary Table SD1 for details). GROMACS 4.6 program (36) with AMBER99sb*ILDN-parmbsc0-χ_{OL3} + AMBER99ATP/ADP force fields were used. This force field is based on AMBER99 force field (37), and additionally implements corrections to protein backbone (38,39), protein side-chain (40), RNA backbone (41), and glycosidic torsions (42) as well as parameters for ATP and ADP (43) and Mg^{2+} (44). Water molecules were described by the TIP3P model (45). All the molecular dynamics simulations were performed on the isothermal-isobaric ensemble, using the stochastic velocity rescaling thermostat at 300 K (46) and the Berendsen barostat with an isotropic pressure coupling of 1 bar (47). The systems were simulated for 1μs each, with a 2 fs time step, using periodic boundary conditions, the LINCS algorithm (48) to constrain bonds, and the particle-mesh Ewald method (49,50) to account for long-range electrostatics. The resulting setup for the apo simulations is thus similar to the one used in (51). The conformations obtained after 200 ns were extracted and used with randomized velocities as starting points for further control simulations of 200 ns length each. The root-mean-square deviations after structural alignment (RMSD) (52) of the

long MD and of the control simulations were used to verify the stability of the simulated systems.

Monitoring conformational changes

To monitor structural fluctuations we used two order parameters based on a combination of RMSDs from the crystal structures. In reference (27) the difference between squared RMSD from open and closed structure was used to monitor the conformational change. We here used a similar procedure but we considered a subset of atoms from the peptide and the ssRNA that are mainly involved in the conformational change. The full list of atoms is provided in Supplementary Data (Supplementary Figure SD1 and Supplementary Table SD2). We observe that the difference between squared RMSDs is closely related to the progression variable S used in the context of path collective variables (53,54). We thus also introduce a variable Z that measures the distance from a hypothetical transition path obtained as a linear interpolation between the two experimental structures. S and Z are thus defined as $S = \frac{R_o^2 - R_c^2}{2R_{oc}}$ and $Z = \frac{R_o^2 + R_c^2}{2} - S^2 - \frac{R_{oc}^2}{4}$. Here R_o and R_c are the RMSDs from the open (PDB: 3O8C) and closed (PDB: 3O8R) structures respectively, and R_{oc} is the RMSD between the open and the closed structure.

Additionally, we performed a principal components analysis (PCA) using the same set of atoms (55). The PCA was made based on a single trajectory obtained by concatenating all the six simulations. Individual simulations were then projected on the eigenvectors corresponding to the two largest eigenvalues.

Hydrogen-bond analysis

Hydrogen bonds are computed based on distance-angle geometric criteria (56). The cut-off radius for the distance donor-acceptor is of 3.5 Å and the cut-off angle between acceptor-donor-hydrogen ($0°$ is the strongest interaction) is $30°$. The *g_hbond* tool of GROMACS 4.6 program was used for the calculations, and the standard errors were estimated through a binning analysis.

Stacking interactions

In order to monitor stacking interactions between nucleobases and between aromatic aminoacids and nucleobases, we used two different procedures. Intra RNA stacking was annotated by the baRNAba tool (57), which assigns one oriented bead to each nucleotide to describe RNA structural properties. For stacking between aromatic aminoacids and nucleobases we used a geometric criterion where the two residues were considered as stacked if the distance between their center of mass was less than 5 Å and the angle between the two planes was less than $30°$. This latter analysis has been performed for selected interactions that are observed in the crystal structure, namely Y241 with ligand adenine and for W501 with U7. Distances and angles were computed with PLUMED (58).

Electrostatic interactions

Electrostatic interactions between protein-RNA and protein-ligand were monitored by computing the Debye–Hückel interaction energy (59). The calculation was made using PLUMED (58) with the implementation described in ref (60). An ionic strength of 0.1 M was selected that corresponds to a screening length of ≈ 10 Å.

Total enthalpy calculations

Enthalpy is defined in the isothermal-isobaric ensemble as $H = \langle U \rangle + p\langle V \rangle$. Here U is the potential energy, p the external pressure, and V the volume of the simulated box. The enthalpy values were computed here with the *g_energy* tool of GROMACS 4.6 program. For energy differences to be meaningful, simulations corresponding to different conformations (open or closed) of the same system were prepared so as to contain exactly the same number and types of atoms. Since total energy can be affected by numerical details, all simulations were performed using identical settings on identical computers. The first 200 ns of MD were performed using GPUs+CPUs, and the remaining 800 ns using CPUs only. We observed no significant difference between results obtained with GPUs+CPUs and results obtained with CPUs only. We also recall that total energy calculation can be affected by statistical errors mostly due to fluctuations of solvent contributions. We here computed errors using a binning analysis, and run simulations long enough for this error to be <5 kcal/mol.

Targeted molecular dynamics

To estimate the relative stability of the closed and open conformations for the apo/ADP/ATP structures, we performed short targeted MD starting from the closed structures (61) following a protocol similar to the one used in ref. (27). Namely, we applied a time dependent harmonic restraint of stiffness 500 kcal/(mol Å^2) to the RMSD from the corresponding open structures. RMSD was computed using all heavy atoms of protein and RNA. The center of the restraint was moved from a given initial value to zero during a 20 ns long simulation. During the first ns the center was left at its original value, whereas during the following 19 ns the center was moved linearly in time to zero. To avoid any bias, we chose the initial value to be equal to RMSD between the closed structure after equilibration and the open crystal in all the three cases, and we repeated every simulation three times using a different random seed. This resulted in a total of 27 independent simulations, corresponding to 3 different systems (apo/ADP/ATP), 3 different steering protocols (starting from the initial RMSD of apo/ADP/ATP structures) and 3 different random seeds. The relative stability of the closed and open structures was estimated by measuring the work performed during the pulling.

