Supplementary Information for:

The Conformational Analysis and Proton Transfer of the Neuraminidase Inhibitors: A Theoretical Study

Zhiwei Yang,[a] Gang Yang,[a] Yuangang Zu,[a] Yujie Fu,[a] Lijun Zhou[a, b]

[a] *Key Laboratory of Forest Plant Ecology, Ministry of Education, Northeast Forestry University, Harbin 150040, P. R. China*

[b] *Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China*

1. Docking and molecular dynamics

1.1. System preparations

The analysis of the influenza virus neuraminidases (NA) indicates the active-site residues are well conserved in all NA sub-types [1-5]. The N9-subtype NA crystal structure complexed with 2, 3-didehydro-2-deoxy-*N*-acetylneuraminic acid (PDB entry: 1F8B) was obtained at 1.8 Å high resolution [6]. The N9-subtype NA structure was used as in the previous works [7, 8]. For convenience, it is named as NA throughout this work.

As shown in Figure S1, all the hetero-atoms, except the Ca^{2+} ion and two crystal water molecules near the active site, were removed from the NA crystal structure. With the aid of the Biopolymer module and manual verification [6, 9-12], the hydrogen atoms were added to the protein systems. The protonation states for the histidine residues and other titratable groups were determined at physiological pH and manually verified. The side chain of Asn294 in NA was rotated so that its Oδ1 and Nδ2 atoms in the amide group form hydrogen bonds instead of repulsive interactions with the nearby Ala246 O and Arg292 Nε2 atoms. This correction will lead to a lower root mean square deviation and an improved agreement with the overall crystal structure [11]. The NA structure was then neutralized with chloride anions.

The docking and molecular dynamics (MD) simulations were carried out with the InsightII 2005 software packages [12], using the consistent-valence force-field (CVFF). As the previous literatures indicated, the CVFF force-field obtained comparable results with other force-fields and was widely used in the ligand-protein interacting systems [11, 13-23], including the NA-related ones [11]. The conjugated gradient algorithm was used to optimize the NA structure (Discover 3.0 module). The convergence criterion was set to 0.01 kcal mol^{-1} $Å^{-1}$.

1.2. Docking

The Binding-site module is a suite of programs capable of identifying and characterizing the protein active sites, binding sites and functional residues. The NA active site was identified with this module. The docking experiments between the NA proteins and inhibitors were performed with the advanced docking program Affinity, using a combination of Monte Carlo (MC) type and simulation annealing (SA) method [24]. Note that the neutral and zwitterionic isomers of BA were docked respectively into the NA active sites. The potential function was assigned using the CVFF force-field and the non-bonded interactions were described by the Cell-Multipole approach. The NA-BA complexes were solvated in a large sphere of TIP3P water molecules [25] with the radius of 35.0 Å. Chloride anions were added to neutralize the docked systems [7-9, 11]. The docked complexes were selected on the basis of interaction energies and geometrical matching qualities, which were further energy-minimized using the conjugated gradient method. The convergence criterion was set to 0.01 kcal mol^{-1} $Å^{-1}$.

1.3. Molecular dynamics (MD)

The MD simulations were performed on the energy-minimized nBA-NA and zBA-NA complexes, using the CVFF force-field in Discover 3.0 module. The

canonical ensemble (NVT) was used and the non-bonded interactions were described with the Cell-Multipole approach. The simulation temperatures are 300.0 K (normal temperature), which was controlled by the Langevin thermostat [26]. The integration of the classical equations of motion was achieved using the Verlet algorithm. During the MD simulations, the regions within a sphere of 10.0 Å around the inhibitors were allowed to move freely whereas the rests were held rigid. An integration step of 1.0 fs was used and every MD simulation was run for 5.0 ns [11, 27, 28]. The snapshots of the MD trajectories were saved at 1.0-ps intervals.

1.4. Results and discussion

The backbone-atom root-mean-square deviations (RMSD) in Figure S2 and the total energies in Figure S3 indicated that all of the systems are stable during the 5.0 ns molecular dynamics (MD) trajectories, since approximately 0.5 ns [11, 27, 28]. The nBA-NA and zBA-BA complexes were analyzed as the averages of the last 4.5 ns (0.5ns~5.0 ns). In some cases, the average data of the 0.5ns~1.0 ns periods were also given and confirmed the equilibrium since about 0.5 ns.

It was found from Figures S4 and S5 that the orientations of the neutral and zwitterionic isomers of BA are quite different at the NA active sites. nBA is characterized by its carboxyl group towards Gly348 and Tyr406 as well as the polar contacts of its guanidine group with Glu276. The distances are approximated at 2.62 (2.61) Å for the nBA:O1 and Gly348:O pair, 2.85 (2.86) Å for the nBA:O2 and Tyr406:OH pair, respectively. Note that the values in parentheses were taken from the beginning 1.0 ns MD trajectories. The N-acetyl and 3-pentyloxyl groups of nBA are partially outside of the NA active site. Two hydrogen bonds are formed between nBA and NA (Figure S5a and Table S1). The carboxyl group is involved in both hydrogen bonds, an indication of its important role during the binding process.

