Sampling protein motion and solvent effect during ligand binding

Vittorio Limongelli[a,b,c,1], Luciana Marinelli[c], Sandro Cosconati[d], Concettina La Motta[e], Stefania Sartini[e], Laura Mugnaini[e], Federico Da Settimo[e], Ettore Novellino[c], and Michele Parrinello[a,b,1]

[a]Department of Chemistry and Applied Biosciences, Computational Science, Eidgenössiche Technische Hochschule (ETH), Zürich, Switzerland; [b]Institute of Computational Science (ICS), Università della Svizzera Italiana, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland; [c]Dipartimento di Chimica Farmaceutica e Tossicologica, Università di Napoli "Federico II", Via D. Montesano, 49, I-80131 Naples, Italy; [d]Dipartimento di Scienze Ambientali, Seconda Università di Napoli, Via Vivaldi 43, I-81100 Caserta, Italy; and [e]Dipartimento di Scienze Farmaceutiche, Università di Pisa, Via Bonanno, 6, I-56126 Pisa, Italy

Edited by Michael L. Klein, Temple University, Philadelphia, PA, and approved November 2, 2011 (received for review July 28, 2011)

An exhaustive description of the molecular recognition mechanism between a ligand and its biological target is of great value because it provides the opportunity for an exogenous control of the related process. Very often this aim can be pursued using high resolution structures of the complex in combination with inexpensive computational protocols such as docking algorithms. Unfortunately, in many other cases a number of factors, like protein flexibility or solvent effects, increase the degree of complexity of ligand/protein interaction and these standard techniques are no longer sufficient to describe the binding event. We have experienced and tested these limits in the present study in which we have developed and revealed the mechanism of binding of a new series of potent inhibitors of Adenosine Deaminase. We have first performed a large number of docking calculations, which unfortunately failed to yield reliable results due to the dynamical character of the enzyme and the complex role of the solvent. Thus, we have stepped up the computational strategy using a protocol based on metadynamics. Our approach has allowed dealing with protein motion and solvation during ligand binding and finally identifying the lowest energy binding modes of the most potent compound of the series, 4-decyl-pyrazolo[1,5-*a*]pyrimidin-7-one.

ADA | well-tempered metadynamics | ligand/protein docking | path collective variables | reweighting algorithm

Adenosine Deaminase (ADA) regulates the purine metabolism by catalyzing the irreversible hydrolysis of adenosine to inosine and 2′-deoxyadenosine to 2′-deoxyinosine. Thus, this enzyme plays a crucial role in many pathologies such as inflammation, some types of cancer, and others which are strictly connected to the physiological level of these nucleosides (1–5). Despite great efforts in developing ADA inhibitors, only Pentostatin is currently in clinical use (**I** in Fig. S1) (3). However, recent progress has been reported by Terasaka et al. and by some of us who have developed a new generation of nonnucleoside ADA inhibitors (**II**, **III**, and **IV** in Fig. S1) (6–11). Unfortunately, the understanding at molecular level of the ligand/ADA interaction is hampered by the pronounced ability of the active site to accommodate different inhibitors and by the crucial role played by water molecules during ligand binding. A rational drug design is further complicated by the fact that in response to different inhibitors, ADA can assume either an open or a closed conformation by changing the position of the H3 α-helix (Thr57-Ala73). The open conformation corresponds to the apo-form (PDB ID code 3iar) and is preferred when a nonnucleoside inhibitor is bound (12, 13). In this case, the active site presents a hydrophilic subsite S0 and three hydrophobic subsites F0, F1, and F2 (Fig. 1*A*) (13). The S0 subsite is defined by the structural gate formed by a β-strand (Leu182-Asp185) and two leucine side chains attached to the H3 α-helix while the F0 site is formed by the hydrophobic side chains of the H3 α-helix. In the open form the H3 α-helix assumes a conformation that exposes two additional hydrophobic subsites in the upper part the helix, namely F1 and F2 (13). The

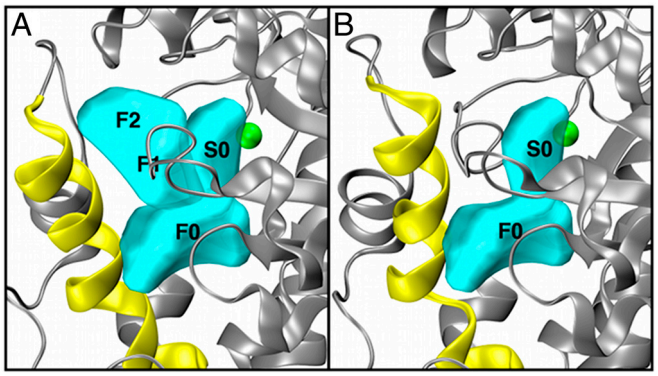


Fig. 1. Representation of ADA in the open (*A*) and closed form (*B*).

binding of a ground-state inhibitor, like the hydroxylated form of adenosine (HDPR) (**V** in Fig. S1) (14), induces a conformational change that leads to the closed form. In this case, the gate is closed, the F1 and F2 sites are no longer accessible and the active site is composed only by S0 and F0 (Fig. 1*B*).

While the structural difference between the two conformations is believed to be the result of binding to different inhibitors, the specific details of the enzyme conformational changes are still unclear. For instance, Kinoshita, et al. (15) have recently reported that the ground-state inhibitor EHNA (**VI** in Fig. S1), expected to bind the closed form, has been cocrystallized rather surprisingly in the open one. In fact, experimental evidences suggest that the n-hexyl group of EHNA interacts with the narrow hydrophobic entrance of the ADA active site destabilizing the closed form. Another example of the difficulty of predicting a ligand binding model comes from the inhibitor FR221647, which has also been found to bind to the open conformation in spite of having been designed as a binder of the closed form (12).

In this confusing scenario, a practical approach is to find a lead compound with whatever strategy works (high-throughput screening, ligand-based design etc.), understand at molecular level the lead/protein recognition mechanism and finally, if necessary, try to rationally improve the lead.

Author contributions: V.L., L. Marinelli, C.L.M., F.D.S., E.N., and M.P. designed research; V.L., C.L.M., S.S., and L. Mugnaini performed research; V.L., C.L.M., and S.S. contributed new reagents/analytic tools; V.L., C.L.M., and M.P. analyzed data; and V.L., L. Marinelli, S.C., C.L.M., F.D.S., E.N., and M.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. E-mail: vittoriolimongelli@gmail.com or parrinello@phys.chem.ethz.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1112181108/-/DCSupplemental.

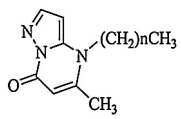
Here we have carried out the first two steps in this strategy by applying a combined experimental and theoretical approach. First we discovered a new class of inhibitors combining a number of functional groups that are present in many of the reported inhibitors (Table 1). Among these new inhibitors, the most potent is 4-decyl-pyrazolo[1,5-*a*]pyrimidin-7-one, here named **3b**. In the lack of further experimental information we performed an extensive series of calculations aimed at understanding the ligand/protein molecular recognition process. We first performed standard docking calculations, which are the most economical way of obtaining molecular information on ligand/protein interaction. Unfortunately, all the docking attempts failed due to the ADA flexibility and the active role of the water molecules. Thus we had to step up the computational strategy and use metadynamics (16, 17). Metadynamics has already been applied with success to difficult docking processes (18–20). After extensive calculations, the binding mode of **3b** has been understood revealing also the important role played by protein motion and solvent during the ligand binding. We are comforted in our conclusions by a number of experimental data and the similarity of our binding mode with that of other ADA inhibitors.

Results

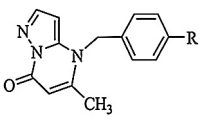
Ligand-Based Design. Prompted by the promising results of the 4-aminopyrazolo[3,4-*d*]pyrimidine (APPs) series (compound **III** in Fig. S1) as ADA inhibitors, recently identified by some of us (9, 10), and using our synthetic expertise in the pyrazolopyrimidines field, we decided to further investigate this scaffold. Thus by modifying the heterocyclic ring, a new series of compounds bearing a pyrazolo[1,5-*a*]pyrimidin-7-one system has been developed. Our investigation started synthesizing compounds **3a–c**, carrying a *n*-nonyl, *n*-decyl and *n*-undecyl chain, respectively, in the position 4 of the heterocyclic core. As we demonstrated in the APPs series (9), a long, linear, and hydrophobic alkyl chain can interact favorably with the enzyme active site. In fact, tested on the bovine spleen ADA, compounds **3a–c** have been found to potently inhibit ADA, showing K_i values in the nanomolar range (Table 1). Moreover, an activity trend similar to the parent APPs was observed as the ranking of inhibitory potency being $n = 8$ (**3a**) $< n = 9$ (**3b**) $> n = 10$ (**3c**). In order to further investigate the role of the 4-alkyl chain in the interaction with the active site of the enzyme, derivatives **7** and **8**, bearing bulkier and more rigid groups at position 4, were likewise synthesised and tested. In **7** the two phenyl rings are connected through an amide linker while in **8** the linker is an ureic residue. Both these functions are present in other potent ADA inhibitors (7, 12). As summarized in Table 1, neither **7** nor **8** showed any appreciable inhibitory activity, proving that in the pyrazolo[1,5-*a*]pyrimidin-7-one series the preferred substitution pattern is represented by the flexible alkyl chain.

In order to rationalize these experimental data and to proceed with a structure-based lead optimization, it was necessary to understand at molecular level the binding mechanism to ADA of the most potent compound of the series, **3b**. This aim was achieved through an extensive computational study described in details in the following paragraphs.

Table 1. ADA Inhibition Data of Derivatives 3a–c, and 6–8



3a-c



6-8

N°	n	R	K_i (nM) *
3a	8		27.00 ± 0.95
3b	9		1.49 ± 0.03
3c	10		3.12 ± 0.09
6		NH_2	n.a.†
7		$NHCOC_6H_4pCF_3$	n.a.
8		$NHCONHC_6H_4pCF_3$	n.a.
(+)-EHNA			1.14 ± 0.10

*The K_i values are means ± SEM (standard error of the mean).
†n.a.: non active. Inhibition occurred at a concentration higher than 10 μM.

Molecular Docking Studies. We first performed a series of docking simulations of **3b** in the ADA active site using the AutoDock program (vs. 4.0) (21, 22). Both the open and closed conformation were explored applying different docking settings (see *SI Text*). We verified that AutoDock correctly predicts the experimental pose for EHNA and HDPR in their open and closed form, respectively. The docking calculations of **3b** in the ADA closed form gave among others a highly populated cluster in which **3b** binds similarly to the nucleoside inhibitor HDPR. This cluster presents also the lowest energy solution. In this pose **3b** coordinates the zinc ion through the oxygen of the pyrimidinone ring (see Fig. S2*A*). Similarly to HDPR, **3b** H-bonds with Asp295 and Glu217 in the S0 site through the oxygen and nitrogen atoms. On the other hand, the n-decyl chain only partially fills the region which, in the case of HDPR, would be occupied by the ribose ring and points towards the hydrophobic region formed by the Leu62 Phe65 and Leu105 residues at the F0 site.

At first glance this is a plausible result, however in the light of recent experiments (15), which found the ground-state inhibitor EHNA to bind the open form of ADA and not the closed one as expected, and being this ligand, at variance from the other compounds of its class, not directly involved in the coordination of the metal, we decided to substantiate the proposed binding pose with further experiments. Thus, we investigated the ability of **3b** to coordinate the zinc ion performing a number of UV spectroscopy measurements. The UV spectra were recorded upon addition of increasing amount of zinc(II) ion into a homogeneous solution of compound **3b** (see *SI Text*). Surprisingly, no appreciable change in optical density of the absorption was observed, thus indicating that **3b** is not able to coordinate the metal.

This experimental result lead us to carry out additional docking calculations using a water molecule as the fifth coordination group of zinc. The open form of ADA was used in all these attempts and furthermore, some water molecules known to be important for ligand binding, such as those involved in the interactions of EHNA with Gly184 and Glu217 (15), were also explicitly taken into account in some of the docking calculations. Unfortunately, in all these docking attempts, the ligand has never been found to interact with residues known to be important for ligand binding in ADA, such as Asp19, Gly184, and Glu217. Furthermore, none of the water molecules, when included in the docking calculation, has been involved in the ligand binding. These reasons and the poor convergence of the docking calculations have casted doubts on the reliability of these results (see *SI Text* for details).

We ascribe this failure to the fact that docking algorithms take into account enzyme flexibility and solvent effects in a limited way. In particular, the solvent can be taken into account either in an implicit way or by adding, as done here, an appropriate number of explicit water molecules. However, the position of these waters does not change during the calculations and if it is not known beforehand, one risks to make things worse. In the ADA case, while the position of the water coordinating the zinc ion can be defined with a good level of precision, the other water molecules, like those involved in the binding of some of the ADA inhibitors (14, 15, 23), can occupy different positions and assume a number of orientations in response to different inhibitors. Thus, the inclusion of waters through an explicit but fixed model might

be not helpful and could lead to rather inaccurate results as we found to be the case here.

Well-Tempered Metadynamics. The failure of the docking calculations suggests using atomistic simulations in explicit water in order to fully include solvent and protein flexibility. Unfortunately, standard molecular dynamics (MD) can access only a time scale of hundreds of nanoseconds, while processes like ligand/protein docking usually take from microseconds to hundreds of seconds. Thus, the use of enhanced sampling is mandatory. Among the emerging techniques, metadynamics has shown to be useful in the study of ligand docking (18–20). We used here this technique in its new variant named well-tempered metadynamics, which enhances sampling allowing the reconstruction of the free-energy profile of the process of interest by adding an adaptive bias on a selected numbers of collective variables (CVs) (17). The choice of CVs represents a crucial point because, to construct the Free-Energy Surface (FES), the CVs must describe all the slow modes relevant for the process under study (24–26).