RESULTS

Structural analysis

The linearized path variables S and Z are used here to monitor conformational changes in our MD trajectories. As it
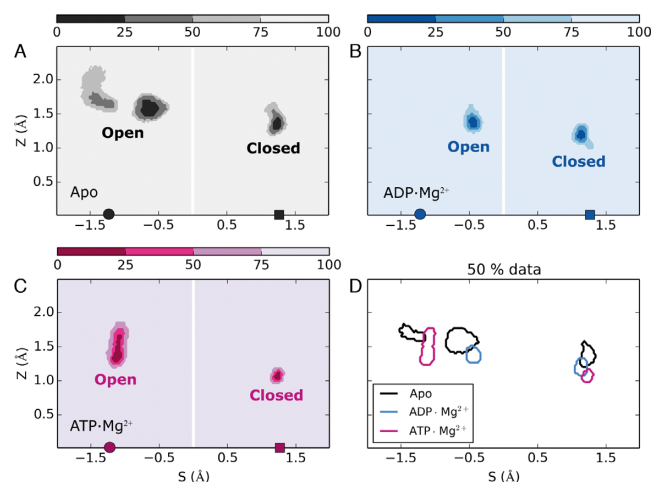
Figure 2. Projection of trajectories on linear path variables for apo (**A**), ADP (**B**) and ATP (**C**) simulations. The contour graphs indicate regions containing the indicated fraction of the analyzed conformations (25, 50 and 75% as indicated). Open and closed regions are well defined and do not overlap between them. Panel (**D**) summarizes the contour lines corresponding to 50% of the data. The location of open and closed reference structures (PDB: 3O8C, 3O8R) is indicated by a circle and a square, respectively.

Table 1. Average number of intra-solute hydrogen bonds for each of the six reported simulations. Detailed counts between selected solute groups are reported in Supplementary Table SD4–SD6

Groups		Total	ΔHB_{oc}
Apo	Closed	349.4	
	Open	348.8	-0.6
	3O8C	320	
ADP·Mg^{2+}	Closed	353.1	
	Open	353.6	0.5
ATP·Mg^{2+}	Closed	362.0	
	Open	348.5	-13.5
	3O8R	329	

can be observed in Figure 2, all the simulations are stable (see also Supplementary Figure SD2 and the time series of RMSD in Supplementary Figure SD3) and remain near to the starting experimental structure ($Z \lesssim 2\text{Å}$) without undergoing significant transitions to the other conformation. The protein/RNA complex is more flexible in the open structure than in the closed one, as it can be appreciated by the larger fluctuations of the Z value. Moreover, one can observe that, in presence of the ligand, fluctuations are partly reduced both in the open and in the closed structure.

The distribution of the S variable indicates that both the apo and the ADP open systems explore regions that are slightly toward the closed reference structure. For the apo system a bimodal distribution is observed (Figure 2) with one peak corresponding to a lower distance from the closed structure. This peak is however sampled only transiently in the first half of the simulation, as it can be seen from Supplementary Figure SD2. In the second half of the trajectory the protein opens again.

We also computed the gap between D1 and D2 as the distance between the centers of mass of D1 and D2 (see Supplementary Table SD3). The interdomain gap is ~28 Å for the open conformations and ~26 Å for the closed ones.

Similar results were obtained by the PCA and can be seen in Supplementary Figure SD4.

Hydrogen Bonds

We computed the number of hydrogen bonds that are formed during the MD simulation by several groups of solute atoms, including the three protein domains, RNA and, when present, ligand. Detailed results are reported in Supplementary Data Supplementary Table SD4–SD6, whereas the total counts are summarized in Table 1.

In general, the values obtained from the simulations starting from the experimental structures are very similar to those obtained from the crystal structures (columns open-apo versus column 3O8C in Supplementary Table SD4; columns closed-ATP versus column 3O8R in Supplementary Table SD6). However, there are a few discrepancies that must be commented. In particular, all the counts are slightly higher in the MD trajectories when compared to the crystal structures. Interestingly, there are a few interactions visible in the MD between domains D1 and D2 for the open structure, which are not present in the crystal structures. This is due to a slight closure of the open structure that is observed during MD and allows contacts to be formed between highly flexible loops. More precisely, transient hydrogen bonds are formed between an acceptor of the Arginine finger (Q460) and donors from the DExH box (E291, H293). Additionally, hydrogen bonds between domain D1 and D3 are observed in the simulated trajectories of both the open and closed structures between residues T305-R512 and S297-E493. These contacts are not present in the reference crystal structure.

We notice that in both cases the donor-acceptor distances and angles in the crystal structures are slightly above the threshold that we used to identify hydrogen bonds and a small fluctuations of these bonds increases the hydrogen count in the MD. This explanation is valid also for the increased number of intradomain D1-D1 and D2-D2 hydrogen bonds.

Our simulations can also provide an insight on possible intermediate structures that have not been crystallized, namely closed-apo/ADP and open-ADP/ATP. We first compare the open and closed structure in the absence of the ligand. These two structures present a very similar number of hydrogen bonds. Although the total number of hydrogen bond is not a rigorous estimator of structural stability, it is expected to give a large contribution to the interaction energy. Thus, this result suggests that the open and the closed structures might have a comparable stability in the absence of the ligand. Notably, the number of D1-D2 hydrogen bonds that are missing in the open structure are compensated by an equivalent number of intradomain D2-D2 bonds (see Supplementary Table SD4). At the same time, protein-RNA bonds are overall maintained during the simulation, and some D2-RNA bonds are replaced with D3-RNA bonds upon opening. Very similar considerations can be made for the ADP case (see Supplementary Table SD5). On the contrary, when ATP is present the

closed structure forms approximately 15 hydrogen bonds more than the open one. This is due to a combination of several effects (see Supplementary Table SD6), including the formation of D1-D2 contacts, the formation of new D2-D2 hydrogen bonds, and the interaction of the ligand with D2. A detailed structural analysis of the binding pocket is presented below. The increased number of hydrogen bonds suggests that ATP stabilizes the closed structure.

