Supporting Information

Revealing the Transient Conformations of a Single Flavin Adenine Dinucleotide by Aerolysin Nanopore

Meng-Yin Li, [a] Ya-Qian Wang, [a] Yi-Lun Ying [a] and Yi-Tao Long*[b]

a   School of Chemistry and Molecular Engineering, East China University of Science and Technology, Shanghai, 200237, P. R. China

b   State Key Laboratory of Analytical Chemistry for Life Science, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, 210023, P. R. China

*Correspondence to: yitaolong@nju.edu.cn

Table of Contents

Reagents and chemicals

Trypsin-EDTA, NaCl, Na_3PO_4, HCl, decane (anhydrous, ≥99%) and flavin adenine dinucleotide were purchased from Sigma-Aldrich Co., Ltd. (St. Louis, MO). 1, 2-Diphytanoyl-sn-glycero-3-phosphocholine (chloroform, ≥99%) was purchased from Avanti Polar Lipids, Inc. (Alabaster, AL). The reagents and materials we used are all of the analytical grade. The solutions in our experiments were preparing with ultrapure water (18.2 MΩ cm at 25 °C) from a Milli-Q system (Billerica, MA).

Single channel recording

Wild-type proaerolysin was produced by expressing in BL21 pLysS *E. coli* and purified by using a Ni column. Then the proaerolysin was activated to form monomeric aerolysin by trypsin-EDTA at 37 ℃ temperature for 1 h. The lipid bilayer was formed from 1, 2-diphytanoyl-sn-glycero-3-phosphocholine (Avanti Polar Lipids) and spanning a 50 μm orifice in a Delrin bilayer cup (Warner Instruments, Hamden, CT). We conducted a series of nanopore experiments in the 1.0 M KCl, 10 mM Tris, pH 8.0, 1.0 mM EDTA, with different gradient concentration of acetonitrile (0%, 5%, 10%, 15%, 20%). Both *cis* and *trans* compartments of the recording chamber contained 1.0 mL buffer solution. The potential is applied by using Ag/AgCl electrodes. The FAD was added to the *cis* compartment to a final concentration of 2.0 μM. The experiments in our work were conducted at 24 ± 2 °C.

Data acquisition and analysis

The current traces were measured with a patch clamp amplifier (Axon 200B equipped with a Digidata 1440A A/D converter, Molecular Devices, USA) with the *cis* chamber connected to ground. The amplified signal (arising from the ionic current passing through the pore) was low-pass filtered at 5 kHz and sampled at 100 kHz by running the Clampex 10.4 software (Molecular Devices, USA). The data analysis was performed by using Mosaic software[1] and Origin-Lab 8.0 (Origin-Lab Corporation, Northampton, MA)

Fitting of the distribution

We fitted mixed-Gaussian model to the distribution of I/I_0 and ψ_{N-O-N}. A linear super-position of univariate Gaussians is given as

$$p(x) = \sum_{n=1}^{N} \pi_n N(x|\mu_n, \Sigma_n)$$

where *N* is the number of Gaussians, π_n is the mixing coefficient, μ_n, Σ_n are the mean and variance of each Gaussian distribution, respectively. Parameters above can be calculated using Expectation Maximization (EM) technique. The EM algorithm is an iterative optimization technique which contains Estimation step and Maximization step. After initializing the parameters, these two steps are iteratively evaluated for optimizing the log likelihood of the mixed model as

$$\ln p(x \mid \mu, \Sigma, \pi) = \sum_{k=1}^{K} \ln \left\{ \sum_{n=1}^{N} \pi_n N(x_k \mid \mu_n, \Sigma_n) \right\}$$

where k is the number of events. In Estimation step, the responsibilities γ_n is evaluated using the current parameter values by the equation below

$$\gamma_n(x) = \frac{\pi_n N(x \mid \mu_n, \Sigma_n)}{\sum_{j=1}^{N} \pi_j N(x \mid \mu_j, \Sigma_j)}$$

Then, the current responsibilities are used to re-estimate the parameters which are derived by

$$\mu_n = \frac{\sum_{k=1}^{K} \gamma_n(x_k) x_k}{\sum_{k=1}^{K} \gamma_n(x_k)}$$

$$\Sigma_n = \frac{\sum_{k=1}^{K} \gamma_n(x_k)(x_k - \mu_n)^2}{\sum_{k=1}^{K} \gamma_n(x_k)}$$

$$\pi_n = \frac{1}{K} \sum_{k=1}^{K} \gamma_n(x_k)$$

In this work, we implemented the above optimization procedure using the Gaussian mixture model in *scikit-learn*[2] , which is a machine learning package in python. The optimized number of Gaussians was determined by evaluating the minimum AIC value.

Estimation of effective charge

z_{inside} is the number of effective charges of FAD inside the aerolysin pore and is estimated by $\tau = \tau_0 \exp (-z_{inside} eV/k_b T)$,[3] where τ is the duration of FAD traversing through aerolysin, τ_0 is a diffusive relaxation time associated with the duplex, e is the magnitude of the elementary charge, k_b is the Boltzmann constant, T is the and V is the applied potential. τ is calculated by fitting the duration to an exponential curve (Figure S3, 4).

Molecular dynamics simulations

Model Construction

The FAD coordinates were separated from the Protein Data Bank (entry 3M13).[4] The structure of FAD was described by the parameters generated from CGenFF program[5,6]. For the simulations in aqueous solution, the FAD

molecule was solvated with pre-equilibrated TIP3P water molecules[7] to a box of 70 Å x 70 Å x 70 Å using VMD[8]. K^+ and Cl^- ions were added to the solution to neutralize the FAD molecule and achieve the concentration of 1.0 M. For the simulations in the mixed solution of water and acetonitrile, the mixed solution was firstly constructed and equilibrated. The mixed solutions with acetonitrile concentration of 5%, 10%, 15% and 20% were built with Packmol program[9] according to the ratio of molecule numbers between water and acetonitrile at each concentration (detailed setup see in Table S). The parameters of acetonitrile were also generated by CGenFF[5,6] program. The mixed solution at each concentration was equilibrated in the NVT ensemble 10 ns and then in the NPT ensemble for 40 ns.