The change from closed to open form is mostly determined by the movement of the H3 α-helix and because it represents a slow mode, use of a CV that is able to describe this movement is necessary (see *Methods* for details). Recently, a number of path CVs (PCV) have been developed in our research group (27, 28), which are able to reconstruct the lowest free-energy path that connects an initial and a final state. In the ADA case, assuming that the initial and the final state correspond to ADA with the H3 α-helix in the closed and open conformation, respectively, a PCV is constructed with the cartesian coordinates of the alpha carbons of the residues that determine the α-helix movement. This type of CV has been successfully applied in recent studies on complex protein motion (29, 30) and ligand/protein docking (20).

In order to specify the position of the ligand relative to the enzyme, we used a distance CV which measures the distance between the center of mass of the pyrazolo[1,5-*a*]pyrimidin-7-one ring and that of a group of protein atoms (see Table S1).

Under the acceleration of metadynamics, the ligand is able to explore the whole binding site moving from one free-energy minimum to the next and then going back into the previous one, thus overcoming the large free-energy barriers which are encountered during the binding process. Thanks to these recrossings between different basins, the calculated FES is accurate and quantitatively well characterized with the deepest basin corresponding to the preferred binding mode of the ligand inside the protein. We underline once more that side chain flexibility, the motion of the H3 α-helix, and solvation are fully taken into account. It is essential for the convergence of the calculations the use of the PCV which accelerates the H3 α-helix motion, allowing a continuos change from one conformation to the other.

The Energy Minima. The full exploration of the binding site took approximately the equivalent of 0.3 μs. In Fig. 2 three different minima, **A**, **B**, and **C**, can be identified. Basin **A** is approximately 3 kcal/mol deeper than the other two minima. Here, the ligand is into the innermost binding site while in **B** and **C** it is located in more external sites. A detailed discussion on basin **B** and **C** is given in the *SI Text* while the movie showing the ADA binding site exploration by the ligand can be found as Movie S1.

Basin A. In this basin just as in the case of the EHNA/ADA complex, the protein is in the open conformation (Fig. 3*A*). Here, the ligand engages a number of favorable interactions with the surrounding residues. In particular, the oxygen atom of the pyrazolo [1,5-*a*]pyrimidin-7-one H-bonds to the NH backbone and the hydroxyl group of Ser103, while the nitrogen atom of the pyrazole ring interacts with the Cys153 side chain (Fig. 2). Additionally, the oxygen of the pyrimidinone moiety is involved in a H-bond interaction with Asp19 *via* a water molecule. Besides these polar interactions, the pyrazolo[1,5-*a*]pyrimidin-7-one ring engages several hydrophobic contacts with the surrounding residues. In fact, the methyl group of the pyrimidinone moiety points towards the hydrophobic side chains of Phe65 and Met69, while the aromatic ring lies close to His17 and His214 where a π-π and a T-shaped interaction, respectively, are possible.

The atoms of the *n*-decyl chain were not used in defining the distance CV (see Table S1) thus different conformations of this flexible chain are to be found in the FES minima. After clustering the ligand tail conformations, two main families have been found. In one, Aa, the tail is placed in the hydrophobic F0 subsite where multiple contacts are possible with residues such as Phe61, Leu62, Phe65, Leu106, Trp117, Met155, Pro116, and His157. In the other, Ab, the *n*-decyl chain points towards the upper part of the H3 α-helix filling the F1 and F2 sites where different hydrophobic contacts are engaged with residues like Phe61, Leu58, and Thr269 (see Fig. S3 *A* and *B*). In both cases the pyrazolo[1,5-*a*] pyrimidin-7-one ring is always placed in the same position in the active site and because the protein assumes always the open conformation, the alkyl tail can occupy all these hydrophobic pock-

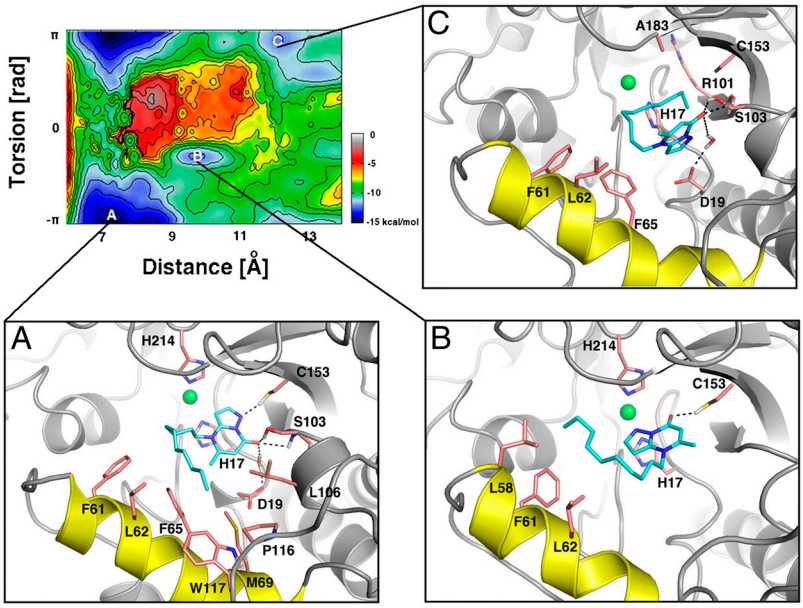


Fig. 2. Representation of the FES of the **3b** binding process to ADA with the three main minima shown as insets, (*A*, *B* and *C*). The FES is represented as function of the distance and dihedral CVs using isosurfaces of 1 kcal/mol. It is important to underline that no bias has been added to the dihedral CV and the statistics collected during the metadynamics simulations on the torsion CV is used to generate the final FES using the reweighting protocol (31). In the insets the ligand and the main interacting residues are displayed as licorice while the protein is represented as cartoon with the H3 α-helix colored in yellow. The ligand/Asp19 water bridge interaction is also shown while only polar hydrogens interacting with the ligand are displayed for clarity.

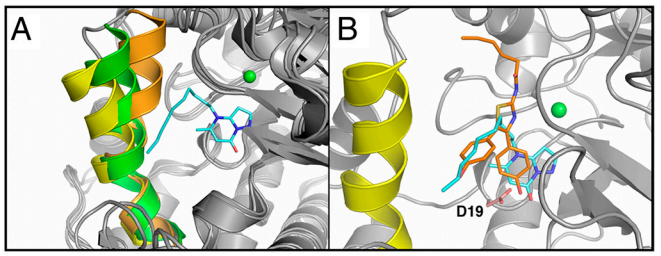


Fig. 3. (*A*) Overlapping of the conformations of ADA found by metadynamics in basin **A** (yellow), ADA complexed with EHNA (green, PDB ID code 2z7g) (15) and ADA complexed with HDPR (orange, PDB ID code 1krm) (14). The metadynamics ADA conformation (yellow) is similar to that ligated to EHNA (open form, green) with a low rmsd value of 1.84 Å calculated for the H3 α-helix backbone atoms. (*B*) Overlapping in the ADA binding site of the **3b** binding conformation calculated by metadynamics (cyan) and the X-ray binding conformation of FR104783 (orange) (PDB ID code 1wxy) (13).

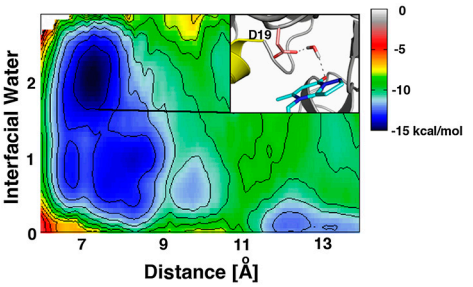


Fig. 4. Representation of the FES of the **3b**/ADA binding process as function of the distance and the interfacial water CVs using isosurfaces of 1 kcal/mol. Looking at the FES it is clear that when the ligand is in basin **A** (distance CV range of 7–8 Å) at least one water bridge interaction is present between **3b** and Asp19 as shown in the inset picture. The energetic contribution to the ligand binding of this interaction is about 5.6 kcal/mol. This is computed as the free-energy difference between the states when the water bridge interaction is present (Interfacial Water CV > 1) and those when this interaction is not formed (Interfacial Water CV < 0.2), considering the ligand in basin **A** (distance CV range of 7–8 Å).

ets. Due to the high flexibility of this long alkyl chain, it is natural to hypothesize that Aa and Ab are equally energetically favorable although the one with the alkyl chain placed at the bottom of the H3 α-helix, Aa (Fig. S3*A*), seems to be more probable as a consequence of the higher hydrophobicity of this region. The relative stability of these two different conformations has been further studied and discussed in details in the following paragraph.

The Alkyl Tail. In order to assess the stability of the two binding conformations Aa and Ab found at basin **A**, one with the alkyl chain pointing towards the F0 site and the other one to the F1 and F2 sites, respectively, MD simulations over 10 ns long have been performed. The results showed a different behavior in the two cases. The binding conformation Aa was very stable during the whole simulation with a low average rmsd value for the ligand (see Fig. S3*C*) and with all the above described ligand/protein interactions fully conserved. In contrast, Ab first maintains its original position in the binding site but, after approximately 3.5 ns of simulation, it moves the alkyl tail down along the H3 α-helix transforming in Aa (see the lower graph in Fig. S3*C*).

The different behaviors of Aa and Ab observed during the MD simulations prompted us to study in more detail the stability of these two conformations. Thus, we performed a new metadynamics run using as CV a distance that is able to distinguish the two tail conformations (see Table S1). Once the calculations converged, the FES was calculated along this CV (CV2) and the distance CV used in the previous metadynamics simulations (CV1) (see Fig. S3*D*) using the reweighting algorithm very recently developed in our group by Bonomi, et al. (31). Thanks to this algorithm, once the free energy of the metadynamics simulation is converged, the FES can be recalculated on CVs different from those used originally in the metadynamics run. The FES shows two minima corresponding to Aa and Ab with the Aa basin wider and about 1.3 kcal/mol deeper than that of Ab. In line with the MD study, these results suggest that although both the poses are possible, Aa has a greater stability.

The Solvent Role. In the binding mode that characterizes basin **A**, a water molecule plays a relevant role in the interaction between **3b** and Asp19 highlighting the importance of the solvent in ligand binding. In order to assess the relevance of this water bridge interaction, the FES along the distance and the interfacial water CV (see *SI Text*) (32) has been calculated using the reweighting protocol (31). Looking at the FES in Fig. 4, it is worth noting that the ligand/ADA complex in basin **A**, corresponding approximately to the distance CV range of 7–8 Å, is always characterized by the presence of at least one interfacial water between the ligand and Asp19 and that the energetic contribution of this water bridge interaction to the ligand binding is about 6 kcal/mol (Fig. 4). This is evidence of the important role of water similarly to what has been found in other ADA inhibitors like EHNA or HDPR (14, 15). A water at similar position but not interacting directly with the ligand can be seen also in the X-ray structure of EHNA/ADA complex (15). Our results suggest that this water has a more than structural role. If in a docking calculation we now include explicitly this water and orient the Asp19 side chain so as to interact with **3b** *via* this water, AutoDock predicts conformations very similar to that found by metadynamics among the most favorable poses (see Fig. S4*A*).

The Asp19 residue is involved also in the binding of other classes of inhibitors. This interaction occurs either through a direct H-bond, like in the case of HDPR (PDB ID code 1krm) (14), or *via* a water bridge interaction as in Coformycin (PDB ID code 3ewc) (23). Furthermore, it is worth mentioning that Kinoshita et al. have suggested that the interactions of EHNA with His17 and Asp19 are anchor points for correctly placing the inhibitor in the active site (15). As described above, both these residues play a role in the **3b** binding mode. We stress that none of this information was used in setting up the simulation protocol.

Experimental Data Rationalization. Our results rationalize the experimental data on the pyrazolo[1,5-*a*]pyrimidin-7-one series (Table 1), which show that the activities of **3a** and **3c** are lower than that of **3b**. We believe that neither **3a** nor **3c** can fill comfortably the F0 site, the one because the alkyl tail is too short, the other because it is too long. We plan to perform further experiments to validate our hypothesis. For the time being we limit ourselves to docking calculations using the protein conformation found by metadynamics and taking explicitly into account the bridging water.

For **3b** AutoDock finds that the pose predicted by metadynamics is among the best scoring ones (see Fig. S4*A*). Note that the docking algorithm did not find the Ab conformation suggesting once more that the preferred and most stable binding conformation is Aa.

In the case of **3a** and **3c**, docking calculations yielded different results. In fact, docking of **3a** showed a number of different binding possibilities and only in the second most populated cluster can be found a pose comparable to that of **3b**. However, in this pose the pyrazolo[1,5-*a*]pyrimidin-7-one ring is translated and rotated by 180 ° thus losing some favorable interactions with ADA (see Fig. S4*B*). In contrast, docking on **3c** found a binding mode very similar to that of **3b** with the pyrazolo[1,5-*a*]pyrimidin-7-one ring slightly displaced from its usual position but still able to form all the main interactions with the protein (see Fig. S4*C*). These results clearly validate our hypothesis that due to the loss or the weakening of some favorable interactions with the protein a

shorter alkyl chain (**3a**) affects the ADA affinity more than the longer one (**3c**) in line with the experimental data.