RNA binding cleft

The ssRNA interacts with all the three protein domains. Overall, the contacts between domain D1 and RNA are maintained in the closed and open conformations, whereas the contacts between domain D2 and RNA are either shifted by one nucleotide or missing in the closed structure. Most of the contacts are with the RNA backbone, consistently with the fact that the helicase can process RNA irrespectively of its precise sequence. Residues K272, T269 and V232 from domain D1 interact with phosphate oxygens from U8, U7 and U6, respectively, in all the six simulations. In the open structures, T416, T411 and K371 from domain D2 interact with phosphate oxygens from U5, U4 and U4, respectively. In the closed structures, interaction of T416 with RNA is not present and both T411 and K371 interact with phosphate oxygens from U5. R393 sidechain from domain D2 interacts with phosphates from U6 and U5 in all the open structures. Conversely, in the closed-ATP an interaction with phosphates from U7 and U6 is formed after ~50 ns of simulation. This interaction is missing in the ADP and apo simulations (see Supplementary Figure SD5). Something similar happens to residue N556, which interacts with a nucleobase in all the open structures and in the closed-ATP one. Finally, we notice that residues W501 (see Figure 3) and V432 (see Supplementary Figure SD6) act as gates located at the 3′ terminus and 5′, respectively, of the simulated oligonucleotide.

RNA stacking interactions

In the closed conformation the RNA strand shows a larger bending so that some of the stacking interactions between the nucleobases are lost. We recall that the 5′ terminal base (U3) is lagging outside of the protein in the closed structure, and thus is not interacting with the neighboring base U4. This stacking is never formed during the MD of the closed structure. Stacking interactions between U4 and U5 and between U5 and U6 are initially formed in the open crystal structure and not formed in the closed crystal structure. During MD, they appear to be transiently formed also in the latter case, although with a lower frequency when compared with the open structures. Stacking between U6 and U7 appears to be stable in all conditions, since the interaction of these nucleotides with D1 and D3 domains is very similar in the open and closed structures. Finally, it is important to notice that stacking between the 3′ terminal bases U7 and U8 is never formed. This is due to the stable stacking interaction between U7 and W501 side chain.
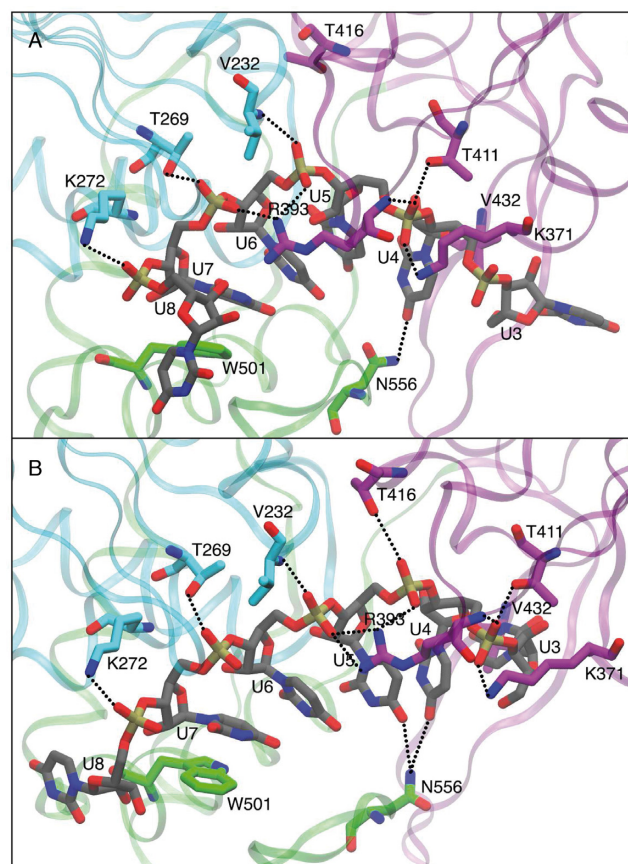

Figure 3. Representation of the RNA binding cleft. Snapshots depicted correspond to the RMSD centroids of trajectories as obtained by the clustering algorithm discussed in ref. (62), using a cutoff radius of 1.25 Å . RNA chain and aminoacids interacting with RNA are highlighted in sticks representation. Residues from D1, D2 and D3 are depicted in cyan, purple and green, respectively, RNA chain is shown in gray color. Snapshots are shown for closed-ATP (**A**) and open-ATP (**B**) simulations. Equivalent graphs for the ADP and apo simulations are shown in Supplementary Figure SD5.