In zBA-NA, the carboxyl group of zBA is in the deprotonated form and redirects towards Arg371 and Lys432 (Figures S4b and S5b). The salt-bridge interactions are formed between the carboxylate anion (COO^-) of zBA and the positively charged pocket consisting of Arg118, Arg292 and Arg371, in agreement with the previous docking results [7, 10, 11, 29-32]. As a result of this strong interactions, three stable hydrogen bonds of O1-Asn347:ND2, O1-Arg371:NH2 and O2-Lys432:NZ are formed with their distances approximating 2.11 (2.11), 1.73 (1.79) and 2.47 (2.63) Å, respectively (Figure S5b and Table S1). It distinguishes from the binding mode of nBA whose carboxyl group is far away from the positively charged pocket consisting of three arginine residues. The guanidine group of zBA is in the protonated form and oriented towards the negatively charged pocket consisting of Asp151 and Glu119. The N-acetyl group sits at the entrance of the hydrophobic pocket containing Trp178, Ile222 and Arg152. In addition, the hydrophobic 3-pentyloxyl group matches the hydrophobic pocket with Glu276 and Glu277, consistent with the previous results [7, 10, 11, 29-32].

The NA active-site residues that have interaction energies below -1.0 kcal mol^{-1} were colleted in Table S2. It was found that Ile149, Asp151 and Arg371 have no interactions with nBA whereas have electrostatic interactions with zBA, especially Arg371 where the electrostatic interaction energy is calculated to be -51.20 (-48.19) kcal mol^{-1} (Table S2). Note that the values in parentheses were taken from the beginning 1.0 ns MD trajectories. The vdW interactions of nBA and the electrostatic interactions of zBA contribute to 66.12% (66.36%) and 89.52% (88.59%). Accordingly, the vdW interactions play a larger role in the binding process of the neutral BA isomer whereas the electrostatic interactions play a dominant in the binding process of the zwitterionic BA isomer. The zwitterionic rather than neutral

isomer of BA matches finely with the NA active sites and therefore is responsible for the bioactivity [5, 10, 11, 29-32]. From the above geometric and energy analysis it was also found that the docked complexes have been sufficiently equilibrated within 1.0 ns MD simulations [11].

1.5. Summary

Docking and MD simulations were used to study the binding modes of the neutral and zwitterionic isomers of BA at the NA active sites. It was found from the geometric and energy analysis that the docked complexes have been equilibrated since about 0.5 ns of the MD simulations. The neutral and zwitterionic isomers of BA have quite distinct orientations at the NA active sites. The neutral binding is characterized by its carboxyl group directing towards Gly348 and Tyr406 as well as the polar contacts of its guanidine group with Glu276, while the zwitterionic binding is characterized by the carboxylate anion directing towards the positively charged pocket consisting of Arg118, Arg292 and Arg371. The interaction energies are summed to 45.39 and -160.64 kcal mol^{-1} for nBA and zBA, respectively. The vdW interactions play a larger role in the binding process of nBA whereas the electrostatic interactions play a dominant role in the binding process of zBA. The zwitterionic BA isomer matches finely with the NA active sites and is responsible for the bioactivity.

2. Details of population analysis

The populations of nBA to zBA were calculated with the Boltzmann statistical expressions [33] shown below:

$$P_{nBA} = \frac{1}{1 + \exp(-\Delta E_{rel}/RT)} \quad \text{(S1)}$$

$$P_{zBA} = \frac{\exp(-\Delta E_{Rel}/RT)}{1 + \exp(-\Delta E_{rel}/RT)} \quad \text{(S2)}$$

Where P_{nBA} and P_{zBA} represent the populations of the neutral and zwitterionic BA isomers, respectively. R and T refer to gas constant and temperature, equal to 8.314 J mol^{-1} K^{-1} and 298.15 K (normal temperature), respectively.

On the basis of Eqs. (S1) and (S2), the proportion of nBA to zBA (ρ) can be obtained with

$$\rho = \frac{P_{nBA}}{P_{zBA}} = \exp(\Delta E_{rel}/RT) \quad \text{(S3)}$$

According to the relative energy (ΔE_{rel}) of 24.76 kcal mol^{-1} obtained at B3LYP/6-311++G(d,p)//6-31G(d) level of theory, the proportion of nBA to zBA (ρ) approaches infinity. That is, the BA molecule in the prepared state should exist exclusively in the neutral form.

Using the expressions similar to Eqs. (S1) - (S3), the proportion of nBA(s) to zBA(s) in solution can also be obtained. The relative energy of zBA(s) vs. nBA(s) (E_{rel}) is calculated to be -7.53 kcal mol^{-1} at B3LYP/6-311++G(d,p)//6-31G(d) level of theory, and therefore the proportion of nBA(s) to zBA(s) approximates zero. That is, the BA molecule in aqueous solutions is predominated by the zwitterionic isomer.