For the simulations in bulk solution, the FAD molecule was solvated with the mixed solution equilibrated above. Then, the systems were equilibrated in NVT ensemble 20 ns and then in the NPT ensemble for 80 ns.

For the simulations in confined space, the "open", "stack" and "quasi stack-1" conformation of FAD were placed in the R1 region of aerolysin. The R1 region is cut from the equilibrated all atom aerolysin module as described in our previous work[10]. Then, the FAD-R1 systems were solvated in 1.0 M KCl, 1.0 M KCl and 5% ACN-KCl mixed solution for "open", "stack" and "quasi stack-1" conformation, respectively. After that, the systems were equilibrated in NVT ensemble 20 ns and then in the NPT ensemble for 80 ns.

System Equilibrium.
All MD simulation were carried out using NAMD[11] program with the analysis and visualization using VMD[8]. The parameters of FAD and acetonitrile generated from CGenFF[5,6] were combined to CHARMM36 force filed. For the equilibrium of mixed solution, after a 10000-step minimization, the mixed solution was equilibrated in the NVT ensemble at 295 K for 10 ns. Next, the equilibration simulation was performed in the NPT ensemble also at 295K and 1 atm for 40 ns. The mixed solution before and after the equilibrium at each concentration was illustrated in Fig. S. For the simulation for the FAD system, after a 10000-step minimization, FAD system was simulated in the NVT ensemble at 295 K for 20 ns. Afterwards, the simulation lasted 80 ns in the NPT ensemble at 295K and 1 atm. All NVT ensemble simulation were performd with Langevin dynamics for temperature control,[12] while the NPT ensemble simulations were carried out with a Langevin piston for pressure control and Langevin dynamics for temperature control.[12] The integration time step chosen was 1 fs for mixed solution equilibrium and 2 fs for FAD simulation. Periodic boundary conditions were employed in all directions. van der Waals energies were calculated using a 12 Å cutoff and the particle-mesh-Ewald (PME) method was used to treat long-range electrostatics.[13]

For the simulations in bulk solution, the last 60-ns simulations in NPT ensemble were used for analysis to ensure the stable of the FAD structure. To evaluate

the conformation of FAD, we chose the angle formed by N_{9A} in adenine, O_{3P} between two phosphorus atoms and N_{10} in flavin (Fig. S) to represent the conformational characteristics. This angle in each frame was calculated according to the trajectory simulated above.

For the simulations in confined space, the angle were also calculated from the last 60-ns simulations. Furthermore, to confirm the stability of each conformation, RMSD of the FAD molecule was calculated using RMSD Trajectory Tool plugin in VMD.

"stack" "open"

Volume: 671.5 Å^3
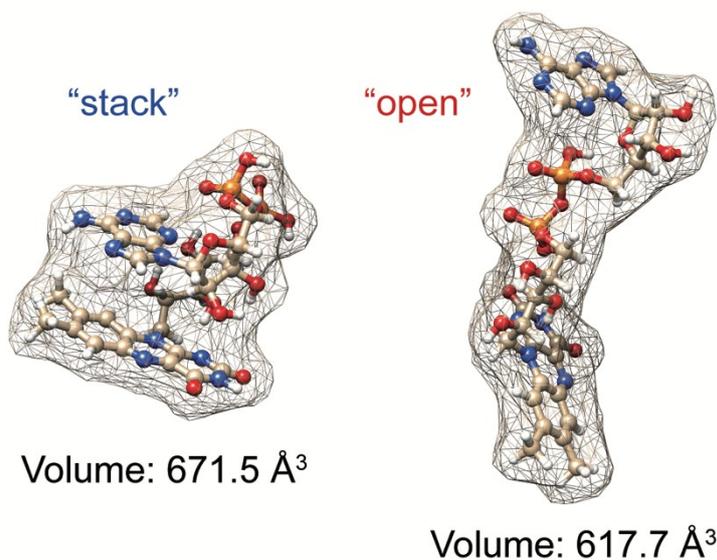
Volume: 617.7 Å^3

Figure S1 The calculated volume of stacked and opened FAD. The volume of stacked FAD is about 671.5 Å^3, while it of opened FAD is about 617.7 Å^3.
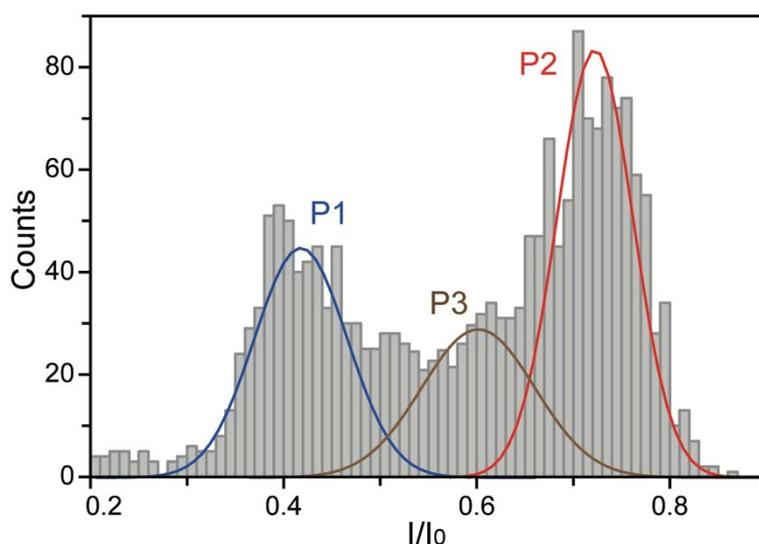
Figure S2 I/I_0 histograms of FAD blockages by K238Q mutant aerolysin at + 80 mV. The data were acquired in 1.0 M KCl, 10 mM Tris, and 1.0 mM EDTA at pH 8.0 in the presence of 2.0 μM FAD. The region near K238 is considered as one of two sensing regions inside aerolysin pore according to our previous