Comparison with Other ADA Inhibitors. Comparing the binding modes of different ADA inhibitors can be problematic because this enzyme is very complex and similar ligands can bind to the protein in different manners. Still we have found it useful to compare the **3b** binding mode with the FR104783 (**VII** in Fig. S1) X-ray pose (PDB ID code 1wxy) (13). As can be seen in Fig. 3*B*, there is a clear similitude, in fact one of the two phenol rings of FR104783 occupies the bottom of the S0 site just like the pyrazolo[1,5-*a*]pyrimidin-7-one ring of **3b**, and in both cases they are H-bonded to Asp19. Furthermore, in FR104783, the second phenol ring occupies the hydrophobic F0 subsite just like the alkyl tail of **3b**. The alkyloxy branch of FR104783 occupies the F1 and F2 subsites in a manner similar to the Ab pose of **3b** suggesting the utility of targeting this region in future lead optimization studies.

Other findings are further validated by a comparison with other crystallographic structures. For instance, the open conformation assumed by the protein when **3b** is in the basin **A** is similar, albeit slightly more open, to what has been found in the EHNA/ADA complex (Fig. 3*A*) (15). Based on the metadynamics results, we can a posteriori say that the failure of all the preliminary docking trials in which we used the open form of ADA either hydrated (with three water molecules) or dry, is due to: (*i*) a slightly different conformation of the H3 α-helix; (*ii*) the flexibility of the side chains which allows different ligands to bind; and (*iii*) the user defined choice of water molecules to be considered in the calculation. As we have seen in our docking samplings, the latter choice is difficult and without a priori knowing the position occupied by the waters, one can easily incur into errors.

Validation of the Method. The success of the metadynamics simulations is strongly related to the choice of the CVs and when a relevant CV is neglected the reconstruction of the FES fails (20, 33). This implies the necessity of running a number of metadynamics simulations to find out which CVs are needed for the process under study. With the aim of assessing the reproducibility of our results, we have carried out two additional metadynamics simulations using CVs different from the original one (see *SI Text*, Table S1, Fig. S5). The results have indicated that the H3 α-helix motion (PCV) and the ligand move (distance CV) are the most relevant degrees of freedom acting on the system and that the inclusion of other CVs, like the interfacial water CV, is not necessary to achieve the final results. Without knowing a priori which are the most important degrees of freedom in a ligand/protein docking study, it is advisable to use standard CVs, such as distance or torsion CVs, to achieve the best results in the shortest time. Eventually, as we have shown in the *Solvent Role* paragraph, one can construct, using the reweighting algorithm, the FES as a function of different CVs thus investigating the role of these degrees of freedom in the event under investigation. Another intriguing possibility is to use a very recently developed formalism, namely orthogonal space random walk (34). This algorithm is able to couple the CVs motion with the system dynamics through structural relaxation, thus allowing the exploration of hidden degrees of freedom. In doing so, CVs that are neglected during the docking study, such as protein motion or solvent effect, can be taken into account. This algorithm has been already used to sample small protein motion (35) and it would be interesting to apply it in more complex studies such as ligand/protein docking.

Discussion

As discussed by Leach, et al. (36), docking methodologies seem to have reached a plateau in recent years and are waiting for an important breakthrough. In this study on a series of novel ADA inhibitors we have been confronted with some of the major docking hurdles and overcome them through the use of an advanced computational technique. Particularly, a metadynamics-based protocol has been employed to unveil the binding pose of the most potent compound of the series **3b**, taking into account effects such as protein flexibility and solvation whose inaccurate description renders docking not reliable. Our findings are in agreement with a number of experimental data and supported by extensive calculations. Among the encouraging experimental facts we mention the similitude of the **3b** binding mode with the X-ray docking pattern of other ADA inhibitors. Furthermore, the solvent has been found to play a fundamental role in the binding of our inhibitor. In fact, a water molecule mediates the important ligand interaction with Asp19, that is known to be involved in the binding of many other potent ADA inhibitors.

Certainly in terms of speed and computational cost, docking protocols remain the first choice technique in lead discovery and optimization, nevertheless in some cases, like ADA, the use of more advanced techniques is necessary. Unfortunately, metadynamics-based protocols are computationally expensive and are not applicable for high-throughput screening purposes. For instance, in the present case, considering only the final **3b**/protein docking metadynamics, the calculation required about 10^5 CPU hours on High Performance Computing cluster, thus being feasible if performed on just a few ligands. However, the high scalability of molecular dynamics simulation codes and the increasing computational power can help to achieve results in more reasonable time.

The information derived from this study will be very useful for the following lead optimization strategies. In fact, a low inhibitory activity, nanomolar or picomolar, is a prerequisite to achieve potent enzymatic inhibition and to develop drug candidates. However, as very recently discussed by Gleeson, et al. (37), the level of activity of molecules is not the only determining factor to achieve drug efficacy because even other properties of molecules (e.g., log P, pharmacokinetics profile) play crucial roles. Our results provide the possibility to rationally optimize the potent lead compound **3b**, representing an important opportunity to improve the protein/ligand interaction and the pharmacokinetics properties of the new designed compounds.

Methods

Metadynamics Simulations. Before doing metadynamics simulations the **3b**/ADA complex was equilibrated through 5 ns MD under NPT conditions at 1 atm and 300 K using the parmff99SB Amber force field (38, 39), as implemented in version 2.7 of the NAMD code (40). For the zinc subsite the bound model was used adding to the force field the bond and angle parameters reported by Coi, et al. (41). The Amber charges were applied to all the protein and waters atoms except for the ligand and the zinc subsite. For these the restrained electrostatic potential (RESP) charges were used (see *SI Text*). All the simulations were carried out in explicit solvent using the TIP3P water model (42) using periodic boundary conditions. The PLUMED plugin was used to carry out metadynamics calculations with the NAMD code (43).

It is important to underline that because the ligand starting conformation does not affect the final results of a well converged metadynamics simulation, we can arbitrarily decide to use one of the poses predicted by the docking program. In our case, a pose from the most populated cluster in an open form docking run was taken as the starting conformation.

The estimation $F(s,t)$ at time t of the free-energy surfaces $F(s)$ as a function of the CV s was determined by metadynamics (15) in its new well-tempered variant (16), using the following formula:

$$F(s,t) = -\frac{T+\Delta T}{\Delta T}V(s,t),$$

where $V(s,t)$ is the bias potential added to the system and T is the temperature of the simulation. ΔT is the difference between the fictitious temperature of the CV and the temperature of the simulation. The bias potential is made up by the sum of the Gaussians deposited along the trajectories of the CVs. Thanks to this new formalism, one can increase barrier crossing and facilitate the exploration in the CVs space by tuning ΔT. A Gaussians deposition rate of 0.5 kcal/mol per picosecond was initially used and gradually decreased on the basis of the adaptive bias with a ΔT of 3,300 K.

The Path Collective Variables. In order to simulate the movement of the H3 α-helix, we used the path CV (PCV) (27) using as metrics the rmsd of the Cα atoms of residues from Ile50 to Pro70. The PCVs are extremely powerful whenever one wants to study a transition between two states A and B. As state A we chose the X-ray conformation of ADA in the closed state (PDB ID code 1krm) (14), while state B represents the open form (PDB ID code 2z7g) (15). Now, let $S(R)$ be a reduced representation of a generic configuration R. If the choice of S is appropriate, we would expect the reactive trajectories to be bundled in a narrow tube around the path. To trace this path, we follow the procedure of Branduardi et al. introducing the two variables $s(R)$ and $z(R)$:

$$s(R) = \frac{1}{P-1}\frac{\sum_{l=1}^{P}(l-1)e^{-\lambda||S(R)-S(l)||^2}}{\sum_{l=1}^{P}e^{-\lambda||S(R)-S(l)||^2}},$$

$$z(R) = -\frac{1}{\lambda}\ln\left(\sum_{l=1}^{P}e^{-\lambda||S(R)-S(l)||^2}\right),$$

which measure the intercept and the distance of a microscopic configuration R from the reference path $S(l)$, respectively. P is the number of frames that define $S(l)$, $||S(R)—S(l)||$ is calculated as the rmsd and λ is proportional to the inverse of the mean square displacement between successive frames. As the structure of the H3 α-helix has only slight differences between the open and closed conformation, we described the transition between the closed and open conformation of the enzyme only using two frames $S(1)$ and $S(2)$, being the closed (state A) and the open conformation (state B), respectively. The λ value was set to 0.75 Å^{-2}. We performed metadynamics only in the space of $s(R)$ while $z(R)$ was constrained to $z(R) < 2$Å^2. This choice gives the possibility for the system to explore conformations different from the reference states, avoiding at the same time the loss of the alpha helical secondary structure.

In addition to $s(R)$, we have also used a distance and, in some metadynamics runs, a torsion CV to describe the different conformations of the ligand during the simulations (see Table S1).

Additional Materials and Methods information is given in the *SI Text* and in *Table S2*, *Table S3*, and *Table S4*.

ACKNOWLEDGMENTS. The authors thank Massimiliano Bonomi, Anna Berteotti, and Antonio Randazzo for useful discussions and acknowledge funding from the European Union (Grant ERC-2009-AdG-247075) and Swiss National Supercomputing Center—(CSCS) under project ID s233.

1. Fischer A, Hacein-Bey-Abina S, Cavazzana-Calvo M (2010) 20 years of gene therapy for SCID. *Nat Immunol* 6:457–460.
2. LePage G-A, Worth I-S, Kimball A-P (1976) Enhancement of the antitumor activity of arabinofuranosyladenine of 2′-Deoxy-coformicin. *Cancer Res* 36:1481–1485.
3. Flinn I-W, et al. (2000) Long-term follow-up of remission duration, mortality, and second malignancies in hairy cell leukemia patients treated with pentostatin. *Blood* 96:2981–2986.
4. Law W-R, Valli V-E, Conlon B-A (2003) Therapeutic potential for transient inhibition of Adenosine Deaminase in systemic inflammatory response syndrome. *Crit Care Med* 31:1475–1481.
5. Oskouie F-G, et al. (2011) High levels of Adenosine Deaminase on dendritic cells promote autoreactive T cell activation and diabetes in nonobese diabetic mice. *J Immunol* 186:6798–6806.
6. Terasaka T, et al. (2004) Structure-based design and synthesis of non-nucleoside, potent, and orally bioavailable Adenosine Deaminase inhibitors. *J Med Chem* 47:2728–2731.
7. Terasaka T, et al. (2004) Structure-based design, synthesis, and structure-activity relationship studies of novel non-nucleoside Adenosine Deaminase inhibitors. *J Med Chem* 47:3730–3743.
8. Terasaka T, et al. (2005) Rational design of non-nucleoside, potent, and orally bioavailable Adenosine Deaminase inhibitors: predicting enzyme conformational change and metabolism. *J Med Chem* 48:4750–4753.
9. Da Settimo F, et al. (2005) Novel, highly potent Adenosine Deaminase inhibitors containing the pirazolo[3,4-d]pyrimidine ring system, synthesis, structure-activity relationships, and molecular modeling studies. *J Med Chem* 48:5162–5174.
10. Antonioli L, et al. (2007) Inhibition of Adenosine Deaminase attenuates inflammation in experimental colitis. *J Pharmacol Exper Ther* 322:435–442.
11. La Motta C, et al. (2009) Exploiting the pyrazolo[3,4-*d*]pyrimidin-4-one ring system as a useful template to obtain potent Adenosine Deaminase inhibitors. *J Med Chem* 52:1681–1692.
12. Terasaka T, Kinoshita T, Kuno M, Nakanishi I (2004) A highly potent non-nucleoside Adenosine Deaminase inhibitor: efficient drug discovery by intentional lead hybridization. *J Am Chem Soc* 126:34–35.
13. Kinoshita T, et al. (2005) Structural basis of compound recognition by Adenosine Deaminase. *Biochemistry* 44:10562–10569.
14. Kinoshita T, Nishio N, Nakanishi I, Sato A, Fujii T (2003) Structure of bovine Adenosine Deaminase complexed with 6-hydroxy-1,6-dihydropurine riboside. *Acta Crystallogr D* 59:299–303.
15. Kinoshita T, Tada T, Nakanishi I (2008) Conformational change of Adenosine Deaminase during ligand-exchange in a crystal. *Biochem Biophys Res Commun* 373:53–57.
16. Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci USA* 99:12562–12566.
17. Barducci A, Bussi G, Parrinello M (2008) Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett* 100:020603.
18. Gervasio F-L, Laio A, Parrinello M (2005) Flexible docking in solution using metadynamics. *J Am Chem Soc* 127:2600–2607.
19. Masetti M, Cavalli A, Recanatini M, Gervasio F-L (2009) Exploring complex protein-ligand recognition mechanisms with coarse metadynamics. *J Phys Chem B* 113:4807–4816.
20. Limongelli V, et al. (2010) Molecular basis of cyclooxygenase enzymes (COXs) selective inhibition. *Proc Natl Acad Sci USA* 107:5411–5416.
21. Morris G-M, et al. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19:1639–1662.
22. Huey R, Morris G-M, Olson A-J, Goodsell D-S (2007) A semiempirical free energy force field with charge-based desolvation. *J Comput Chem* 28:1145–1152.
23. Ho M-C, et al. (2009) Structural and metabolic specificity of methylthiocoformycin for malarial Adenosine Deaminases. *Biochemistry* 48:9618–9626.
24. Gear C-W, Kevrekidis I-G, Theodoropoulos C (2002) 'Coarse' integration/bifurcation analysis via microscopic simulators: micro-Galerkin methods. *Comput Chem Eng* 26:941–963.
25. Hummer G, Kevrekidis I-G (2003) Coarse molecular dynamics of a peptide fragment: free energy, kinetics, and long-time dynamics computations. *J Chem Phys* 118: 10762–10773.
26. Parrinello M (2008) *Physical Biology From Atoms to Medicine*, ed AH Zewail (Imperial College Press, London), pp 247–265.
27. Branduardi D, Gervasio F-L, Parrinello M (2007) From A to B in free energy space. *J Chem Phys* 126:054103.
28. Bonomi M, Branduardi D, Gervasio F-L, Parrinello M (2008) The unfolded ensemble and folding mechanism of the C-terminal GB1 beta-hairpin. *J Am Chem Soc* 130:13938–13944.
29. Berteotti A, et al. (2009) Protein conformational transitions: the closure mechanism of a kinase explored by atomistic simulations. *J Am Chem Soc* 131:244–250.
30. Pfaendtner J, Branduardi D, Parrinello M, Pollard T-D, Voth G-A (2009) Nucleotide-dependent conformational states of actin. *Proc Natl Acad Sci USA* 106:12723–12728.
31. Bonomi M, Barducci A, Parrinello M (2009) Reconstructing the equilibrium Boltzmann distribution from well-tempered metadynamics. *J Comput Chem* 30:1615–1621.
32. Pietrucci F, Marinelli F, Carloni P, Laio A (2009) Substrate binding mechanism of HIV-1 protease from explicit-solvent atomistic simulations. *J Am Chem Soc* 131:11811–11818.
33. Laio A, Gervasio F-L (2008) Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep Prog Phys* 71:126601–126622.
34. Zheng L, Chen M, Yang W (2008) Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems. *Proc Natl Acad Sci USA* 105:20227–20232.
35. Lee S, Chen M, Yang W, Richards NGJ (2010) Sampling long time scale protein motions: OSRW simulation of active site loop conformational free energies in formyl-CoA:oxalate CoA transferase. *J Am Chem Soc* 132:7252–7253.
36. Leach AR, Shoichet BK, Peishoff CE (2006) Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. *J Med Chem* 49:5851–5855.
37. Gleeson MP, Hersey A, Montanari D, Overington J (2011) Probing the links between in vitro potency, ADMET and physicochemical parameters. *Nat Rev Drug Discov* 10:197–208.
38. Cornell W-D, et al. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 117:5179–5197.
39. Hornak V, et al. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65:712–725.
40. Phillips J-C, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802.
41. Coi A, Tonelli M, Ganadu M-L, Bianucci A-M (2006) Binding free energy calculations of Adenosine Deaminase inhibitors. *Bioorg Med Chem* 14:2636–2641.
42. Jorgensen W-L, Madura J-D (1983) Solvation and conformation of methanol in water. *J Am Chem Soc* 105:1407–1413.
43. Bonomi M, et al. (2009) PLUMED: a portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* 180:1961–1972.