ATP/ADP binding pocket

As expected, the interactions in the binding pocket are heavily affected by the protein conformation (open versus closed) and by the presence and the kind of the ligand (apo versus ADP versus ATP). As it can be appreciated in Figure 4A, the closed-ATP structure after 1μs is still very similar to the crystal structure. More precisely, interactions between ligand, Walker A (residues 204–211), DExH box (residues 290–293), and arginine finger (residues 460–467) are preserved in the MD simulation. The observed stability of this conformation is consistent with the fact that it is among those that have been crystallized (25). On the contrary, the open-ATP structure experiences a reorganization of the binding pocket after 2 ns. In particular, a bending of the ATP is observed which leads to the coordination of the Mg^{2+} cation to α, β and γ phosphates, which corresponds to an alternative metastable conformation of the ATP·Mg^{2+} complex (63). Moreover, a contact between Mg^{2+} and D290 (DExH box) is formed and the contact between Mg^{2+} and S211 is now mediated by a water molecule.
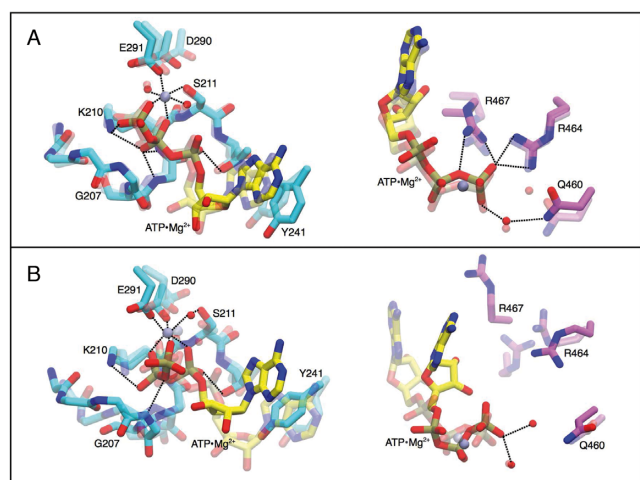
Figure 4. Snapshots of the ATP/ADP binding site. Aminoacids from D1(cyan) and D2 (purple) and ligand (yellow) are shown as sticks, the magnesium ion is shown as a lilac sphere. Closed-ATP (**A**) and open-ATP (**B**) systems are shown. Transparent representation indicates the initial structure, solid color indicates the structure after 1 μs of molecular dynamics. Coordination with Mg^{2+} cation and hydrogen bond interactions are shown with the black dashed line. Equivalent representations for ADP are shown in Supplementary Figure SD7.

When ADP replaces ATP (see Supplementary Figure SD7), both closed and open conformations present fluctuations and changes in the interactions between ligand, residues and water molecules. Notably, in the open conformation the Mg^{2+} ion forms initially a pentacoordinated species which after a few ns is replaced by a hexacoordinated species coordinated with three water molecules, the α and β phosphates and S211. This latter arrangement is stable for the entire simulation. We observe that pentacoordinated Mg^{2+} is not expected in this system (64) and that the force field employed for Mg^{2+} (44) can recover the correct hexacoordinated species.

We also notice that whereas in the open-ATP case the ligand exhibits a poor interaction with domain D2, in the open-ADP case there are several interactions appearing due to a rearrangement of R467 that interacts with phosphate β. This result is also supported by the hydrogen bond analysis in Table 1. Finally, we observe that in closed-ADP and open-ATP, the stacking between the adenine from the ligand and Y241 is formed for a small fraction of the time. This suggests that stability of the ligand is enhanced in closed-ATP structure relatively to the other cases.

Electrostatic interactions

Debye–Hückel interaction energies (G^{DH}) were computed discarding the initial 200 ns of simulation. These numbers cannot provide quantitative estimates of the binding affinities but can provide a qualitative ranking for the different protein conformations.

We computed the interaction between the peptide and the RNA so as to have a proxy for the RNA binding energy (see Table 2). The RNA-protein interaction is slightly larger (∼1 kcal/mol) in the closed conformations when compared with the open ones, irrespectively of the presence of the lig-

Table 2. Electrostatic interaction computed as Debye–Hückel energies (G^{DH}). Protein* denotes Protein-RNA complex. Errors were computed from binning analysis (bin width: 80ns). Errors lower that 0.05 kcal/mol are shown as "0.0"

	ΔG^{DH} (kcal/mol)	
RNA-protein	Closed	Open
Apo	− 4.9 ± 0.0	− 3.7 ± 0.2
ADP·Mg^{2+}	− 5.0 ± 0.1	− 3.9 ± 0.0
ATP·Mg^{2+}	− 5.5 ± 0.1	− 4.7 ± 0.1
RNA-Ligand	Closed	Open
ADP·Mg^{2+}	− 0.1 ± 0.1	− 0.2 ± 0.0
ATP·Mg^{2+}	− 0.2 ± 0.1	− 0.2 ± 0.1
Ligand-Protein*	Closed	Open
ADP·Mg^{2+}	− 0.4 ± 0.0	− 1.1 ± 0.0
ATP·Mg^{2+}	− 0.8 ± 0.0	− 0.1 ± 0.0

and. However, we notice that in the presence of ATP the RNA-protein interaction is stronger. By considering separately the direct RNA-ligand interaction (also shown in Table 2) it can be appreciated that this change is due to a rearrangement of the protein and not to a direct RNA-ligand electrostatic coupling. The contribution of the individual aminoacids located at distance <6 Å from the RNA is presented in Supplementary Table SD7 where it can be appreciated that in the closed conformations the most interacting residue is K371 whereas for the open structures the most interacting residue is R393. An exception is given by the closed-ATP where also the R393 is the most interacting residue. This difference is consistent with the rearrangement of the RNA binding cleft discussed above.

We also computed the interaction between the ligand and the RNA-protein complex, which is also reported in Table 2. Remarkably, ATP interacts more with the protein in its closed conformation, whereas ADP interacts more with the protein in its open conformation. Though being far from quantitative, this result is qualitatively consistent with the fact that the affinity for ATP is larger in the closed form.

Enthalpies

Average enthalpies were computed discarding the initial 200 ns of each trajectory. This initial part is thus considered here as equilibration. We only computed enthalpy differences between simulations including exactly the same number and types of atoms, corresponding to two alternative conformations (open versus closed). From the results reported in Table 3, it is observed that open conformation has a systematically lower enthalpy than closed ones. This difference is more marked when ATP is present. We observe that statistical errors, as obtained from binning analysis (bin width: 80 ns), are very low for all the cases. Estimation of the enthalpy values for the first and second halves of the simulations are also reported in Supplementary Table SD8 and the values are consistent with the reported ones from the entire trajectory.