3. Intrinsic reaction coordinate calculations

The intrinsic reaction coordinate (IRC) calculated results are shown in Figure S7. The starting point is nBA4. The elongations of the O1-H1 distance and then the O5-H9 distance cause the monotonous increases of the total energy until TS4a, the first transition state. That is, the hydronium ion (H_3O^+) composed of the O5, H1, H8 and H9 atoms is not the transition state but the saddle point. The sequential elongation of the O5-H9 distance causes the total energy to go downhill until the formation of the intermediate IN4 (Figure 3a). The second transition state TS4b is situated with the

slight elongation of the O7-H13 distance (1.132 Å). With the continuing elongation of the O7-H13 distance, the H15 atom will spontaneously transfer to the guanidine N2 atom and directly forms the product zBA4. That is, the hydronium ion containing the O8, H13, H14 and H15 atoms is a saddle point of the proton transfer process, just as the one of the O5, H1, H8 and H9 atoms. It indicates that the two middle water molecules (i.e., H9O6H10 and H11O7H12) rather than the boundary water molecules (i.e, H1O5H8 and H14O8H15) are capable of capturing the protons and form the hydronium ions, thus acting as the transition states.

References

[1] M. von Itzstein, W. Y. Wu, G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. Van Phan, M. L. Smythe, H. F. White and S. W. Oliver, *Nature*, **1993**, *363*, 418-423.

[2] P. Sears and C. H. Wong, *Angew. Chem. Int. Ed.*, **1999**, *38*, 2300-2324.

[3] J. S. Oxford, P. Novelli, A. Sefton and R. Lambkin, *Antivir. Chem. Chemother.*, **2002**, *13*, 205-217.

[4] P. Ward, I. Small, J. Smith, P. Suter and R. Dutkowski, *J. Antimicrob. Chemother.*, **2005**, *55 (Suppl. S1)*, i5-i21.

[5] M. von Itzstein, *Nat. Rev. Drug. Discov.*, **2007**, *6*, 967-974.

[6] B. J. Smith, P. M. Colman, M. Von Itzstein, B. Danylec and J. N. Varghese, *Protein Sci.*, **2001**, *10*, 689-696.

[7] G. Yang, Z. W. Yang, Y. G. Zu, X. M. Wu and Y. J. Fu, *Internet. Electron. J. Mol. Des.*, **2008**, *7*, 97-113.

[8] G. Yang, Z. W. Yang, X. M. Wu and Y. G. Zu, *Chin. Comput. Appl. Chem.*, **2008**, *25*, 409-414.

[9] I. D. Wall, A. R. Leach, D. W. Salt, M. G. Ford and J. W. Essex, *J. Med. Chem.*,

1999, *42*, 5142-5152.

[10] M. Fornabaio, P. Cozzini, A. Mozzarelli, D. J. Abraham and G. E. Kellogg, *J. Med. Chem.*, **2003**, *46*, 4487-4500.

[11] P. Bonnet and R. A. Bryce., *Protein Sci.*, **2004**, *13*, 946-957.

[12] InisghtII, Version **2005**, Accelrys Inc. San Diego, USA.

[13] H.-C. Siebert, M. Gilleron, H. Kaltner, C.-W. von der Lieth, T. Kozár, N. Bovin, E. Y. Korchagina, J. F. G. Vliegenthart and H.-J. Gabius, *Biochem. Biophys. Res. Commun.*, **1996**, *219*, 205-212.

[14] G. D. Szklarz and J. R. Halpert, *Drug. Metab. Dispos.*, **1998**, *26*, 1179-1184.

[15] C. A. Bush, M. Martin-Pastor and A. Imbery, *Annu. Rev. Bioph. Biom.*, **1999**, *28*, 269-293.

[16] V. Frecer, R. Rizzo and S. Miertus, *Biomacromolecules*, **2000**, *1*, 91-99.

[17] B. C. Braden, F. A. Goldbaum, B.-X. Chen, A. N. Kirschner, S. R. Wilson and B. F. Erlanger, *Proc. Natl. Acad. Sci. U. S. A.*, **2000**, *97*, 12193-12197.

[18] M. Fuhrmans, A. G. Milbradt and C. Renner, *J. Chem. Theory Comput.*, **2006**, *2*, 201-208.

[19] Y. H. Zhou, Q. C. Zheng, Z. S. Li, Y. Zhang, M. Sun, C. C. Sun, D. Si, L. Cai, Y. Guo and H. Zhou, *Biochimie*, **2006**, *88*, 1457-1465.

[20] S. Costantini, G. Colonna and A. M. Facchiano, *Comput. Biol. Chem.*, **2007**, *31*, 196-206.