Supporting Information

Limongelli et al. 10.1073/pnas.1112181108

SI Text

Additional Docking Trials. Once the docking in the Adenosine Deaminase (ADA) closed form was seen to fail in providing any convincing result we performed a number of extra docking simulations in the ADA open form using as protein coordinates the EHNA/ADA complex (PDB ID code 2z7g) (4). Unfortunately, the docking calculations did not yield any reasonable result. In fact, the two poses, one with the lowest binding score and the other belonging to the most populated cluster, do not place the ligand in the catalytic site but in more external solvent exposed regions. Particularly, the best scoring pose occupies the upper part of the H3 α-helix H-bonding with the backbone NH of Leu58 and engaging a number of hydrophobic contacts with the side chains of Leu58, Phe61, Thr187, and Thr269 (Fig. S2*B*). Similarly to the best scoring pose, also in the most frequent one the ligand is placed far from the catalytic site making polar contacts in the external part of the enzyme with the His157 and Gln158 side chains and hydrophobic contacts with Phe65 (Fig. S2*C*). The fact that in none of the docking solutions **3b** interacts with residues like Asp19, Gly184, and Glu217, involved in the binding of other ADA inhibitors, and the low convergence of the docking calculations has casted doubts on the reliability of these results.

Therefore, once again further docking runs were carried out. This time three water molecules, that mediate the interaction of EHNA with Gly184 and Glu217, were included in the protein structure. This choice was done because in our opinion the omission of these functional waters in the previous calculations could be at the basis of the failure. The best ranked conformation and the pose representing the most populated cluster occupies similarly the ADA binding site (Fig. S2 *D* and *E*). In fact, the former presents the pyrazolo[1,5-*a*]pyrimidin-7-one ring involved in H-bonds with the His157 side chain and the backbone NH of Asp185, while the alkyl tail points towards residues such as Phe61 and Phe65. On the other hand, the most occurred pose shows the pyrazolo[1,5-*a*]pyrimidin-7-one ring H-bonding with the His157 and Gln158 side chains and the *n*-decyl chain pointing the inner part of the enzyme where hydrophobic interactions with Phe65 are possible. Even in this last attempt, the ligand has never been found to interact with residues known to be important for ligand binding and furthermore, none of the water molecules, explicitly included in the docking calculation, has been involved in the ligand binding suggesting us to discard also these solutions.

The Energy Minimum Basins B and C Found by Metadynamics. ***Basin B.*** This minimum is approximately 3.1 kcal/mol higher than basin **A** (see Fig. 2 in the main text). In this pose the pyrazolo[1,5-*a*]pyrimidin-7-one ring is placed very close to the zinc ion although a coordination of the metal is not possible due to the presence of the zinc coordinating water. At this site only a T-shaped interaction with His17 and a H-bond with Cys147 are possible, while all the other interactions of the nucleotide-like ring described in pose **A** are lost. In this minimum the enzyme has been found in the open form which allows the ligand alkyl tail to engage hydrophobic contacts with residues present in the upper part of the H3 α-helix like Leu58, Phe61 and Leu62. The higher energy value of basin **B** and the smaller number of interactions at this site suggest to consider **B** as an intermediate pose of the ligand during the binding process.

Basin C. The energy minimum **C** (see Fig. 2 in the main text) is energetically comparable to **B**. At variance from **A** and **B**, here the pyrazolo[1,5-*a*]pyrimidin-7-one ring is located in the external part of the enzyme while the alkyl chain points towards the active site. The O-7 of the pyrazolo[1,5-*a*]pyrimidin-7-one ring H-bonds with the Ser103 side chain and a water molecule, while the N-1 interacts *via* a water bridge interaction with Asp19. Also in this case ADA has been found in the open conformation and while the C1′, C2′, C3′, and C4′ atoms of the ligand tail interacts with Phe61, Leu62 and Phe65, the rest of the chain points deep in the active site interacting with His17, Arg101, Cys153, and Ala183. It is interesting to note that this binding conformation is similar to some of those predicted by AutoDock when the ADA open form is used. Actually, in the poses predicted by AutoDock the pyrazolo[1,5-*a*]pyrimidin-7-one ring visits an even more external part of the enzyme, however, the overall volume occupied by **3b** in the catalytic site is almost the same. Basin **C** is 3.3 kcal/mol higher than the deepest minimum **A** suggesting a good stability of this pose. The external position of this pose in the active site leads us to suppose that **C** might represent a presite for the entry or the exit of the ligand from the catalytic site.

Additional metadynamics simulations. Two extra metadynamics simulations have been carried out to assess the results of the original metadynamics run. In the first one, a distance and a torsion collective variable (CV) (Table S1) have been used with no path CV (PCV) applied to the system. In this run, while the minimum **A** has been found, the basins **B** and **C** disappeared in the Free-Energy Surface (FES) (Fig. S5*B*). Actually, the minimum **A** is slightly shifted to higher distance CV values (Fig. S5) reflecting slight differences in pose **A** between the two matadynamics runs. In fact, in the simulation without the PCV the pyrazolo[1,5-*a*]pyrimidin-7-one ring cannot optimally interact with Asp19, Ser103 and Cys153. This is due to the fact that no PCV has been used in this run thus not allowing to fully explore the protein flexibility. Therefore these results have shown that the use of the PCV to sample the protein flexibility is needed to find the correct binding mode and to get an accurate estimation of the free energy. Consequently, the subsequent metadynamics simulation has been carried out using again the PCV. In order to assess the independency of the results from the choice of the distance CV, this time a distance CV different from that of the original metadynamics has been used (see Table S1). In order to facilitate the comparison of the results of the different simulations, once the calculations converged, the FES has been recalculated through the reweighting algorithm (13) using the distance and the torsion CV used in the original run (Fig. S5*C*). Looking at the FES, one deepest energy minimum is present, which has the same CVs values of **A**. In fact, in this minimum the ligand interacts with the protein similarly to what was found in the original metadynamics sampling. Particularly, the pyrazolo[1,5-*a*]pyrimidin-7-one ring interacts directly with Ser103 and Cys153 and through a water molecule with Asp19, while the *n*-decyl chain occupies the F0 subsite. It is interesting to note that basin **B** and **C** are again present in the FES. The similarities of these results with those of the original run underlines that it is essential to include a CV describing the H3 α-helix flexibility. Failing to do so allows only to reproduce some of the FES features. If instead we explicitly use the PCV that describes the H3 α-helix motion then we find a good level of reproducibility even if different ligand/protein distance CV is used. Finally, the constant presence of **A** across all these calculations highlights the reproducibility of our results despite the different CVs setting used, thus supporting the reliability of our findings.

Synthesis of pyrazolo[1,5-a]pyrimidinones. The target compounds were prepared as outlined in the Fig. S6. Reaction of the commercially available 3-aminopyrazole **1** with the suitable alkyl bromide, in the presence of anhydrous K_2CO_3, gave a mixture of different products, which were separated by flash chromatoghaphy. The pure compounds were identified by analytical and spectral data, proving that through reaction with both 1-bromononane and 1-bromoundecane the desired N-alkyl-1*H*-pyrazol-3-amines **2a,c** were obtained with the corresponding di-alkylated by products **2d,e**, while by using 1-bromodecane the target N-decyl-1*H*-pyrazol-3-amine **2b** was obtained with its regioisomer **2f**. Reaction of **2a–c** with ethyl-3-oxobutanoate gave the target pyrazolo [1,5-*a*]pyrimidin-7-ones **3a–c** and their regioisomers pyrazolo[1,5-*a*]pyrimidin-5-ones **4a–c**, which were separated by flash cromatography and characterized by spectral data. Alkylation of 3-aminopyrazole **1** with 1-(chloromethyl)-4-nitrobenzene afforded a mixture of mono and di-alkylated products, **2g–i**, obtained pure through flash chromatoghaphy. Cyclization of **2g** with ethyl-3-oxobutanoate provided the target heterocycle **5**. This latter was then subjected to catalytic hydrogenation, performed under atmospheric pressure and room temperature in the presence of Pd/C, to obtain the key intermediate 4-(4-aminobenzyl)-5-methylpyrazolo[1,5-*a*]pyrimidin-7(4*H*)-one **6**. Reaction of **6** with 4-(trifluoromethyl)benzoyl chloride, in the presence of triethylamine, let to the inhibitor **7**, while treatment with 1-isocyanato-4-(trifluoromethyl)benzene gave the inhibitor **8**. Both reactions were conducted under microwave irradiation, which greatly enhanced the speed of the reaction simplifying the experimental protocol.

Adenosine Deaminase inhibition assay. The in vitro activity of ADA (type IX from bovine spleen, 150–200 U/mg, Sigma) was determined spectrophotometrically by monitoring for 2 min at 262 nm the decrease in absorbance resulting from deamination of adenosine into inosine, catalyzed by the enzyme. The change in adenosine concentration/min, in absence and in the presence of the newly synthesized compounds, was determined using a Beckman DU-64 kinetics software program (Solf Pack TM Module) following a previously reported protocol (13).

Chemistry. Melting points were determined using a Reichert Köfler hot-stage apparatus and are uncorrected. Infrared spectra were recorded with a FT-IR spectrometer Nicolet/Avatar in Nujol mulls. Routine ^{1}H NMR spectra were recorded in DMSO-d_6 solution on a Varian Gemini 200 spectrometer operating at 200 MHz. Evaporation was performed in vacuo (rotary evaporator). Anhydrous sodium sulfate was always used as the drying agent. Analytical TLC was carried out on Merck 0.2 mm precoated silica gel aluminum sheets (60 F-254), with visualization by irradiation with a UV lamp. Flash chromatography was performed with Merck Silica gel 60 (230–400 mesh ASTM). The microwave-assisted procedures were carried out in sealed vessels using a CEM/Discover LabMate 220 VAC/50 Hz microwave system. Elemental analyses were performed by our Analytical Laboratory and agreed with theoretical values to within ±0.4%.

3-Aminopyrazole, 1-(chloromethyl)-4-nitrobenzene, ethyl-3-oxobutanoate, 4-(trifluoromethyl)benzoyl chloride, 1-isocyanato-4-(trifluoromethyl)benzene and the alkyl bromides used to obtain compounds **2a–c** were from Sigma-Aldrich. All other chemicals were of reagent grade.