work[10]. The observation that FAD inside K238Q aerolysin still exhibits three populations suggests that the mutation on one of sensing region does not affect the behavior of FAD inside the pore, excluding the multiple binding of FAD to different region of aerolysin.
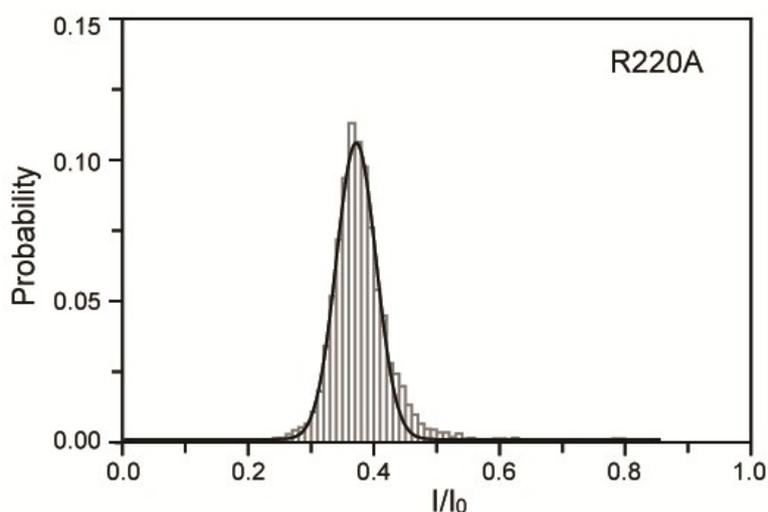
Figure S3 I/I_0 histograms of FAD blockages by R220A mutant aerolysin at + 80 mV. The data were acquired in 1.0 M KCl, 10 mM Tris, and 1.0 mM EDTA at pH 8.0 in the presence of 2.0 µM FAD.
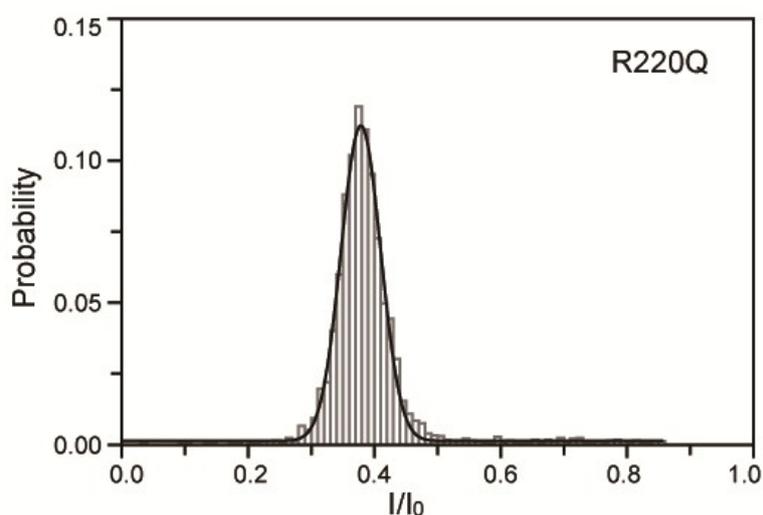

Figure S4 I/I_0 histograms of FAD blockages by R220Q mutant aerolysin at + 80 mV. The data were acquired in 1.0 M KCl, 10 mM Tris, and 1.0 mM EDTA at pH 8.0 in the presence of 2.0 µM FAD. The FAD in R220A and R220Q mutant aerolysin show only one peak in I/I_0, located at the P1 in WT aerolysin, which exhibits much more differentials than that in K238Q mutant pore. This observation indicates that the capturing of opened or linear FAD is significantly limited in R220A and R220Q mutant aerolysin, supporting the importance of R1