[21] A. A. Rodríguez-Menchaca, R. A. Navarro-Polanco, T. Ferrer-Villada, J. Rupp, F. B. Sachse, M. Tristani-Firouzi and J. A. Sánchez-Chapula, *Proc. Natl. Acad. Sci. U. S. A.*, **2008**, *105*, 1364-1368.

[22] X. Zhang, Y. Hu and Z. Yuan, *Biochem. Biophys. Res. Commun.*, **2008**, *370*, 72-76.

[23] D. Beccati, *Investigations of prebiotics and of inter- and intra-molecular glycan-protein interactions.*, G. Ronzoni' Institute for Chemical and Biochemical Research, Milan, Italy, **2009**.

[24] Affinity User Guide, Accelrys Inc., San Diego, **2005**.

[25] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, **1983**, *79*, 926-935.

[26] S. A. Adelman and J. D. Doll, *J. Chem. Phys.*, **1976**, *64*, 2375-2388.

[27] K. M. Masukawa, P. A. Kollman and I. D. Kuntz, *J. Med. Chem.*, **2003**, *46*, 5628-5637.

[28] P. Bonnet and R. A. Bryce, *J. Mol. Graph. Model.*, **2005**, *24*, 147-156.

[29] B. J. Smith, J. L. McKimm-Breshkin, M. McDonald, R. T. Fernley, J. N. Varghese and P. M. Colman, *J. Med. Chem.*, **2002**, *45*, 2207-2212.

[30] V. Stoll, K. D. Stewart, C. J. Maring, S. Muchmore, V. Giranda, Y.-g. Y. Gu, G. Wang, Y. Chen, M. Sun, C. Zhao, A. L. Kennedy, D. L. Madigan, Y. Xu, A. Saldivar, W. Kati, G. Laver, T. Sowin, H. L. Sham, J. Greer and D. Kempf, *Biochemistry*, **2003**, *42*, 718-727.

[31] K. A. Armstrong, B. Tidor and A. C. Cheng, *J. Med. Chem.*, **2006**, *49*, 2470-2477.

[32] J.-J. Shie, J.-M. Fang, S.-Y. Wang, K.-C. Tsai, Y.-S. E. Cheng, A.-S. Yang, S.-C. Hsiao, C.-Y. Su and C.-H. Wong, *J. Am. Chem. Soc.*, **2007**, *129*, 11892-11893.

[33] J. H. Jensen and M. S. Gordon, *J. Am. Chem. Soc.*, **1991**, *113*, 7917-7924.

Table S1. Important hydrogen bonds formed between NA and the neutral and zwitterionic BA isomers [a,b]

	Donor	Acceptor	Distance (Å)	Angle (°)
nBA-NA	nBA:O1	GlyA348:O	1.67 (1.67)	170.78 (170.12)
	Tyr406:OH	nBA:O2	1.93 (1.94)	165.65 (163.73)
zBA-NA	Asn347:ND2	nBA:O1	2.11 (2.11)	160.26 (162.42)
	Arg371:NH2	nBA:O1	1.73 (1.79)	164.60 (158.06)
	Lys432:NZ	zBA:O2	2.47 (2.63)	136.64 (112.66)

[a] The atom numbering of nBA and zBA can be found in Figure 1.

[b] The values in parentheses were obtained over the beginning 1.0 ns MD trajectories .

Table S2. The vdW energies (E_{vdW}), electrostatic energies (E_{ele}) and summed interaction energies (E_{sum}) between BA and the NA active-site residues [a,b]

	nBA-NA			zBA-NA		
Residues	E_{vdW}	E_{ele}	E_{sum}	E_{vdW}	E_{ele}	E_{sum}
Arg118	-0.55 (-0.59)	-1.34 (-1.40)	-1.89 (-1.99)			
Ile149				-2.28 (-2.34)	-0.83 (-0.84)	-3.11 (-3.18)
Asp151				-0.98 (-1.52)	-12.33 (-12.38)	-13.31 (-13.90)
Arg152	-2.55 (-2.69)	-0.98 (-1.24)	-3.53 (-3.93)			
Ala246	-4.88 (-4.87)	-0.05 (0.10)	-4.93 (-4.77)			
Thr247	-2.58 (-2.74)	-0.94 (-0.93)	-3.52 (-3.67)			
Glu276	-1.29 (-1.34)	-1.69 (-1.83)	-2.98 (-3.17)			
Glu277	-1.09 (-1.06)	-1.04 (-1.08)	-2.13 (-2.14)			
Arg292	-4.21 (-4.26)	-0.10 (0.03)	-4.31 (-4.23)	-0.87 (-0.98)	-1.53 (-1.40)	-2.39 (-2.38)
Asn294	-3.94 (-4.09)	-0.06 (-0.07)	-4.00 (-4.16)			
Asn346	-1.92 (-2.11)	0.10 (0.11)	-1.82 (-2.00)			
Asn347	-3.84 (-3.80)	0.25 (0.32)	-3.59 (-3.48)	-4.65(-4.58)	-6.05 (-6.13)	-10.70 (-10.71)
Gly348	-0.86 (-0.84)	-5.10 (-5.15)	-5.96 (-5.99)			
Arg371				-1.22 (-1.51)	-51.20 (-48.19)	-52.42 (-49.70)
Tyr406	-2.30 (-2.31)	-4.43 (-4.42)	-6.73 (-6.73)	-2.19 (-2.15)	0.28 (0.24)	-1.91 (-1.91)
Pro431				-5.24 (-5.41)	1.90 (1.97)	-3.34 (-3.44)
Lys432				0.59 (0.39)	-74.04 (-73.81)	-73.45 (-73.42)
Sum	-30.01 (-30.70)	-15.38 (-15.56)	-45.39 (-46.26)	-16.84 (-18.10)	-143.80 (-140.54)	-160.64 (-158.64)