Enthalpy values are important to describe internal interactions (65). Differences of enthalpy thus provide an indication of relative stability of different conformations, at least

Table 3. Enthalpy differences between open and closed conformations. The values are computed between systems with same number and kind of molecules

Ligand	ΔH_{oc} (kcal/mol)
Apo	-9.7 ± 2.8
ADP·Mg^{2+}	-25.3 ± 3.1
ATP·Mg^{2+}	-54.1 ± 4.1

when their contribution is dominant in the free-energy difference. In the present context, enthalpy calculations suggest that the ATP stabilizes the open structure, in contrast with the observed number of intra-solute hydrogen bonds, the structural analysis of the binding pocket, and the electrostatic interactions estimated within the Debye–Hückel model.

Targeted molecular dynamics

The work performed during the targeted molecular dynamics was used as a proxy for the relative stability of the closed and open structure in the presence and the absence of the ligand. One should consider here that the average work only provides an upper estimate of the free-energy change, and that in principle simulations should be combined using the Jarzynski's equality (66). However, a very large number of simulations would be required to provide a converged estimate of the free-energy change for this complex rearrangement. Additionally, since in the targeted MD an increasingly strong restraint is applied to a large number of atoms, the absolute value of the free-energy change would be largely affected by the decrease in the entropy of the restrained atoms. In spite of these quantitative limitations, the average work can be used as a qualitative tool to rank the free-energy differences for apo, ADP and ATP structures. The results for the different steering protocols are shown in Supplementary Figure SD8. Here, it can be appreciated that the work required to open the apo form is systematically lower than the one required to open the ADP and ATP forms. This result is compatible with the structural analysis of the binding pocket, the hydrogen bonds between solute molecules, and the electrostatic interactions discussed above, indicating that the enthalpy is compensated by a large entropic change. The ligand-induced stabilization of the closed form is also compatible with the fact that the ATP form was crystallized in this conformation.

DISCUSSION

We here presented a detailed analysis of conformational properties of NS3 helicase from HCV in complex with ssRNA, with ATP/ADP, and in the apo form. For each of these three systems, we performed MD simulations starting from both the open and the closed conformation, for a total of six simulations of 1 μs length each. For two of these simulations the starting structure was already available as obtained from X-ray cristallography, whereas the remaining four systems were built by structural alignment. Although this protein/RNA complex has only been crystallized either in the apo form or with ATP, we also analyzed the effect of ATP replacement by ADP. An intermediate structure with

ADP has been reported for NS3 from Dengue virus (35), but has never been characterized for the HCV NS3.

The ATP hydrolysis reaction was not explicitly analyzed here, but has been addressed in the literature for the PcrA helicase (67) and for other motor proteins (see, e.g. (68–70)). The present work is not aimed at giving a mechanistic understanding of hydrolysis itself, but rather at providing an insight on the system's behavior after the hydrolysis reaction has occurred.

Overall, all the observed trajectories were stable on this timescale and did not experience any significative structural change. This provides a validation for the protocol employed to build the four structures not obtained from X-ray. In particular, structural stability was assessed by monitoring combinations of RMSD from the open and closed PDB structures, where it can be appreciated that all systems remained near to the experimental structure on this timescale. Interestingly, the ATP and ADP open structures have a lower enthalpy than the corresponding closed structures, indicating that they have been properly equilibrated. Additionally, we observed that in the presence of ATP the closed structure fluctuated less compared with closed-apo/ADP and its open analogue, suggesting that the closed conformation is stabilized by the presence of ATP. This is consistent with the fact that the crystallized structure with ATP is in its closed conformation (20,25). On the other hand, the simulations from the open structure displayed larger fluctuations. This holds also for the open-apo structure, which has been crystallized (20,25). It is possible that crystal contacts or interactions with the protease domain stabilize the open-apo structure in the experiments.

When ATP is present, the total number of hydrogen bonds formed by the solute is significantly larger in the closed structure when compared with the open one. This is also consistent with the fact that ATP stabilizes the closed conformation. Interestingly, for both the apo and the ADP complex this is not observed, and the number of hydrogen bonds is equivalent in the open and closed structures. This indicates that the free-energy difference ΔG between the open and the closed structures should be larger in the ATP case, as confirmed also by the targeted MD simulations. This ΔG is related to the differential affinity of the ligand in the two structures. Indeed, $\Delta G_{ATP} - \Delta G_{Apo} = -k_B T \log K_d^{(closed)} + k_B T \log K_d^{(open)}$, where k_B is the Boltzmann constant, T is the temperature, and $K_d^{(open)}$ and $K_d^{(closed)}$ are the ATP dissociation constants in the open and closed conformation, respectively. Thus, our result implies that the ATP affinity is larger in the closed form. The difference in the protein-ADP/ATP interaction estimated by a Debye–Hückel model is also consistent with this picture.

Surprisingly, enthalpy calculations has an opposite trend, suggesting that open-ATP is dramatically more stable than closed-ATP. We observe that this difference (≈ 54 kcal/mol) is much larger than its statistical error (≈ 4 kcal/mol) so that the reported MD simulations are sufficiently converged to estimate this trend. We recall that, although enthalpy provides a very important contribution to free-energy changes, it might be not sufficient to evaluate the relative stability of two conformers. Indeed, entropic contributions have been

reported to cancel enthalpic ones in several cases (71). From our simulations, it is not possible to properly estimate the entropic difference which should be ascribed not only to the changes in solute flexibility but also to non-trivial changes in water entropy which are connected with, e.g. hydrophobic effects in the cavity between D1 and D2. Although it is in principle possible to use enhanced sampling methods to directly compute free-energy differences (72,73), this is a complex procedure for concerted conformational changes and is left as a subject of further investigation. From the presented results, it can be inferred that a strong entropic compensation stabilizes the experimentally observed closed-ATP structure.