region. Therefore, the distinguishable current blockages of FAD are caused by different conformations of FAD in R1 region rather than the multiple binding sites.
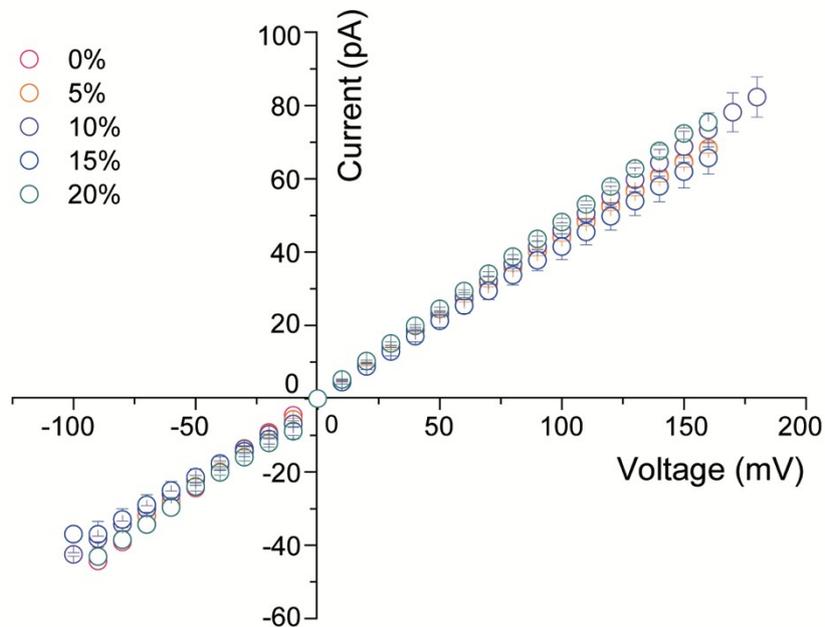
Figure S5 Current-voltage curves of WT aerolysin nanopore. The data were acquired in the condition of 1.0 M KCl, 10 mM Tris, 1.0 mM EDTA, at pH 8.0 with different acetonitrile concentration. Each coloured circle stands for an acetonitrile concentration. Three independent experiments were calculated in each curve.
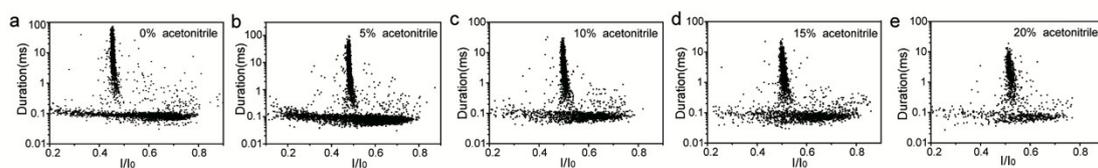
Figure S6 Scatter plots of poly(dA)$_4$ by WT aerolysin at + 80 mv in the 1.0 M KCl (10 mM Tris, 1.0 mM EDTA, pH 8.0) solution with acetonitrile concentration of 0% (a), 5% (b), 10% (c), 15% (d), 20% (e). The presence of poly(dA)$_4$ is 2.0 µM.
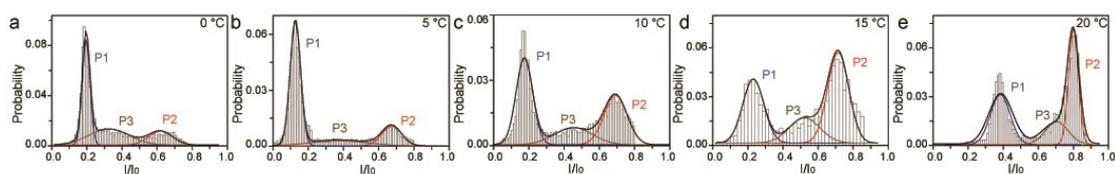
Figure S7 Effects of temperature on the distribution of I/I_0 for the FAD inside aerolysin nanopore. (a) 0 ℃, (b) 5 ℃, (c) 10 ℃, (d) 15 ℃, (e) 20 ℃. The data were acquired in 1.0 M KCl, 10 mM Tris, and 1.0 mM EDTA at pH 8.0 in the presence of 2.0 μM FAD.

a

b

80x80x80 (Å)   69.7x70.4x70.5 (Å)
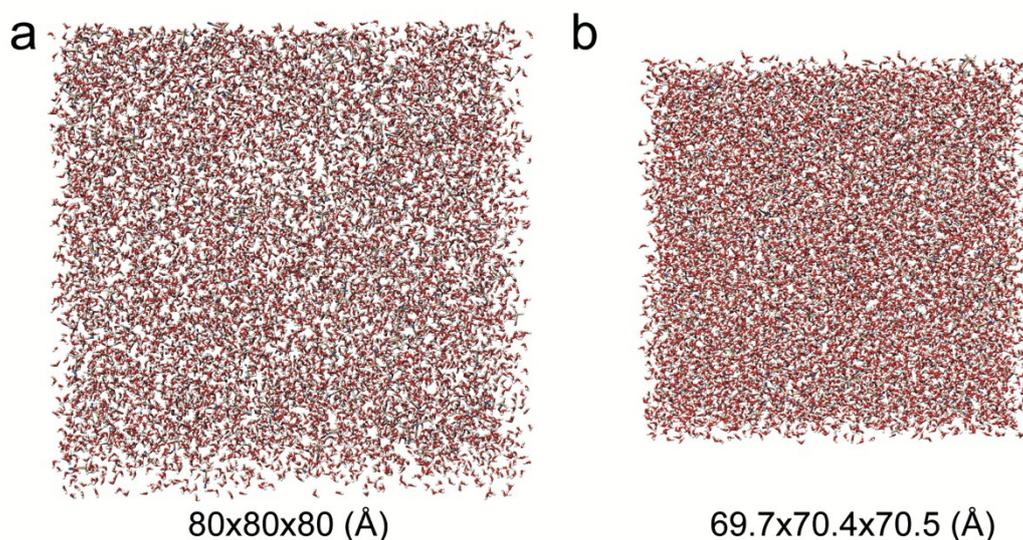
**Figure S8** The equilibrium of the 5% acetonitrile in acetonitrile-water solution. (a) Initial snapshot of mixed solution with 5% acetonitrile. (b) The mixed solution after 50-ns equilibrium at 295K and 1 atm.

a

b

80x80x80 (Å)   72.2x71.8x71.6 (Å)
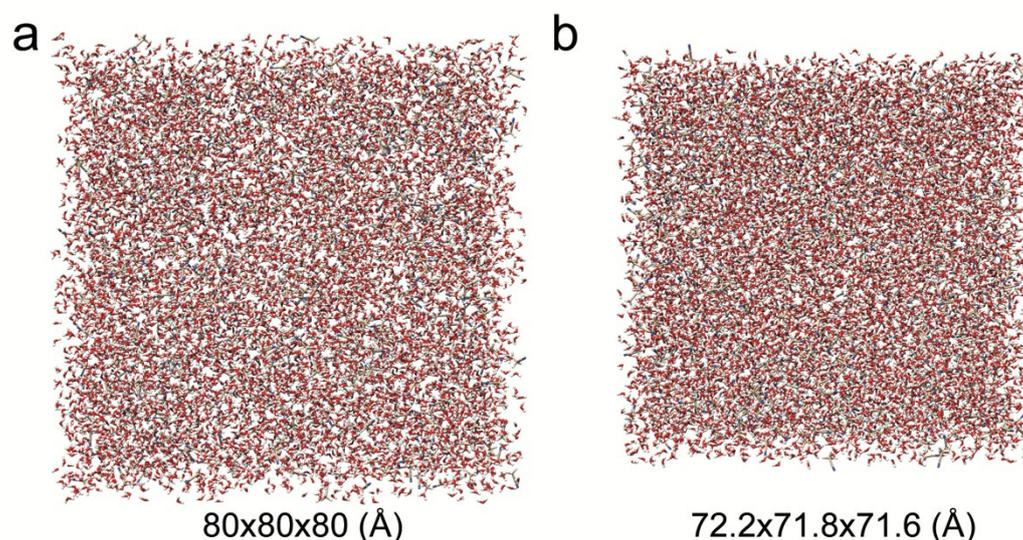
Figure S9 The equilibrium of the 10% acetonitrile in acetonitrile-water solution. (a) Initial snapshot of mixed solution with 10% acetonitrile. (b) The mixed solution after 50-ns equilibrium at 295K and 1 atm.