[a] Energy units in kcal mol^{-1};

[b] The values in parentheses were obtained over the beginning 1.0 ns MD trajectories.

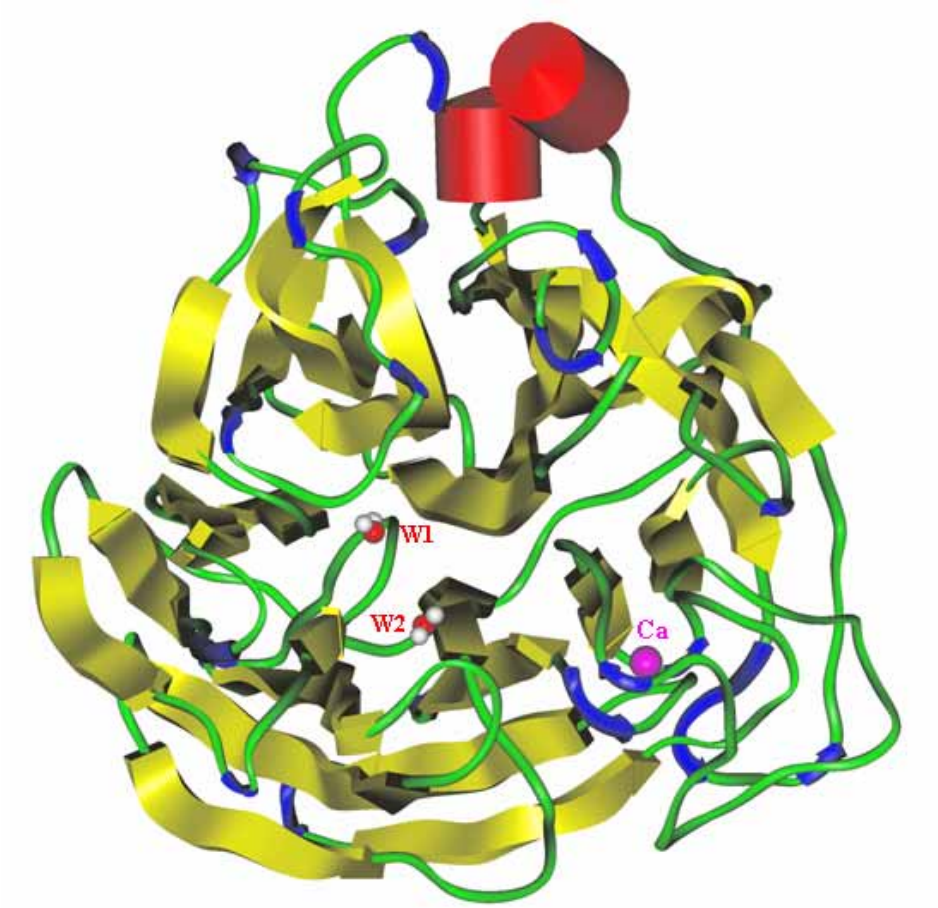


Figure S1. Ribbon diagram of the N9-subtype neuraminidase structure (NA).

Ribbon colors: Helices (including α-, 3_{10}- and π-helix), hydrogen-bonded turns, extended strands and random coils are in red, blue, yellow and green, respectively. The Ca^{2+} ion and two crystal water molecules (W1 and W2) are represented with the ball and stick models.