We also analyzed carefully the pattern of RNA-protein interactions. In our simulations the protein does not interact with sugars in the RNA backbone. This is consistent with the empirical observation that this helicase can also process DNA (7,20). Contacts between T269 and T411 with the phosphate oxygens are different according to the protein conformation and act like hooks with the RNA chain during the translocation mechanism. This feature has been reported for other SF2 RNA-binding helicases such as Vasa drosophila and eIF4A, suggesting a possible common mechanism among several helicases of this superfamily (7). Besides confirming the contacts that were already seen in the experimental structures, our simulations provide an insight on the possible conformations available for the open-ATP and closed-apo structures. Indeed, in the ATP complex we observe different contacts with residues R393 and N556, which should be likely ascribed to an allosteric change induced by the ligand.

We notice that the contact network of the ligand binding pocket is affected to the largest extent by the addition of ATP/ADP or by opening and closing. The closed-ATP complex preserves the contacts between the Motifs I, VI and the ATP observed in the crystal structures. On the contrary, the open-ATP complex presents a rearrangement due to the missing contacts with D2. The above effect is probably an underlying reason for the ATP-dependent stabilization of the closed form. Upon the replacement of the ATP by ADP, which mimics the result of the hydrolysis reaction, more interactions are observed between the ligand and D2. This suggests that the stabilizing effect of ADP on the open form is significantly larger than that of ATP. In other words, after hydrolysis has occurred the closed conformation is destabilized with respect to the open one and a conformational transition toward the open structure could be favored.

As a final remark, we observe that performing molecular dynamics of RNA/protein complexes is a non-trivial task, and only a few studies have been published so far (32). In this work, we tested molecular dynamics on a biologically relevant complex including a medium sized protein, a short ssRNA, and a nucleotide ligand using recent force fields on the µs time scale, which is the state of the art for classical molecular dynamics. The total time including all the systems and all the control simulations is approximately 9 µs. This simulation time allowed us to compute averages with relatively small statistical errors and to perform several consistency checks indicating that our simulations provide a converged conformational ensemble around the equilibrium structures. The experimental snapshots are stable within the investigated time scale, indicating that these force fields might be able to properly characterize such heterogenous complexes. Although enhanced sampling techniques should be used to quantify the relative stability of the simulated conformations, this work provides an extensive characterization of the possible intermediates sampled during NS3 translocation on RNA at atomistic detail.

REFERENCES

1. Pang,P.S., Jankowsky,E., Planet,P.J. and Pyle,A.M. (2002) The hepatitis C viral NS3 protein is a processive DNA helicase with cofactor enhanced RNA unwinding. *EMBO J.*, **21**, 1168–1176.
2. Pyle,A.M. (2008) Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu. Rev. Biophys.*, **37**, 317–336.
3. Ding,S.C. and Pyle,A.M. (2012) Chapter six - molecular mechanics of RNA translocases. *Methods Enzymol.*, **511**, 131–147.
4. Colizzi,F. and Bussi,G. (2012) RNA unwinding from reweighted pulling simulations. *J. Am. Chem. Soc*, **134**, 5173–5179.
5. Schwer,B. (2001) A new twist on RNA helicases: DExH/D box proteins as RNPases. *Nat. Struct. Mol. Biol.*, **8**, 113–116.
6. Raney,K.D., Sharma,S.D., Moustafa,I.M. and Cameron,C.E. (2010) Hepatitis C virus non-structural protein 3 (HCV NS3): A multifunctional antiviral target. *J. Biol. Chem.*, **285**, 22725–22731.
7. Myong,S. and Ha,T. (2010) Stepwise translocation of nucleic acid motors. *Curr. Opin. Struct. Biol.*, **20**, 121–127.
8. Arunajadai,S.G. (2009) RNA unwinding by NS3 helicase: A statistical approach. *PLOS ONE*, **4**, e6937.
9. Cheng,W., Arunajadai,S.G., Moffitt,J.R., Tinoco,I. and Bustamante,C. (2011) Single–base pair unwinding and asynchronous RNA release by the hepatitis C virus NS3 helicase. *Science*, **333**, 1746–1749.
10. Dumont,S., Cheng,W., Serebrov,V., Beran,R.K., Tinoco,I., Pyle,A.M. and Bustamante,C. (2006) RNA translocation and unwinding mechanism of HCV NS3 helicase and its coordination by ATP. *Nature*, **439**, 105–108.
11. Cheng,W., Dumont,S., Tinoco,I. and Bustamante,C. (2007) NS3 helicase actively separates RNA strands and senses sequence barriers ahead of the opening fork. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 13954–13959.
12. Krawczyk,M., Stankiewicz-Drogoń,A., Haenni,A.-L. and Boguszewska-Chachulska,A. (2010) Fluorometric assay of hepatitis C virus NS3 helicase activity. *Methods Mol. Biol.*, **587**, 211–221.
13. Mukherjee,S., Hanson,A.M., Shadrick,W.R., Ndjomou,J., Sweeney,N.L., Hernandez,J.J., Bartczak,D., Li,K., Frankowski,K.J., Heck,J.A. *et al.* (2012) Identification and analysis of hepatitis C virus NS3 helicase inhibitors using nucleic acid binding assays. *Nucleic Acids Res.*, **40**, 8607–8621.

14. Ventura,G.T., Costa,E. C. B.d., Capaccia,A.M. and Mohana-Borges,R. (2014) pH-dependent conformational changes in the HCV NS3 protein modulate its ATPase and helicase activities. *PLOS ONE*, **9**, e115941.

15. Jennings,T.A., Mackintosh,S.G., Harrison,M.K., Sikora,D., Sikora,B., Dave,B., Tackett,A.J., Cameron,C.E. and Raney,K.D. (2009) NS3 helicase from the hepatitis C virus can function as a monomer or oligomer depending on enzyme and substrate concentrations. *J. Biol. Chem.*, **284**, 4806–4814.