a   b

80x80x80 (Å)   74.6x74.6x73.6 (Å)
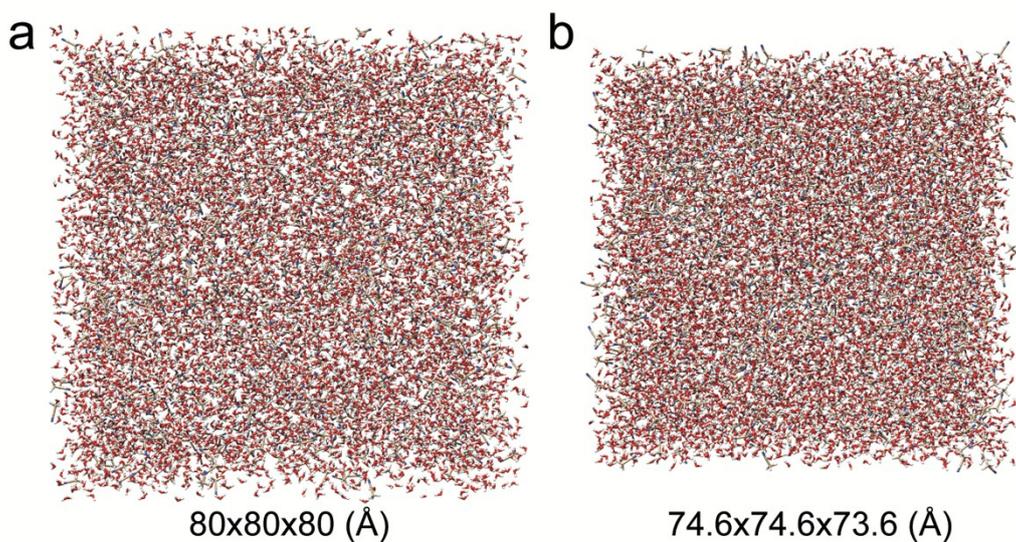
Figure S10 The equilibrium of the 15% acetonitrile in acetonitrile-water solution. (a) Initial snapshot of mixed solution with 15% acetonitrile. (b) The mixed solution after 50-ns equilibrium at 295K and 1 atm.