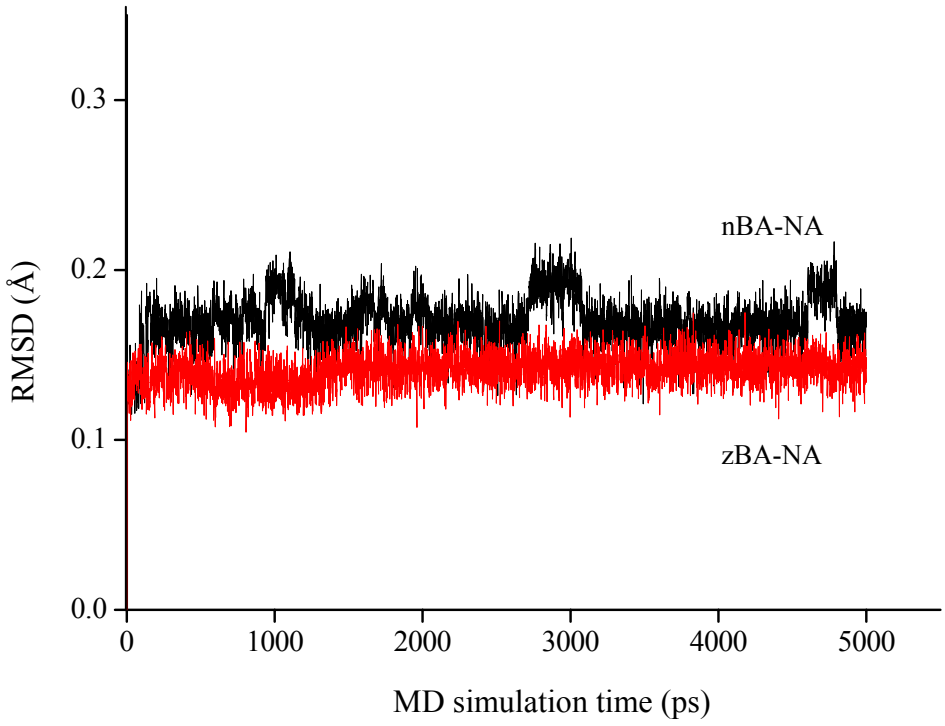


Figure S2. The time-evolution backbone-atom root mean square deviations (RMSD) of the protein structures in the nBA-NA (in black, on the top) and zBA-NA (in red, on the bottom) complexes.

The RMSD values were obtained by the comparing with the energy-minimized structures of docking.

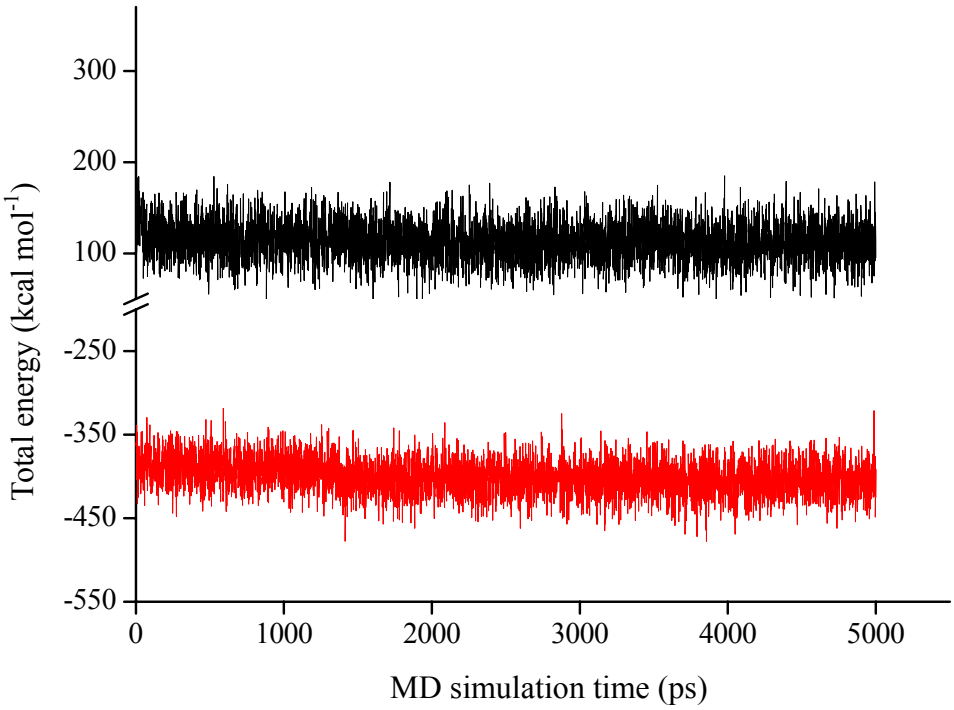
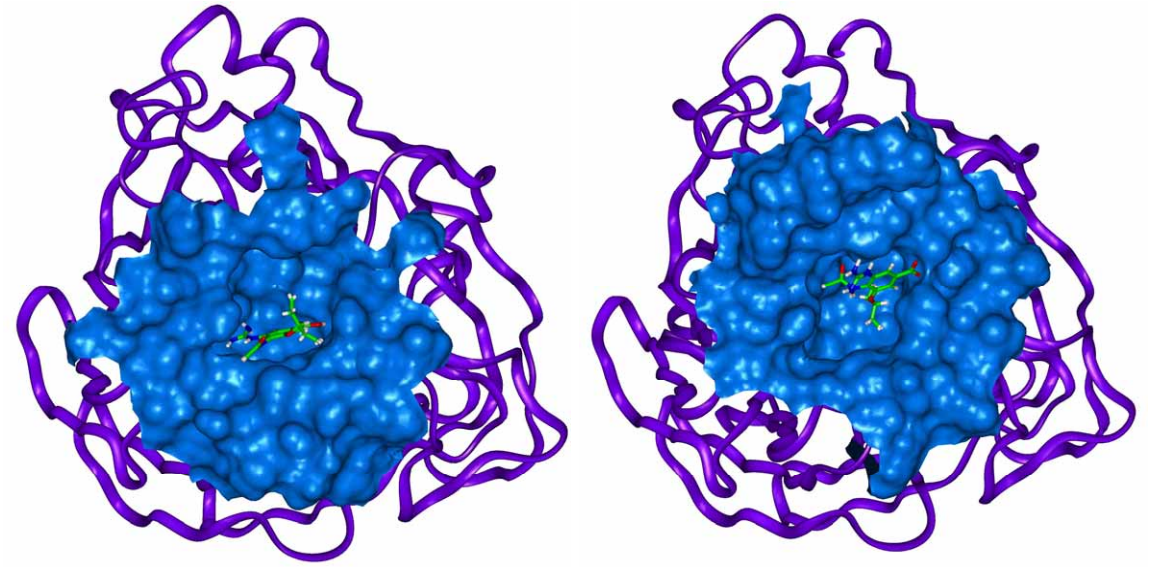