General procedures for the synthesis of alkyl pyrazol-3-amines 2a–f. The suitable alkyl bromide (12 mmol) was added drop wise to a suspension of 3-aminopyrazole **1** (0.830 g, 10 mmol) and anhydrous potassium carbonate (1.66 g, 12 mmol) in 25 mL of dimethylformamide (DMF), and the resulting mixture was stirred at 80 °C until the disappearance of the starting material (TLC, 2–3 h). After cooling, the inorganic material was filtered off and the solution was evaporated to dryness under reduced pressure. The crude obtained was flash chromatographed (eluting system: ethyl acetate/cyclohexane 3/7) to separate the target products **2a–c** from the undesired **2d–e**. Each compound, obtained either as a solid or as a pure oily product, was then characterized by analytical and spectral data (Table S2).

N-(4-nitrobenzyl)-pyrazol-3-amine 2 g, N,1-bis(4-nitrobenzyl)-pyrazol-3-amine 2h. 1-(4-nitrobenzyl)-pyrazol-3-amine 2i. A solution of 1-(chloromethyl)-4-nitrobenzene (2.05 g, 12 mmol) in DMF was added drop wise to a suspension of 3-aminopyrazole **1** (0.830 g, 10 mmol) and anhydrous potassium carbonate (1.66 g, 12 mmol) in 25 mL of DMF, and the resulting mixture was stirred at 50 °C for 1 h. After cooling, the inorganic material was filtered off and the solution was evaporated to dryness under reduced pressure. The crude obtained was flash chromatographed (eluting system: ethyl acetate/cyclohexane 5/5) to separate the target product **2g** from the undesired **2h,i**. Each compound was then recrystallized from the appropriate solvent (Table S2).

General procedures for the synthesis of 4-alkyl-5-methylpyrazolo[1,5-a]pyrimidin-7-ones 3a–c and 4-alkyl-7-methylpyrazolo[1,5-a]pyrimidin-5-ones 4a–c. A mixture of the suitable N-alkylpyrazol-3-amine **2a–c** (1.00 mmol) and ethyl-3-oxobutanoate (0.15 mL, 1.20 mmol) was irradiated with microwaves at 140 °C for 20 min. The cooled residue was diluted with ethyl acetate and flash chromatographed (eluting system: ethyl acetate/cyclohexane 5/5) to separate compounds **3a–c** from the isomers **4a–c**. Each product was then recrystallized from the appropriate solvent (Table S3).

5-methyl-4-(4-nitrobenzyl)pyrazolo[1,5-a]pyrimidin-7(4H)-one 5. N-(4-nitrobenzyl)-pyrazol-3-amine **2g** (0.218 g, 1.00 mmol) and ethyl-3-oxobutanoate (0.150 mL, 1.20 mmol) were mixed thoroughly and irradiated with microwaves at 110 °C for 10 min. The cooled residue was then diluted with toluene and the solid separated was collected and recrystallized from the appropriate solvent (Table S3).

5-methyl-4-(4-aminobenzyl)pyrazolo[1,5-a]pyrimidin-7(4H)-one 6. A suspension of 5-methyl-4-(4-nitrobenzyl)pyrazolo[1,2-*a*]pyrimidin-7(4*H*)-one **5** (0.284 g, 1.00 mmol) and 10% Palladium on Carbon (0.100 mmol) in 250 mL of absolute ethanol was hydrogenated at atmospheric pressure and room temperature until the theoretical uptake of hydrogen was achieved. After filtering the catalyst, the solvent was evaporated to dryness to give a pale yellow solid which was collected and recrystallized from the appropriate solvent (Table S3).

N-(4-((5-methyl-7-oxopyrazolo[1,5-a]pyrimidin-4(7H)-yl)methyl)phenyl)-4-(trifluoromethyl) benzamide 7. 5-Methyl-4-(4-aminobenzyl) pyrazolo[1,5-*a*]pyrimidin-7(4*H*)-one **6** (0.254 g, 1.00 mmol), 4-(trifluoromethyl)benzoyl chloride (0.178 mL, 1.20 mmol), and triethylamine (0.167 mL, 1.20 mmol) were mixed thoroughly and irradiated with microwaves at 140 °C for 10 min. The cooled residue was then diluted with ice-water and the solid separated was filtered and purified by recrystallization from the appropriate solvent to give the target compound, **7** (Table S3).

1-(4-((5-methyl-7-oxopyrazolo[1,5-a]pyrimidin-4(7H)-yl)methyl)phenyl)-3-(4-(trifluoromethyl)phenyl)urea 8. 5-Methyl-4-(4-aminobenzyl)pyrazolo[1,5-*a*]pyrimidin-7(4*H*)-one **6** (0.254 g, 1.00 mmol) and 1-isocyanato-4-(trifluoromethyl)benzene (0.130 mL, 1.20 mmol) were mixed thoroughly and irradiated with microwaves at 140 °C for 10 min. The cooled residue was then diluted with toluene and the solid separated was filtered and purified by recrystallization from the appropriate solvent to give the target compound, **8** (Table S3).

UV absorption spectroscopic analyses. Stock solutions of 100.0 μM $ZnCl_2$, from Sigma-Aldrich, and of 50.00 μM **3b** were prepared by dissolving the required amount of substance in distilled water. For **3b** the solubility was facilitated by using DMSO. Metal coordination experiments were carried out by adding from 0 to 2.0 mL of zinc(II) ion solution, in 0.5 mL increments, to 2.0 mL of 50.00 μM **3b** and diluting to a total volume of 5.0 mL with potassium phosphate pH = 7.2. Titration was monitored by UV-Vis spectroscopy measuring the absorbance of each solution, at ranging from 190 to 260 nm and at $T = 20\,°C$, using a PerkinElmer Lambda 25 spectrophotometer.

Docking simulations. Molecular docking simulations of **3b**, EHNA and HDPR in the three-dimensional X-ray structures of ADA were carried out using the AutoDock software package (version 4.0) (1, 2). The protein coordinates were taken from the HDPR/ADA (PDB ID code 1krm) (3) and EHNA/ADA complex (PDB ID code 2z7g) (4) for the closed and open form, respectively.

Ligands and protein setup. The charges on the ligand were computed using the restrained electrostatic potential (RESP) fitting procedure (5). First the ESP were calculated by means of the Gaussian package (6) using a 6-31G* basis set at Hartree-Fock level of theory and then the RESP charges were obtained by a two-stages fitting procedure (7). The Gasteiger-Marsili partial charges (8) were assigned to the protein atoms except for the zinc ion and its coordinating residues (zinc subsite). These are represented by His12, His15, His211, Asp292 and a water molecule when the ADA open form is considered and His12, His15, His211, Asp292, and HDPR when the closed form is used. For these atoms the RESP charges were obtained using the unrestricted density functional UB3LYP. For the optimization and the subsequent electrostatic potential (ESP) calculation, the following locally dense basis set was used: 6-31G* was assigned to the Zn coordinating side chains, except those atoms directly involved in the zinc coordination. In fact, to better describe charge transfer phenomena, the basis set on these atoms was augmented to 6–31+G*. Conversely, to limit the amount of computational time, the other atoms of the coordinating residues, as well as the acetyl (ACE) and N-methyl (NME) capping groups, were treated with the smaller 3–21G* basis set. The Stuttgart-Dresden ECP (SDD) was used for zinc ion.

The octahedral coordination geometry of the Zn ion for ab initio calculations was obtained from crystallographic templates (PDB ID code 2z7g and 1krm) and the zinc was considered in the +2 oxidation state with low spin state ($S = 0$).

Hence, the system was geometrically optimized constraining the backbone atoms, and using the previously reported locally dense basis set. The model system was then optimized till reaching the default Gaussian03 convergence criterion.

The RESP charges were then calculated on the geometrically minimized Zn subsite. First the ESP were calculated again by means of the Gaussian package (6) and then the RESP charges were obtained by a two-stages fitting procedure (7), fitting first the polar areas by using weak hyperbolic restraints (0.0005 a.u.), and then fitting the remaining areas imposing equivalencies and by using a stronger hyperbolic restraint (0.001 a.u.). In each step, the charges of the standard residues ACE and NME were constrained to their AMBER force field value (9). The electrostatic potential, used as input by the program to compute the RESP charges, was sampled according to the Merz-Singh-Kollman scheme (10, 11), namely using 10 concentric layers at the default level of spacing, a surface density of 6 points/$Å^2$, and adopting the covalent radius of 0.88 Å for zinc as reported on the WebElements server (Winter M. WebElements™ Periodic Table. United Kingdom: University of Sheffield. http://www.webelements.com/). Hence, the computed charges for ligands and Zn subsite were used for further computational calculations.

Docking setup. The docking area has been defined by a box, centered on the cartesian coordinates of the zinc ion. Grids points of $60 \times 60 \times 60$ with 0.375 Å spacing were calculated around the docking area for all the ligand atom types using AutoGrid4. For each ligand, 100 separate docking calculations were performed. Otherwise default docking parameters were applied. The docking conformations were clustered on the basis of rmsd (rmsd = 1.5 Å) between the cartesian coordinates of the ligand atoms and were ranked based on the AutoDock scoring function.

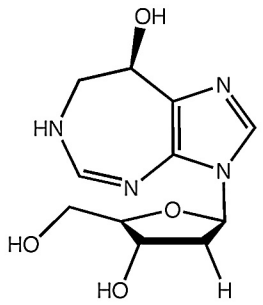
The interfacial water CV. In order to assess the contribution of the water bridge interaction between compound **3b** and Asp19 in the binding process, we have reweighted (12) the free energy along the distance and the interfacial water CV (14). This latter CV is defined as:

$$s_{\mathrm{IntWat}} = \sum_{i}^{n_0} \left(\frac{1 - (\frac{|\mathbf{r}_i - \mathbf{r}_1|}{r_0})^n}{1 - (\frac{|\mathbf{r}_i - \mathbf{r}_1|}{r_0})^m} \right) \left(\frac{1 - (\frac{|\mathbf{r}_i - \mathbf{r}_2|}{r_0})^n}{1 - (\frac{|\mathbf{r}_i - \mathbf{r}_2|}{r_0})^m} \right),$$

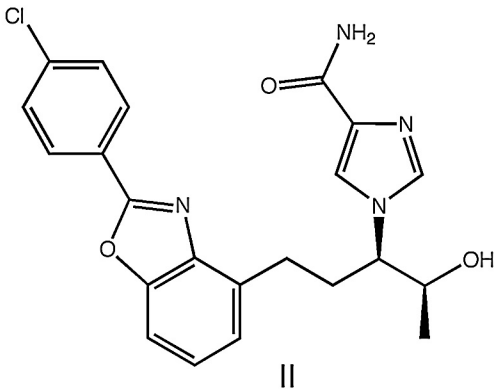
where $|r_i - r_1|$ and $|r_i - r_2|$ are the distances between the n_0 waters oxygens and the oxygen atom of the pyrazolo[1,5-*a*]pyrimidin-7-one ring and one of the carboxylic oxygens of Asp19, respectively, and $r_0 = 4$, $n = 6$ and $m = 12$ are the values used in the switching function to simulate H-bond interaction.

The VMD program (15) was used for visualization and data analysis while the figures were made using the PyMOL software (16).

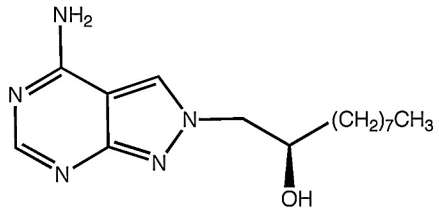
1. Morris G-M, et al. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19:1639–1662.
2. Huey R, Morris G-M, Olson A-J, Goodsell D-S (2007) A semiempirical free energy force field with charge-based desolvation. *J Comput Chem* 28:1145–1152.
3. Kinoshita T, Nishio N, Nakanishi I, Sato A, Fujii T (2003) Structure of bovine Adenosine Deaminase complexed with 6-hydroxy-1,6-dihydropurine riboside. *Acta Crystallogr D Biol Crystallogr* 59:299–303.
4. Kinoshita T, Tada T, Nakanishi I (2008) Conformational change of Adenosine Deaminase during ligand-exchange in a crystal. *Biochem Biophys Res Commun* 373:53–57.
5. Bayly C-I, Cieplak P, Cornell W-D, Kollman P-A (1993) A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: the RESP model. *J Phys Chem* 97:10269–10280.
6. Frisch M-J, et al. (2004) *Gaussian 03*, Revision C.02, (Gaussian, Inc., Wallingford CT).
7. Cornell W-D, Cieplak P, Bayly C-I, Kollman P-A (1993) Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J Am Chem Soc* 115:9620–9631.
8. Gasteiger J, Marsili M (1980) Iterative partial equalization of orbital electronegativity —a rapid access to atomic charges. *Tetrahedron* 36:3219–3228.
9. Cornell W-D, et al. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 117:5179–5197.
10. Singh U-C, Kollman P-A (1983) An approach to computing electrostatic charges for molecules. *J Comput Chem* 5:129–145.
11. Besler B-H, Merz K-M, Kollman P-A (1990) Atomic charges derived from semiempirical methods. *J Comput Chem* 11:431–439.
12. Bonomi M, Barducci A, Parrinello M (2009) Reconstructing the equilibrium Boltzmann distribution from well-tempered metadynamics. *J Comput Chem* 30:1615–1621.
13. Da Settimo F, et al. (2005) Novel, highly potent Adenosine Deaminase inhibitors containing the pirazolo[3,4-d]pyrimidine ring system. Synthesis, structure-activity relationships, and molecular modeling studies. *J Med Chem* 48:5162–5174.
14. Pietrucci F, Marinelli F, Carloni P, Laio A (2009) Substrate binding mechanism of HIV-1 protease from explicit-solvent atomisitc simulations. *J Am Chem Soc* 131:11811–11818.
15. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graphics* 14:33–38.
16. DeLano W-L (2002) The PyMOL molecular graphics system. (DeLano Scientific, San Carlos, CA USA), http://www.pymol.org.