a   b

80x80x80 (Å)   75.5x76.6x76.6 (Å)
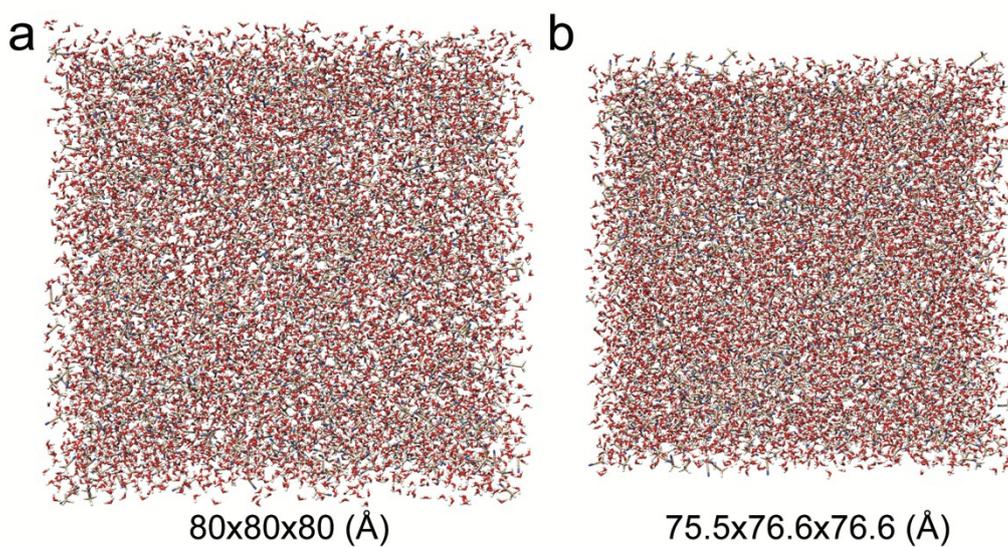
Figure S11 The equilibrium of the 20% acetonitrile in acetonitrile-water solution. (a) Initial snapshot of mixed solution with 20% acetonitrile. (b) The mixed solution after 50-ns equilibrium at 295K and 1 atm.
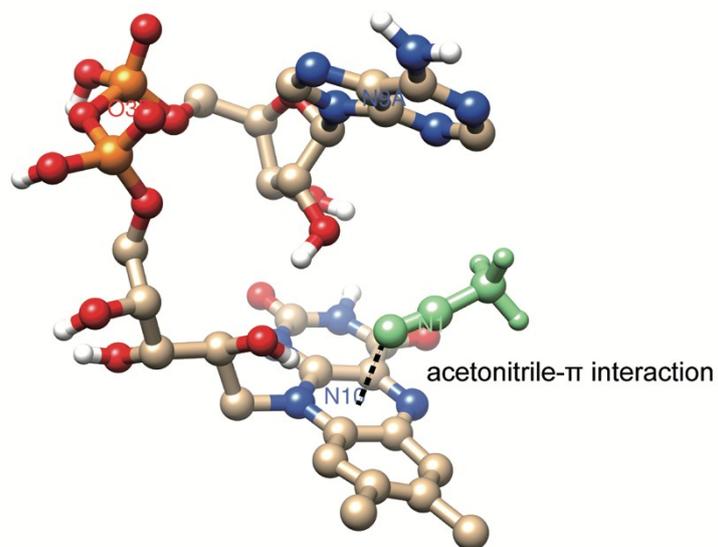
**Figure S12** The interaction between quasi stacked FAD and acetonitrile molecule. Quasi stacked FADs are stabilized by the insertion of acetonitrile molecule into the space between the isoalloxazine ring and the adenine, which could form acetonitrile-π interactions.
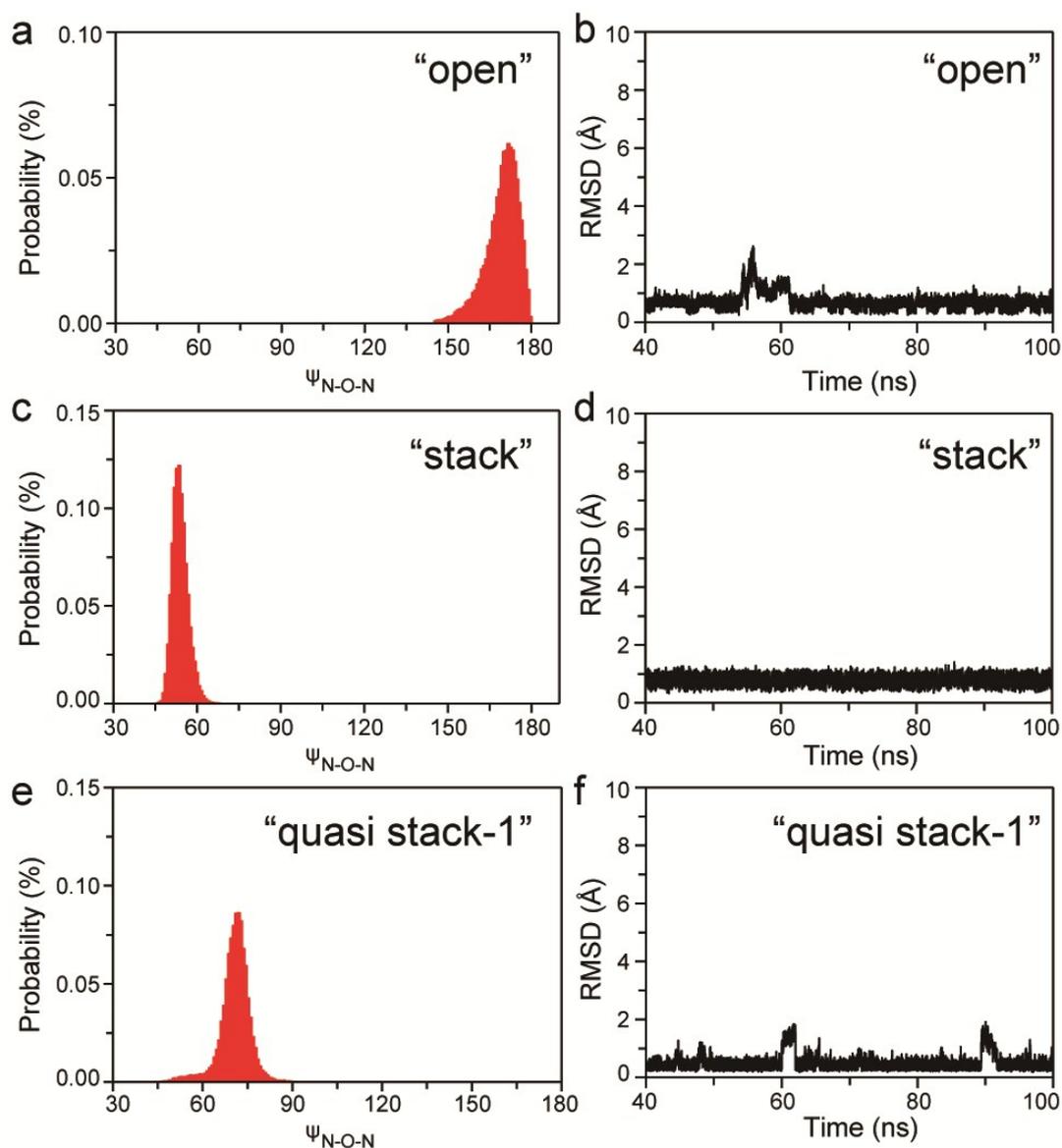