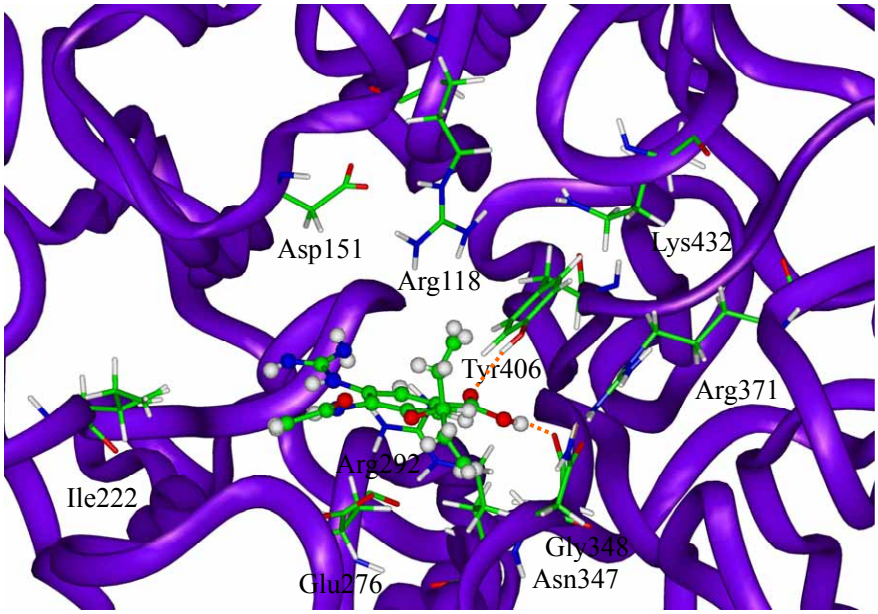
Figure S3. The time-evolution total energies of the nBA-NA (in black, on the top) and zBA-NA (in red, on the bottom) complexes.



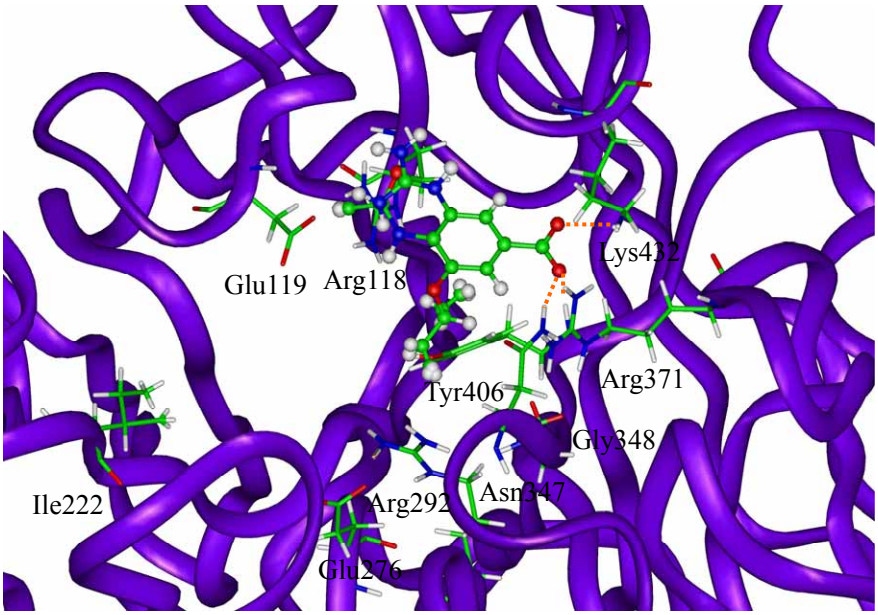
(a) nBA-NA (b) zBA-NA

Figure S4. The neutral and zwitterionic BA isomers at the NA active sites.