16. Sikora,B., Chen,Y., Lichti,C.F., Harrison,M.K., Jennings,T.A., Tang,Y., Tackett,A.J., Jordan,J.B., Sakon,J., Cameron,C.E. *et al.* (2008) Hepatitis C virus NS3 helicase forms oligomeric structures that exhibit optimal DNA unwinding activity in vitro. *J. Biol. Chem.*, **283**, 11516–11525.

17. Singleton,M.R., Dillingham,M.S. and Wigley,D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.*, **76**, 23–50.

18. Jankowsky,E. (2011) RNA helicases at work: binding and rearranging. *Trends Biochem. Sci.*, **36**, 19–29.

19. Rajagopal,V., Gurjar,M., Levin,M.K. and Patel,S.S. (2010) The protease domain increases the translocation stepping efficiency of the hepatitis C virus NS3-4A helicase. *J. Biol. Chem.*, **285**, 17821–17832.

20. Gu,M. and Rice,C.M. (2010) Three conformational snapshots of the hepatitis C virus NS3 helicase reveal a ratchet translocation mechanism. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 521–528.

21. Beran,R. K.F., Serebrov,V. and Pyle,A.M. (2007) The serine protease domain of hepatitis C viral NS3 activates RNA helicase activity by promoting the binding of RNA substrate. *J. Biol. Chem.*, **282**, 34913–34920.

22. Liphardt,J., Dumont,S., Smith,S.B., Tinoco,I. and Bustamante,C. (2002) Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science*, **296**, 1832–1835.

23. Brenner,M.D., Zhou,R. and Ha,T. (2011) Forcing a connection: Impacts of single-molecule force spectroscopy on in vivo tension sensing. *Biopolymers*, **95**, 332–344.

24. Yao,N., Reichert,P., Taremi,S.S., Prosise,W.W. and Weber,P.C. (1999) Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease–helicase. *Structure*, **7**, 1353–1363.

25. Appleby,T.C., Anderson,R., Fedorova,O., Pyle,A.M., Wang,R., Liu,X., Brendza,K.M. and Somoza,J.R. (2011) Visualizing ATP-dependent RNA translocation by the NS3 helicase from HCV. *J. Mol. Biol.*, **405**, 1139–1153.

26. Flechsig,H. and Mikhailov,A.S. (2010) Tracing entire operation cycles of molecular motor hepatitis C virus helicase in structurally resolved dynamical simulations. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 20875–20880.

27. Zheng,W. and Tekpinar,M. (2012) Structure-based simulations of the translocation mechanism of the hepatitis C virus NS3 helicase along single-stranded nucleic acid. *Biophys. J.*, **103**, 1343–1353.

28. Tuckerman,M. (2010) *Statistical Mechanics: Theory and Molecular Simulation*, 1st edn, OUP, Oxford.

29. Yu,J., Ha,T. and Schulten,K. (2007) How directional translocation is regulated in a DNA helicase motor. *Biophys. J.*, **93**, 3783–3797.

30. Bock,L.V., Blau,C., Schröder,G.F., Davydov,I.I., Fischer,N., Stark,H., Rodnina,M.V., Vaiana,A.C. and Grubmüller,H. (2013) Energy barriers and driving forces in tRNA translocation through the ribosome. *Nat. Struct. Mol. Biol.*, **20**, 1390–1396.

31. Whitford,P.C., Blanchard,S.C., Cate,J.H. and Sanbonmatsu,K.Y. (2013) Connecting the kinetics and energy landscape of tRNA translocation on the ribosome. *PLOS Comput. Biol.*, **9**, e1003003.

32. Krepl,M., Havrila,M., Stadlbauer,P., Banáš,P., Otyepka,M., Pasulka,J., Štefl,R. and Sponer,J. (2015) Can we execute stable microsecond-scale atomistic simulations of protein-RNA complexes?. *J. Chem. Theory Comput.*, **11**, 1220–1243.

33. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD – Visual Molecular Dynamics. *J. Mol. Graph.*, **14**, 33–38.

34. Stone,J. (1998) An Efficient Library for Parallel Ray Tracing and Animation. Master's Thesis, Computer Science Department, University of Missouri-Rolla.

35. Luo,D., Xu,T., Hunke,C., Grüber,G., Vasudevan,S.G. and Lescar,J. (2008) Crystal structure of the NS3 protease-helicase from dengue virus. *J. Virol.*, **82**, 173–183.

36. Hess,B., Kutzner,C., van derSpoel,D. and Lindahl,E. (2008) Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, **4**, 435–447.

37. Wang,J., Cieplak,P. and Kollman,P.A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?. *J. Comput. Chem.*, **21**, 1049–1074.

38. Hornak,V., Abel,R., Okur,A., Strockbine,B., Roitberg,A. and Simmerling,C. (2006) Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.*, **65**, 712–725.

39. Best,R.B. and Hummer,G. (2009) Optimized molecular dynamics force fields applied to the helix coil transition of polypeptides. *J. Phys. Chem. B*, **113**, 9004–9015.

40. Lindorff-Larsen,K., Piana,S., Palmo,K., Maragakis,P., Klepeis,J.L., Dror,R.O. and Shaw,D.E. (2010) Improved side-chain torsion potentials for the AMBER ff99SB protein force field. *Proteins: Struct., Funct., Bioinf.*, **78**, 1950–1958.

41. Pérez,A., Marchán,I., Svozil,D., Sponer,J., Cheatham III,T., Laughton,C. and Orozco,M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of [alpha]/[gamma] conformers. *Biophys. J.*, **92**, 3817–3829.