Figure S13 The stability of FAD with different conformations in R1 region of aerolysin. (a)-(f) The backbone angle of FAD molecule (ψ_{N-O-N}) with "open" (a), "stack" (c) and "quasi stack-1" (e) conformation during the last 60 ns simulations. Corresponding RMSD of "open" (b), "stack" (d) and "quasi stack-1" (f) conformations calculated from the trajectories of last 60 ns simulations.
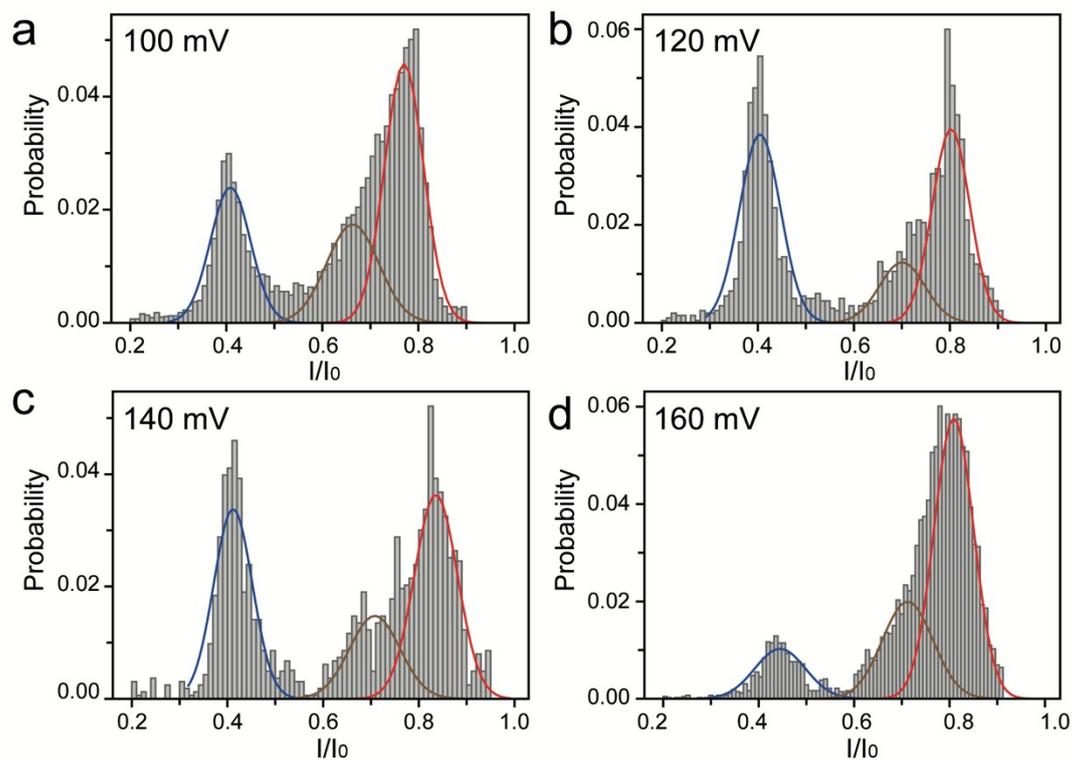
Figure S14 I/I_0 histograms of FAD molecule by WT aerolysin at potential of (a) 100 mV, (b) 120 mV, (c) 140 mV, (d) 160 mV. The data were acquired in 1.0 M KCl, 10 mM Tris, 1.0 mM EDTA, pH 8.0 and in the presence of 2.0 μM FAD.
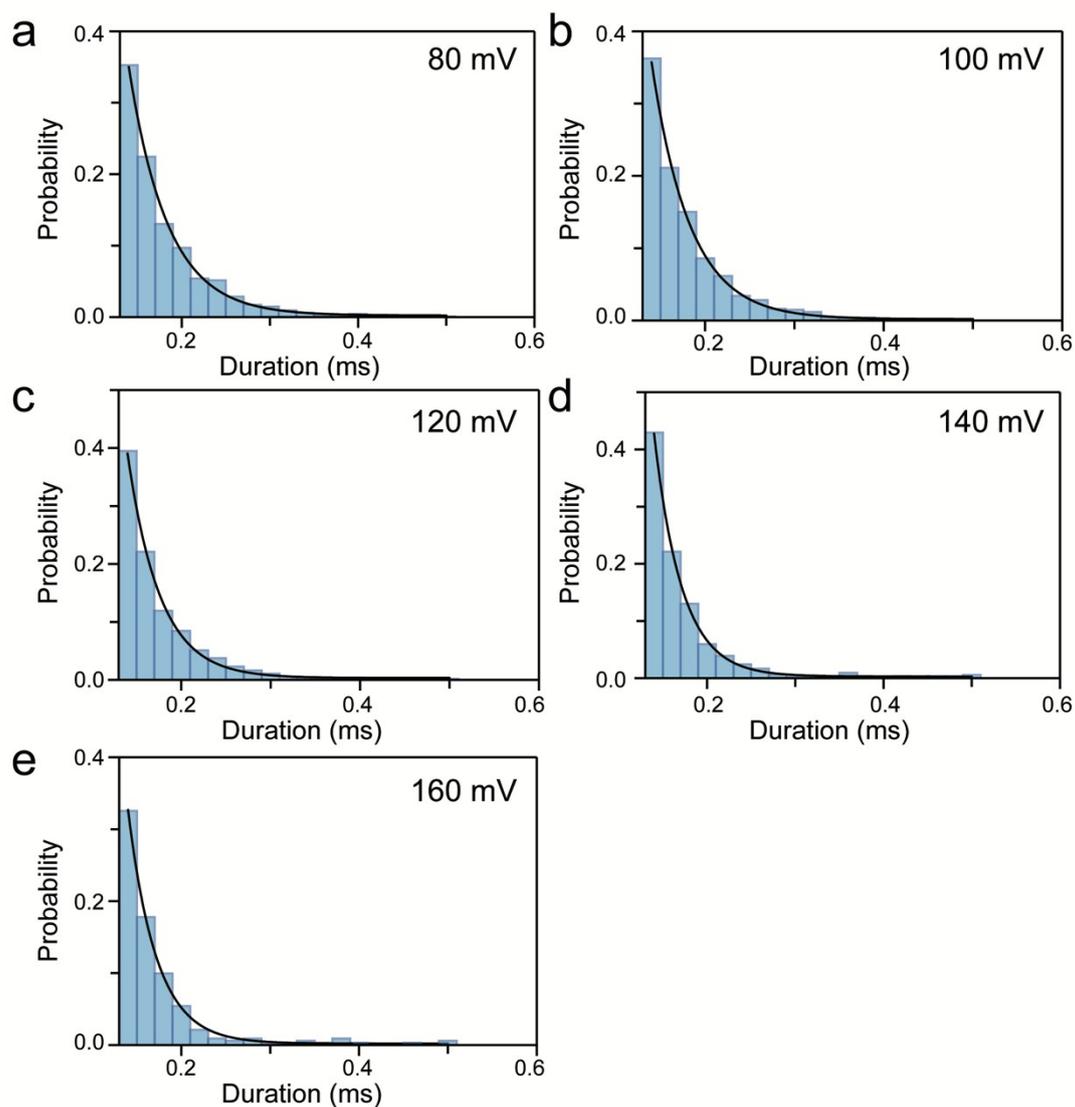
Figure S15 Duration time histograms of P1 of FAD by WT aerolysin at potential of (a) 80 mV, (b) 100 mV, (c) 120 mV, (d) 140 mV, (e) 160 mV. All of the histograms