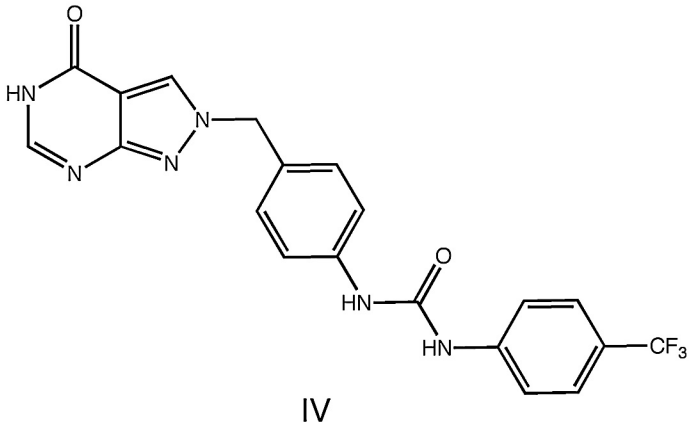
I



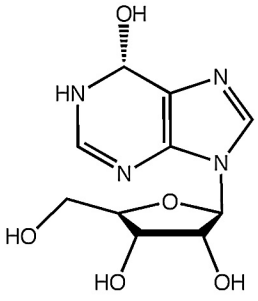
II



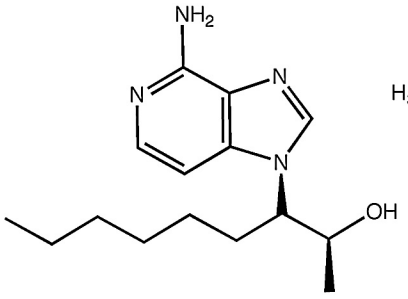
III



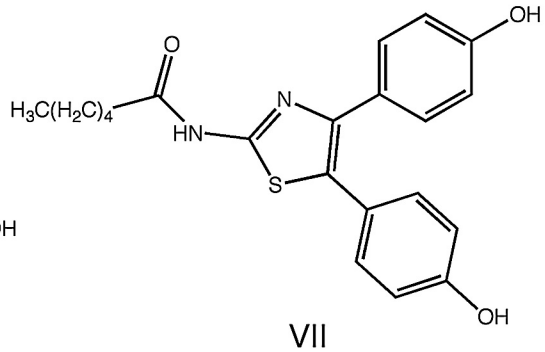
IV



V



VI



VII

Fig. S1. Some of the most potent ADA inhibitors: (**I**) deoxycoformycin (dCF or Pentostatin), (**II**) 1-((3*R*, 4*S*)-1-(2-(4-chlorophenyl)benzo[*d*]oxazol-4-yl)-4-hydroxypentan-3-yl)-1*H*-imidazole-4-carboxamide, (**III**) (*R*)-4-amino-2-(2-hydroxy-1-decyl)pyrazolo[3,4-*d*]pyrimidine, (**IV**) 1-(4-((4-oxo-4,5-dihydropyrazolo[3,4-d]pyrimidin-2-yl)methyl)phenyl)-3-(4-(trifluoromethyl)phenyl)urea, (**V**) (*R*)-6-hydroxy-1,6-dihydropurine ribonucleoside (HDPR), (**VI**) erythro-9-(2-hydroxy-3-nonyl) adenine ((+)-EHNA), (**VII**) *N*-(4,5-bis(4-hydroxyphenyl)thiazol-2-yl)hexanamide (FR104783).

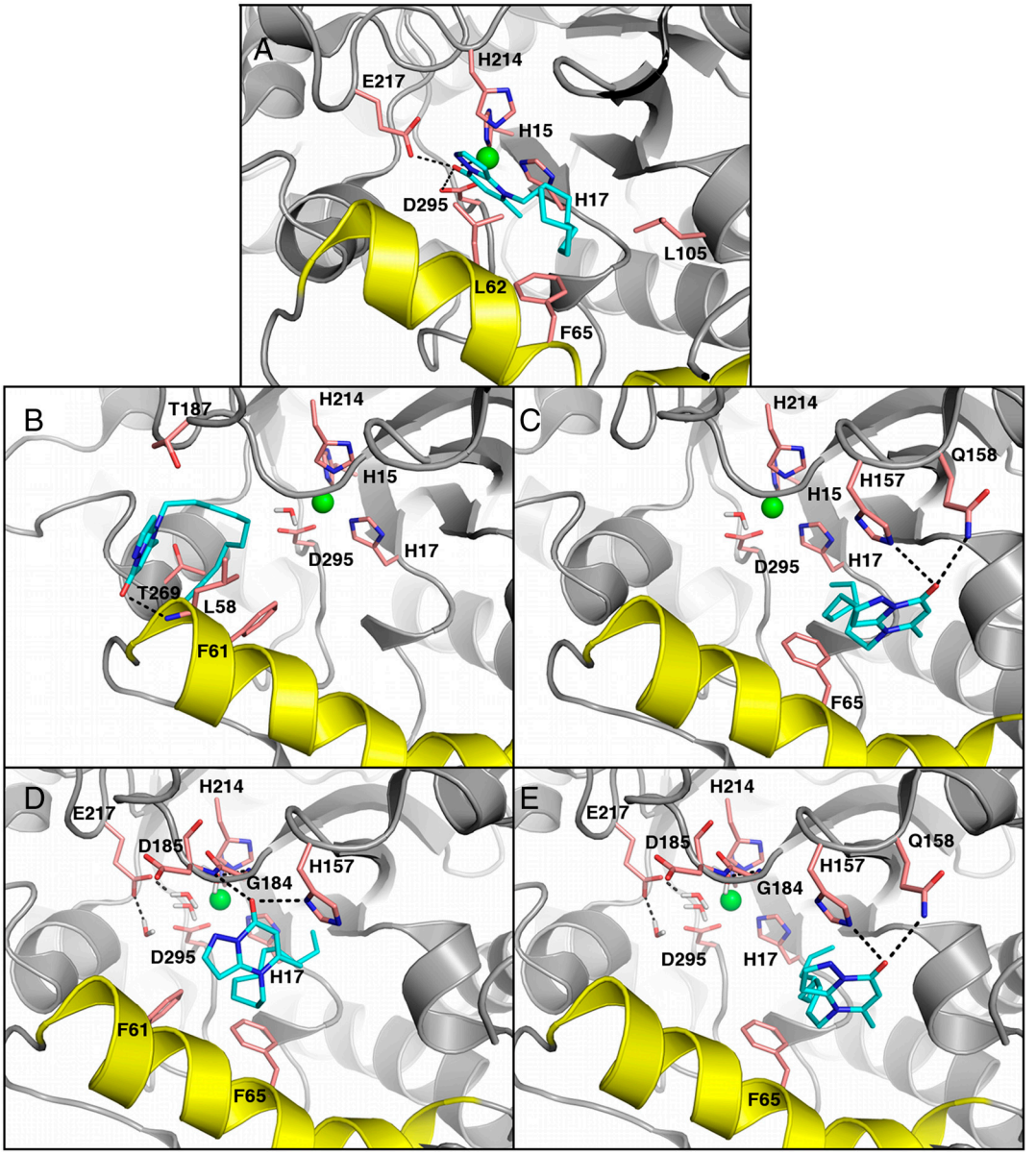


Fig. S2. (*A*) Representation of the **3b** binding conformation found by AutoDock in the closed form of ADA. (*B*) and (*C*) Representation of the **3b** binding conformations found by AutoDock in the ADA open form when a water molecule has been considered as the fifth coordinating group of the zinc ion. (*B*) Shows the ligand conformation representing the best scored cluster while (*C*) that representing the most populated cluster. (*D*) and (*E*) Representation of the **3b** binding conformations found by AutoDock in the ADA open form when the zinc coordinating water and three additional water molecules are considered. As reported in the main text the latter are involved in H-bond bridge interactions between EHNA and ADA (PDB ID code 2z7g). (*D*) Shows the ligand conformation representing the best scored cluster while (*E*) that representing the most populated cluster. The ligand and the main interacting residues are displayed as licorice while the protein is represented as cartoon with the H_3 α-helices colored in yellow. Only heavy atoms are displayed for clarity.

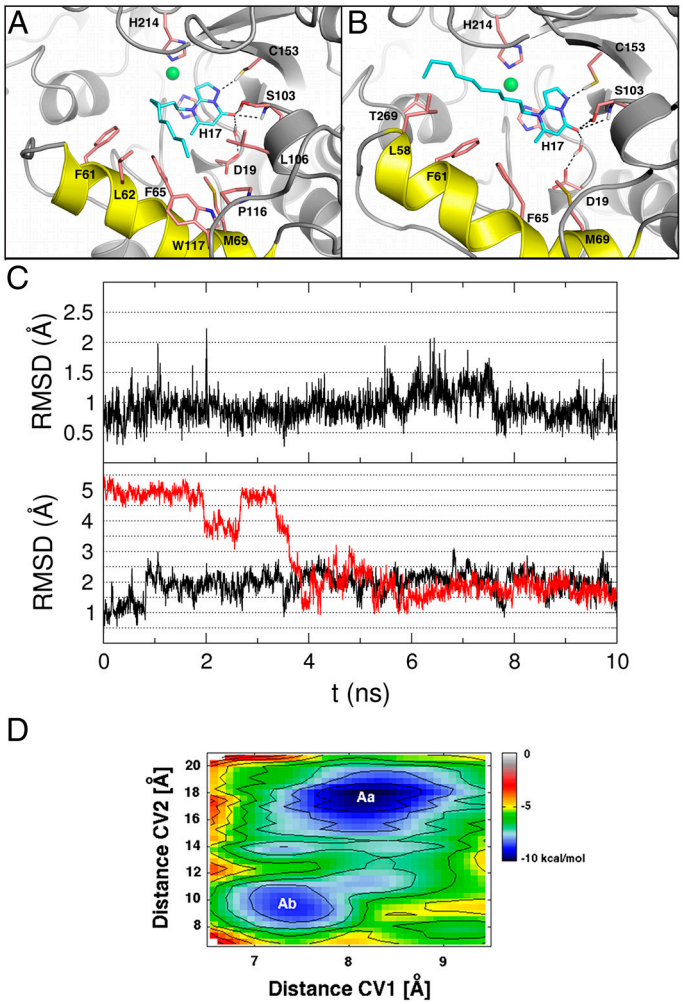


Fig. S3. Representation of the binding mode of the two ligand tail conformation families found by metadynamics. (*A*) Shows that with the *n*-decyl chain interacting with the residues present at the bottom of the H3 α-helix (binding conformation Aa) while (*B*) that with the alkyl tail occupying the upper part of the H3 α-helix (binding conformation Ab). The ligand and the main interacting residues are displayed as licorice while the protein is represented as cartoon with the H_3 α-helix colored in yellow. The ligand/Asp19 water bridge interaction is also shown while only polar hydrogens interacting with the ligand are displayed for clarity. (*C*) (Top graph) Plot of the rmsd of the heavy atoms of the pyrazolo[1,5-*a*]pyrimidin-7-one ring during over 10 ns of MD simulations with the ligand in the Aa binding conformation. The very low average rmsd value of 0.93 Å for the pyrazolo[1,5-*a*]pyrimidin-7-one ring reflects the good stability of this pose. (Lower graph) Plot of the rmsd of the heavy atoms of the ligand in the binding conformation Aa (black lines) and Ab (red lines) during over 10 ns of standard MD. The rmsd values are calculated relatively to the ligand coordinates in the Aa conformation at the beginning of the simulation. The red plots clearly show that the Ab conformation changes in Aa after approximately 3.5 ns. This change is remarked also by the minimum rmsd value of 0.93 Å calculated for Ab relative to Aa. (*D*) Representation of the FES of the ligand conformations in ADA as function of the new distance CV (CV2) and the distance CV used in the original metadynamics simulations (CV1). Isosurfaces of 1 kcal/mol is used. Looking at the FES two minima are found corresponding to the two ligand tail conformations Aa and Ab. The Aa basin is wider and 1.3 kcal/mol deeper than that of Ab with an energetic barrier Aa → Ab of approximately 3.5 kcal/mol between the two minima. These results are in line with the MD study that suggests a greater stability of the Aa conformation (Fig. S3*C*).