The NA proteins are plotted in purple while the residues within 10 Å (free to move) from the centers of the neutral or zwitterionic BA isomer in light blue Connolly surface. Inhibitors are represented with stick models (C in green, O in red, N in blue and H in white, respectively).



(a) nBA-NA



(b) zBA-NA

Figure S5. Interactions of the NA active-site residues with the neutral and zwitterionic BA isomers.

The important hydrogen bonds are labeled in the dashed orange lines.

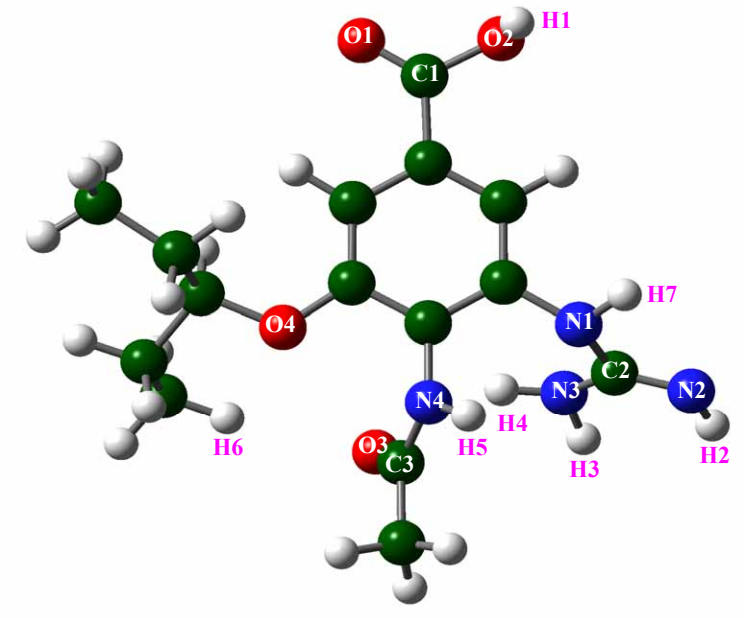


Figure S6. The transition state structure (TSa) of Step 1 in the absence of water

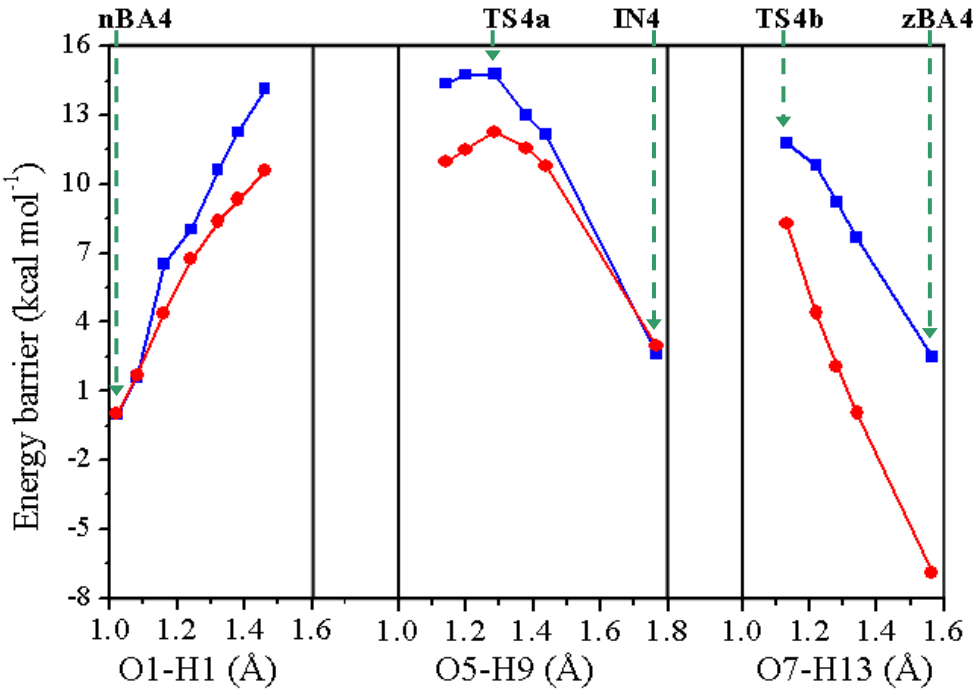


Figure S7. The intrinsic reaction coordinates (IRC) for the proton transfer nBA4 → zBA4 obtained at the B3LYP/6-311++G(d,p)//6-31G(d) theoretical level.

The data of the discrete water and implicit solvation models were plotted in blue and red, respectively.