42. Zgarbová,M., Otyepka,M., Šponer,J., Mládek,A., Banáš,P., Cheatham,T.E. and Jurečka,P. (2011) Refinement of the cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.*, **7**, 2886–2902.

43. Meagher,K.L., Redman,L.T. and Carlson,H.A. (2003) Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.*, **24**, 1016–1025.

44. Allnér,O., Nilsson,L. and Villa,A. (2012) Magnesium ion–water coordination and exchange in biomolecular simulations. *J. Chem. Theory Comput.*, **8**, 1493–1502.

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

46. Bussi,G., Donadio,D. and Parrinello,M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.*, **126**, 014101.

47. Berendsen,H. J.C., Postma,J. P.M., vanGunsteren,W.F., DiNola,A. and Haak,J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**, 3684–3690.

48. Hess,B., Bekker,H., Berendsen,H. J.C. and Fraaije,J. G. E.M. (1997) LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.*, **18**, 1463–1472.

49. Darden,T., York,D. and Pedersen,L. (1993) Particle mesh Ewald: An Ncdotlog(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.

50. Essmann,U., Perera,L., Berkowitz,M.L., Darden,T., Lee,H. and Pedersen,L.G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, **103**, 8577–8593.

51. Ferrarotti,M.J., Bottaro,S., Pérez-Villa,A. and Bussi,G. (2015) Accurate multiple time step in biased molecular simulations. *J. Chem. Theory Comput.*, **11**, 139–146.

52. Kabsch,W. (1976) A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. A*, **32**, 922–923.

53. Branduardi,D., Gervasio,F.L. and Parrinello,M. (2007) From A to B in free energy space. *J. Chem. Phys.*, **126**, 054103.

54. Díaz Leines,G. and Ensing,B. (2012) Path finding on high-dimensional free energy landscapes. *Phys. Rev. Lett.*, **109**, 020601.

55. Amadei,A., Linssen,A. B.M. and Berendsen,H. J.C. (1993) Essential dynamics of proteins. *Proteins: Struct., Funct., Bioinf.*, **17**, 412–425.

56. Luzar,A. and Chandler,D. (1996) Hydrogen-bond kinetics in liquid water. *Nature*, **379**, 55–57.

57. Bottaro,S., Di Palma,F. and Bussi,G. (2014) The role of nucleobase interactions in RNA structure and dynamics. *Nucleic Acids Res.*, **42**, 13306–13314.

58. Tribello,G.A., Bonomi,M., Branduardi,D., Camilloni,C. and Bussi,G. (2014) PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.*, **185**, 604–613.

59. Leach,A. (2001) *Molecular Modelling: Principles and Applications*, 2nd edn, Prentice-Hall, NY, pp. 604–605.

60. Do,T.N., Carloni,P., Varani,G. and Bussi,G. (2013) RNA/peptide binding driven by electrostatics—insight from bidirectional pulling simulations. *J. Chem. Theory Comput.*, **9**, 1720–1730.

61. Schlitter,J., Engels,M., Krüger,P., Jacoby,E. and Wollmer,A. (1993) Targeted molecular dynamics simulation of conformational change-application to the T ↔ R transition in insulin. *Mol. Simul.*, **10**, 291–308.

62. Daura,X., Gademann,K., Jaun,B., Seebach,D., vanGunsteren,W.F. and Mark,A.E. (1999) Peptide folding: When simulation meets experiment. *Angew. Chem. Int. Ed.*, **38**, 236–240.

63. Branduardi,D., Marinelli,F. and Faraldo-Gómez,J.D. (2015) Atomic-resolution dissection of the energetics and mechanism of isomerization of hydrated ATP-Mg2+ through the soma string method. *J. Comput. Chem.*, doi:10.1002/jcc.23991.

64. Hsin,K., Sheng,Y., Harding,M.M., Taylor,P. and Walkinshaw,M.D. (2008) MESPEUS: a database of the geometry of metal sites in proteins. *J. Appl. Crystallogr.*, **41**, 963–968.

65. Hilser,V.J., Gómez,J. and Freire,E. (1996) The enthalpy change in protein folding and binding: Refinement of parameters for structure-based calculations. *Proteins: Struct., Funct., Bioinf.*, **26**, 123–133.

66. Jarzynski,C. (1997) Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, **78**, 2690–2693.

67. Dittrich,M. and Schulten,K. (2006) Pcra helicase, a prototype ATP-driven molecular motor. *Structure*, **14**, 1345–1353.

68. Zhou,Y., Ojeda-May,P. and Pu,J. (2013) H-loop histidine catalyzes ATP hydrolysis in the e. coli abc-transporter hlyb. *Phys. Chem. Chem. Phys.*, **15**, 15811–15815.

69. Hayashi,S., Ueno,H., Shaikh,A.R., Umemura,M., Kamiya,M., Ito,Y., Ikeguchi,M., Komoriya,Y., Iino,R. and Noji,H. (2012) Molecular mechanism of ATP hydrolysis in f1-ATPase revealed by molecular simulations and single-molecule observations. *J. Am. Chem. Soc.*, **134**, 8447–8454.

70. McGrath,M.J., Kuo,I.-F.W., Hayashi,S. and Takada,S. (2013) Adenosine triphosphate hydrolysis mechanism in kinesin studied by combined quantum-mechanical/molecular-mechanical metadynamics simulations. *J. Am. Chem. Soc.*, **135**, 8908–8919.

71. Chodera,J.D. and Mobley,D.L. (2013) Entropy-enthalpy compensation: Role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.*, **42**, 121–142.

72. Dellago,C. and Hummer,G. (2014) Computing equilibrium free energies using non-equilibrium molecular dynamics. *Entropy*, **16**, 41–61.

73. Abrams,C. and Bussi,G. (2014) Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration. *Entropy*, **16**, 163–199.