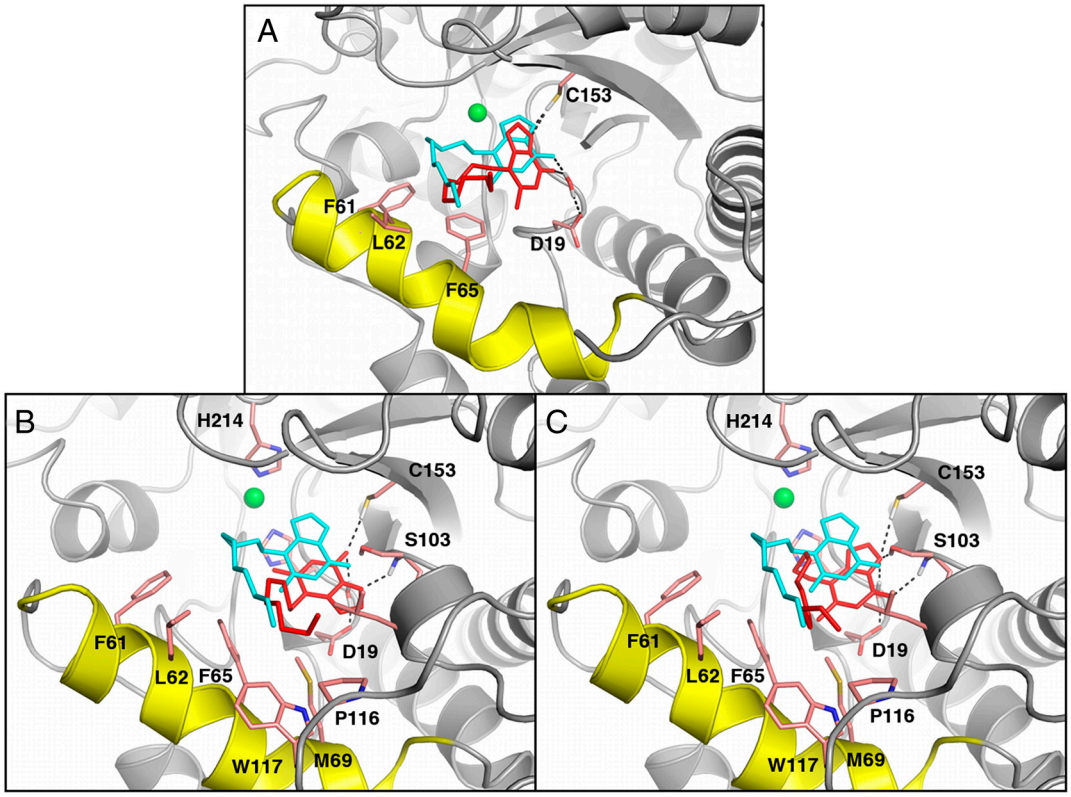


Fig. S4. (*A*) Overlapping in the ADA binding site of the **3b** binding conformations predicted by metadynamics (cyan) and by AutoDock (red). Using the ADA conformation ligated to EHNA (PDB ID code 2z7g) this docking trial has been executed including the water molecule that mediates the interaction between the ligand and Asp19 with the latter optimally oriented for such interaction. (*B*) and (*C*) Comparison of the biding mode of **3b** (cyan) with those calculated by AutoDock for **3a** (b) and **3c** (c). (*B*) If compared with **3b**, **3a** showed the pyrazolo[1,5-*a*]pyrimidin-7-one ring translated and rotated by 180° with the O7 atom involved in a water bridge interaction with Asp19 and a H-bond with Cys153. Due to the ring rotation the hydrophobic interactions between the methyl group of the pyrimidinone moiety and Phe65 and Met69 are lost as well as the stacking interactions with His17 and His214. On the other hand, the new position of the pyrazolo[1,5-*a*]pyrimidin-7-one ring in the active site allows the shorter n-nonyl chain to fill the F0 site in a way similar to the n-decyl chain in **3b**. (*C*) In contrast, **3c** presents a binding mode very similar to that of **3b** with the pyrazolo[1,5-*a*]pyrimidin-7-one ring slightly displaced from its usual position but still able to form all the main interactions with the protein. The ligand and the main interacting residues are displayed as licorice while the protein is represented as cartoon with the H_3 α-helix colored in yellow. The ligand/Asp19 water bridge interaction is also shown while only polar hydrogens interacting with the ligand are displayed for clarity.

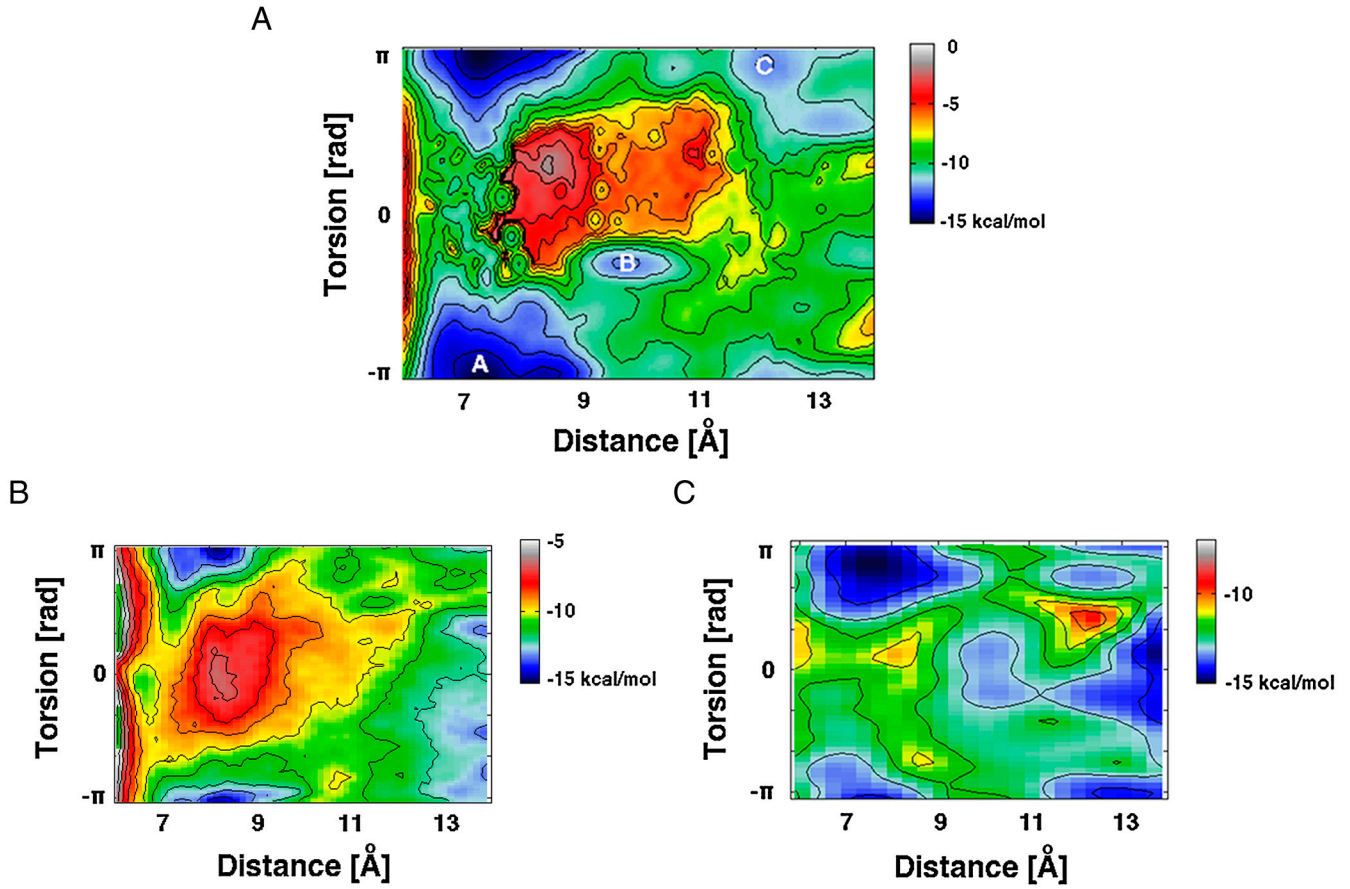


Fig. S5. Comparison between the FES graphs obtained using different CVs settings shown using isosurfaces of 1 kcal/mol. Å. (*A*) FES obtained using the distance and the PCV and reweighted on the torsion CV, (*B*) FES obtained using the distance, and torsion CV without using the path CV and (*C*) FES obtained using a different distance CV and the PCV. In order to facilitate the comparison, in the latter case the FES is recalculated along the distance and torsion CVs used in the original metadynamics. The basin **A** is constantly found through all the metadynamics simulations although this basin is less deep and slightly shifted along the distance CV when no PCV is used (see graph B). Furthermore, basin **B** and **C** have not been found in the simulations without the PCV while they reappear if a distance CV and the PCV are used. These results highlight the importance of simulating the ADA flexibility to properly define the interactions engaged by the ligand with the protein along its binding path.

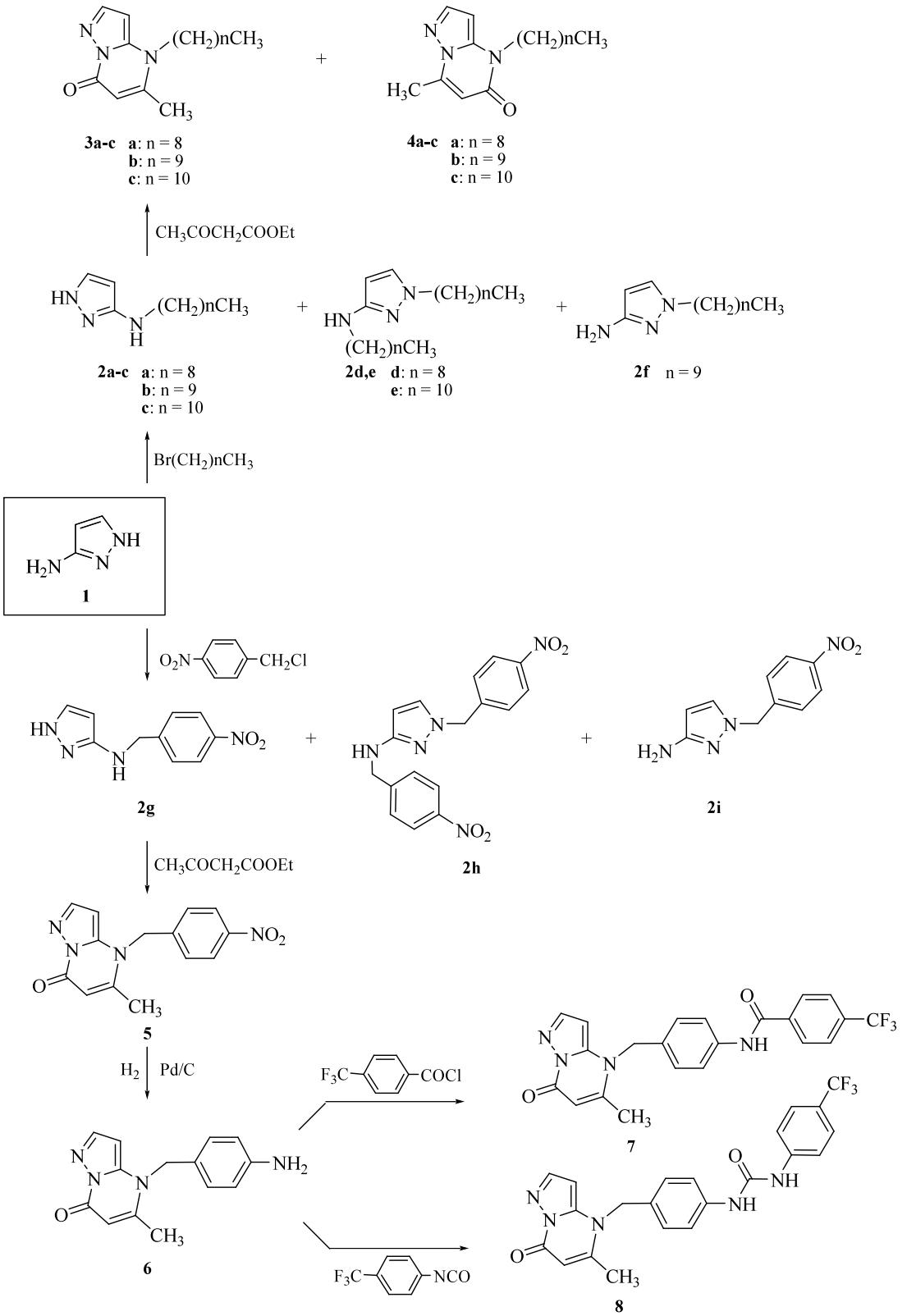
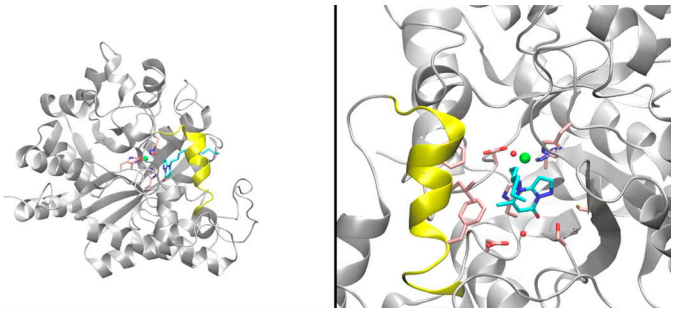


Fig. S6. Synthesis of Pyrazolo[1,5-*a*]pyrimidinone Derivatives.



Movie S1. The movie shows the exploration of the ADA binding site by the inhibitor **3b** simulated through metadynamics calculations.

Movie S1 (MOV)

Table S1. List of atoms that define the different CVs used in the metadynamics simulations

List of atoms used to define the distance and torsion CV. Distance CV is defined as the distance between the center of mass of groups of atoms.

CV type	Atom	Group
Distance	CMR *	1
Distance	Val16Cα	2
Distance	His17Cα	2
Distance	Val100Cα	2
Distance	Arg101Cα	2
Torsion	Val16Cα	-
Torsion	Arg101Cα	-
Torsion	C1′ of the alkyl tail(attached to N4 of the pyrazolopyrimidinone ring)	-
Torsion	N1 of the pyrazolopyrimidinone ring	-

List of atoms used to define the distance CV used in the metadynamics simulations to assess the stability of Aa and Ab conformations. Distance CV is defined as the distance between the center of mass of groups of atoms.

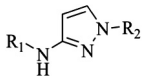
CV type	Atom	Group
Distance	C7′ of the alkyl tail	1
Distance	C8′ of the alkyl tail	1
Distance	C9′ of the alkyl tail	1
Distance	Ser265Cα	2
Distance	Leu268Cα	2

List of atoms used to define the distance CV used in the extra metadynamics simulations. Distance CV is defined as the distance between the center of mass of groups of atoms.

CV type	Atom	Group
Distance	CMR *	1
Distance	Val16Cα	2
Distance	Asp295Cα	2

*CMR is the center of the mass of the ligand pyrazolo[1,5-*a*]pyrimidin-7-one ring

Table S2. Physical properties and spectral data of compounds 2a–i

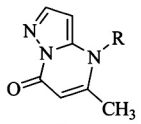


No.	R_1	R_2	Yield (%)	mp (°C)	Recyst. solv. *	Formula†	IR (ν, cm^{-1})	^{1}H-NMR (δ, ppm)
2a	$(CH_2)_8CH_3$	H	71	oil		$C_{12}H_{23}N_3$	3,302; 3,216; 1,664; 1,557	0.86 (t, 3H, CH_3); 1.15–1.32 (m, 12H, CH_2); 1.47 (t, 2H, CH_2); 2.96 (m, 2H, CH_2); 4.99 (bs, 1H, NH, exc.); 5.39 (d, 1H, H_4); 7.28 (d, 1H, H_5); 11.44 (bs, 1H, NH, exc.)
2b	$(CH_2)_9CH_3$	H	62	28–30	EtOH	$C_{13}H_{25}N_3$	3,313; 3,190; 1,552; 1,495	0.86 (t, 3H, CH_3); 1.13–1.30 (m, 14H, CH_2); 1.45 (t, 2H, CH_2); 2.93 (m, 2H, CH_2); 4.99 (bs, 1H, NH, exc.); 5.39 (d, 1H, H_4); 7.28 (d, 1H, H_5); 11.42 (bs, 1H, NH, exc.)
2c	$(CH_2)_{10}CH_3$	H	63	32–34	EtOH	$C_{14}H_{27}N_3$	3,298; 3,206; 1,557; 1,495	0.85 (t, 3H, CH_3); 1.14–1.31 (m, 16H, CH_2); 1.44 (t, 2H, CH_2); 2.93 (m, 2H, CH_2); 4.96 (bs, 1H, NH, exc.); 5.38 (d, 1H, H_4); 7.25 (d, 1H, H_5); 11.43 (bs, 1H, NH, exc.)
2d	$(CH_2)_8CH_3$	$(CH_2)_8CH_3$	15	oil		$C_{21}H_{41}N_3$	3,190; 1,552; 1,526; 1,495	0.85 (t, 6H, CH_3); 1.11–1.38 (m, 24H, CH_2); 1.43 (t, 2H, CH_2); 1.61 (t, 2H, CH_2); 3.06 (m, 2H, CH_2); 3.81 (m, 2H, CH_2); 4.99 (t, 1H, NH, exc.); 5.44 (d, 1H, H_4); 7.28 (d, 1H, N_5)
2e	$(CH_2)_{10}CH_3$	$(CH_2)_{10}CH_3$	21	oil		$C_{25}H_{49}N_3$	3,262; 1,577; 1,557; 1,495	0.85 (t, 6H, CH_3); 1.10–1.36 (m, 32H, CH_2); 1.43 (t, 2H, CH_2); 1.641 (t, 2H, CH_2); 2.92 (m, 2H, CH_2); 3.80 (m, 2H, CH_2); 4.91 (t, 1H, NH, exc.); 5.35 (d, 1H, H_4); 7.29 (d, 1H, N_5)
2f	H	$(CH_2)_9CH_3$	22	oil		$C_{13}H_{25}N_3$	3,329; 3,196; 1,716; 1,634	0.85 (t, 3H, CH_3); 1.12–1.30 (m, 14H, CH_2); 1.64 (t, 2H, CH_2); 3.79 (m, 2H, CH_2); 5.08 (bs, 2H, NH_2, exc.); 5.22 (d, 1H, H_4); 6.99 (d, 1H, H_5)
2g	$CH_2C_6H_4pNO_2$	H	62	118–119	EtOH	$C_{10}H_{10}N_4O_2$	3,443; 3,250; 1,591	4.35 (d, 2H, CH_2); 5.45 (d, 1H, H_4); 5.92 (t, 1H, NH, exc.); 7.31 (d, 1H, H_5); 7.60 (d, 2H, ArH); 8.16 (d, 2H, ArH); 11.49 (bs, 1H, NH, exc.)
2h	$CH_2C_6H_4pNO_2$	$CH_2C_6H_4pNO_2$	21	250–252	EtOH	$C_{17}H_{15}N_5O_4$	3,325; 3,125; 1,649; 1,608	4.36 (d, 2H, CH_2); 5.45 (d, 1H, H_4); 5.34 (s, 2H, CH_2); 5.92 (t, 1H, NH, exc.); 7.32 (d, 1H, H_5); 7.44 (d, 2H, ArH); 7.62 (d, 2H, ArH); 7.81 (d, 2H, ArH); 8.17 (d, 2H, ArH)
2i	H	$CH_2C_6H_4pNO_2$	12	171–173	EtOH	$C_{10}H_{10}N_4O_2$	3,325; 3,125; 1,649; 1,608	5.27 (s, 1H, H_4); 5.34 (s, 2H, CH_2); 5.96 (bs, 2H, NH_2, exc.); 7.31 (s, 1H, H5); 7.60 (d, 2H, ArH); 8.17 (d, 2H, ArH)

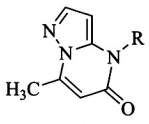
*Compounds *2a*, **d–f** were obtained analytically pure through flash chromatography

†Elemental analyses for C, H, N were within ±0.4% of the calculated values.

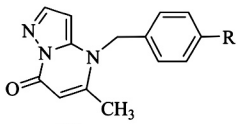
Table S3. Physical properties and spectral data of compounds 3a–c, 4a–c, and 5–8



3



4



5-8

No.	R	Yield (%)	Mp (°C)	Recyst. solv.	Formula	IR (ν, cm-1)	^{1}H-NMR (δ, ppm)
3a	$(CH_2)_8CH_3$	46	63–65	EtOH	$C_{16}H_{25}N_3O$	1,685; 1,659; 1,557	0.85 (t, 3H, CH_3); 1.10–1.29 (m, 12H, CH_2); 1.69 (t, 2H, CH_2); 2.41 (s, 3H, CH_3); 4.02 (t, 2H, CH_2); 5.73 (s, 1H, H_6); 6.40 (d, 1H, H_3); 7.89 (d, 1H, H_2)
3b	$(CH_2)_9CH_3$	45	64–66	EtOH	$C_{17}H_{27}N_3O$	1,690; 1,648; 1,562	0.86 (t, 3H, CH_3); 1.11–1.29 (m, 14H, CH_2); 1.69 (t, 2H, CH_2); 2.42 (s, 3H, CH_3); 4.02 (t, 2H, CH_2); 5.74 (s, 1H, H_6); 6.40 (d, 1H, H_3); 7.89 (d, 1H, H_2)
3c	$(CH_2)_{10}CH_3$	49	76–78	EtOH	$C_{18}H_{29}N_3O$	1,690; 1,659; 1,552	0.86 (t, 3H, CH_3); 1.10–1.30 (m, 16H, CH_2); 1.68 (t, 2H, CH_2); 2.41 (s, 3H, CH_3); 4.02 (t, 2H, CH_2); 5.73 (s, 1H, H_6); 6.40 (d, 1H, H_3); 7.89 (d, 1H, H_2)
4a	$(CH_2)_8CH_3$	35	43–45	EtOH	$C_{16}H_{25}N_3O$	1,680; 1,562	0.85 (t, 3H, CH_3); 1.09–1.27 (m, 12H, CH_2); 1.70 (t, 2H, CH_2); 2.48 (s, 3H, CH_3); 3.91 (t, 2H, CH_2); 6.01 (s, 1H, H_6); 6.22 (d, 1H, H_3); 7.84 (d, 1H, H_2)
4b	$(CH_2)_9CH_3$	35	45–47	EtOH	$C_{17}H_{27}N_3O$	1,685; 1,560	0.85 (t, 3H, CH_3); 1.11-1.29 (m, 14H, CH_2); 1.69 (t, 2H, CH_2); 2.49 (s, 3H, CH_3); 3.94 (t, 2H, CH_2); 6.01 (s, 1H, H_6); 6.22 (d, 1H, H_3); 7.84 (d, 1H, H_2)
4c	$(CH_2)_{10}CH_3$	30	54–56	EtOH	$C_{18}H_{29}N_3O$	1,682; 1,563	0.85 (t, 3H, CH_3); 1.10–1.28 (m, 16H, CH_2); 1.69 (t, 2H, CH_2); 2.49 (s, 3H, CH_3); 3.91 (t, 2H, CH_2); 6.01 (s, 1H, H_6); 6.22 (d, 1H, H_3); 7.84 (d, 1H, H_2)
5	NO_2	82	275 dec.	EtOH	$C_{14}H_{12}N_4O_3$	1,741; 1,726; 1,691	2.33 (s, 3H, CH_3); 5.53 (s, 2H, CH_2); 5.86 (s, 1H, H_6); 6.30 (d, 1H, H_3); 7.42 (d, 2H, ArH); 7.85 (d, 1H, H_2); 8.21 (d, 2H, ArH)
6	NH_2	92	172–173	EtOH	$C_{14}H_{14}N_4O$	3,414; 3,342; 1,673	2.38 (s, 3H, CH_3); 5.14 (bs, 4H, CH_2, NH_2); 5.78 (s, 1H, H_6); 6.37 (d, 1H, H_3); 6.52 (d, 2H, ArH); 6.86 (d, 2H, ArH); 7.86 (d, 1H, H_2)
7	$NHCOC_6H_4$-*p*-CF_3	84	267–269	MeOH	$C_{21}H_{18}N_4O_2$	3,296; 1,675; 1,666	2.38 (s, 3H, CH_3); 5.34 (s, 2H, CH_2); 5.82 (s, 1H, H_6); 6.35 (d, 1H, H_3); 7.16 (d, 2H, ArH); 7.53 (d, 2H, ArH); 7.76 (d, 2H, ArH); 7.86 (d, 1H, H_2); 7.93 (d, 2H, ArH); 10.30 (s, 1H, NH, exc.)
8	$NHCONHC_6H_4$-*p*-CF_3	75	272–275	MeOH	$C_{21}H_{19}N_5O_2$	3,330; 3,283; 1,665	2.38 (s, 3H, CH_3); 5.30 (s, 2H, CH_2); 5.81 (s, 1H, H_6); 6.36 (d, 1H, H_3); 6.97 (d, 2H, ArH); 7.09 (d, 2H, ArH); 7.27 (d, 2H, ArH); 7.43 (d, 2H, ArH); 7.86 (d, 2H, H_2); 8.64 (s, 1H, NH, exc.); 8.72 (s, 1H, NH, exc.)

Table S4. Analytical data of compounds 2a–i, 3a–c, 4a–c, and 5–8

No.	Formula	Calcd. %			Found %		
		C	H	N	C	H	N
2a	$C_{12}H_{23}N_3$	68.85	11.07	20.07	68.62	11.15	20.13
2b	$C_{13}H_{25}N_3$	69.91	11.28	18.81	69.80	11.40	18.64
2c	$C_{14}H_{27}N_3$	70.83	11.46	17.70	70.89	11.63	19.58
2d	$C_{21}H_{41}N_3$	75.16	12.32	12.52	75.06	12.28	12.37
2e	$C_{25}H_{49}N_3$	76.66	12.61	10.73	76.67	12.71	10.64
2f	$C_{13}H_{25}N_3$	69.91	11.28	18.81	69.90	11.40	18.70
2g	$C_{10}H_{10}N_4O_2$	55.04	4.62	25.68	55.18	4.69	25.73
2h	$C_{17}H_{15}N_5O_4$	57.79	4.28	19.82	57.91	4.33	19.95
2i	$C_{10}H_{10}N_4O_2$	55.04	4.62	25.68	55.28	4.73	25.61
3a	$C_{16}H_{25}N_3O$	69.78	9.15	15.26	69.58	9.11	15.36
3b	$C_{17}H_{27}N_3O$	70.55	9.40	14.52	70.65	9.54	14.61
3c	$C_{18}H_{29}N_3O$	71.25	9.63	13.85	71.29	9.74	13.98
4a	$C_{16}H_{25}N_3O$	69.78	9.15	15.26	69.77	9.14	15.20
4b	$C_{17}H_{27}N_3O$	70.55	9.40	14.52	70.50	9.36	14.65
4c	$C_{18}H_{29}N_3O$	71.25	9.63	13.85	71.27	9.59	13.81
5	$C_{14}H_{12}N_4O_3$	59.15	4.25	19.71	59.23	4.38	19.84
6	$C_{14}H_{14}N_4O$	66.13	5.55	22.03	66.26	5.60	21.98
7	$C_{21}H_{18}N_4O_2$	70.38	5.06	15.63	70.28	4.96	15.51
8	$C_{21}H_{19}N_5O_2$	67.55	5.13	18.76	67.52	5.10	18.61