were fit to single Exponential function. The data were acquired in 1.0 M KCl, 10 mM Tris, 1.0 mM EDTA, pH 8.0 and in the presence of 2.0 μM FAD.
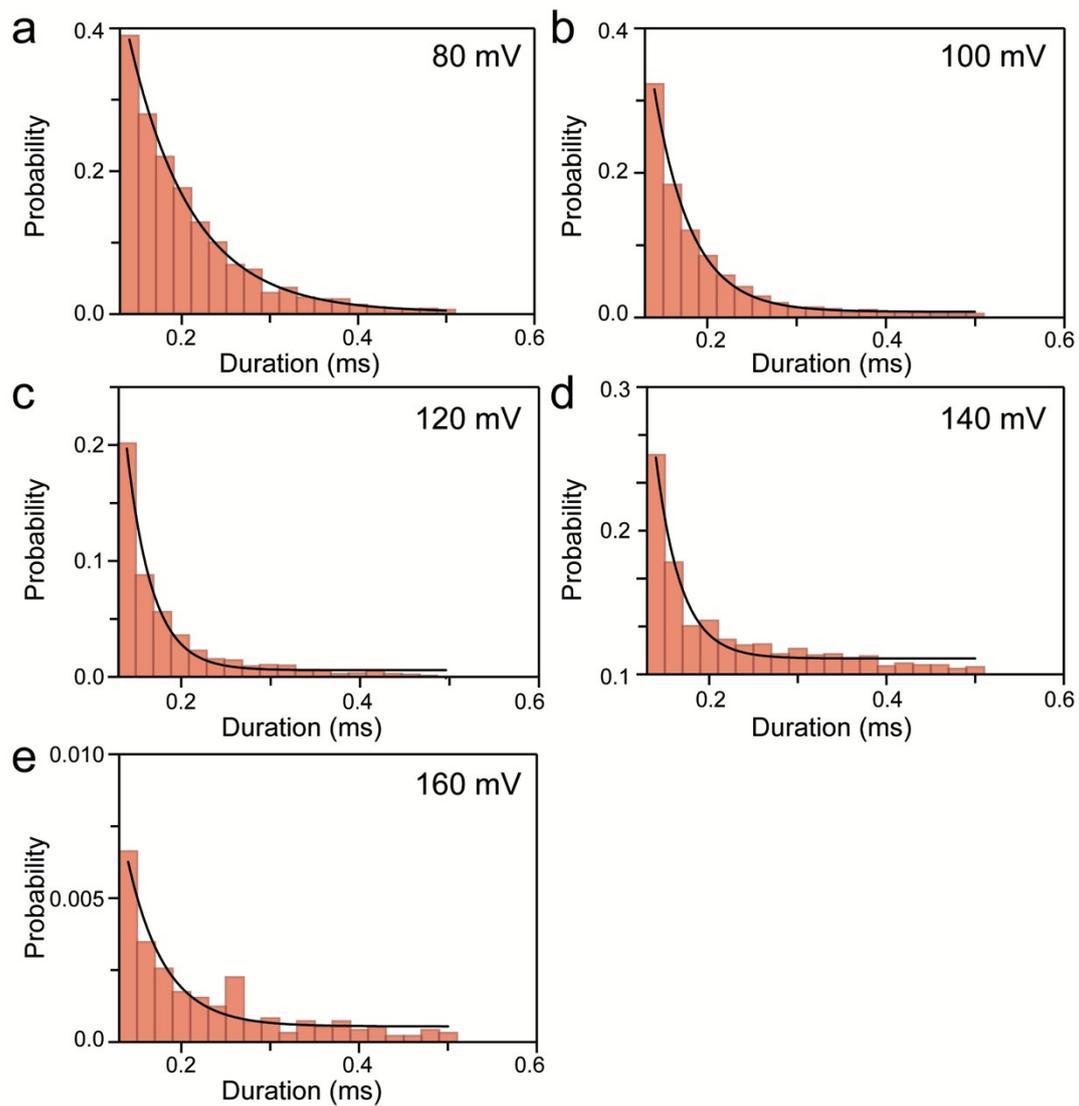
Figure S16 Duration time histograms of P2 of FAD by WT aerolysin at potential of (a) 80 mV, (b) 100 mV, (c) 120 mV, (d) 140 mV, (e) 160 mV. All of the histograms were fit to single Exponential function. The data were acquired in 1.0 M KCl, 10 mM Tris, 1.0 mM EDTA, pH 8.0 and in the presence of 2.0 μM FAD.
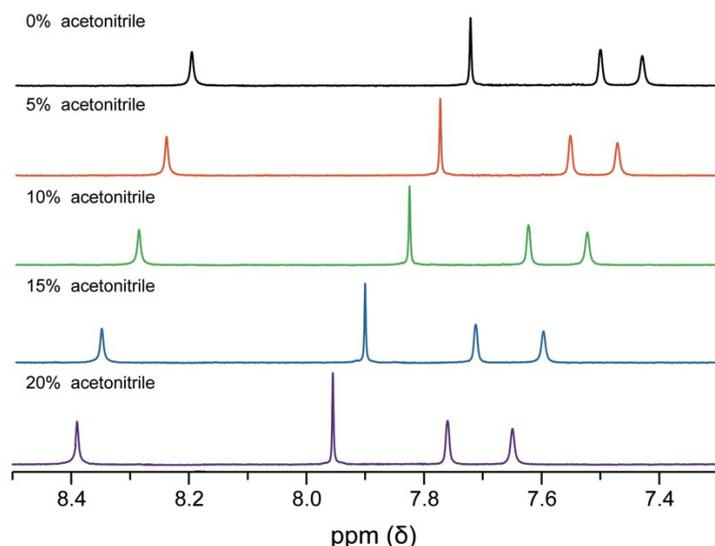
Figure S17 Proton nuclear magnetic resonance spectrum (^{1}H NMR) of the aromatic region of 10mg/mL FAD. The solvent from the top to the bottom is 0%，5%, 10%, 15%, 20% CD$_3$CN in D$_2$O solution, respectively.
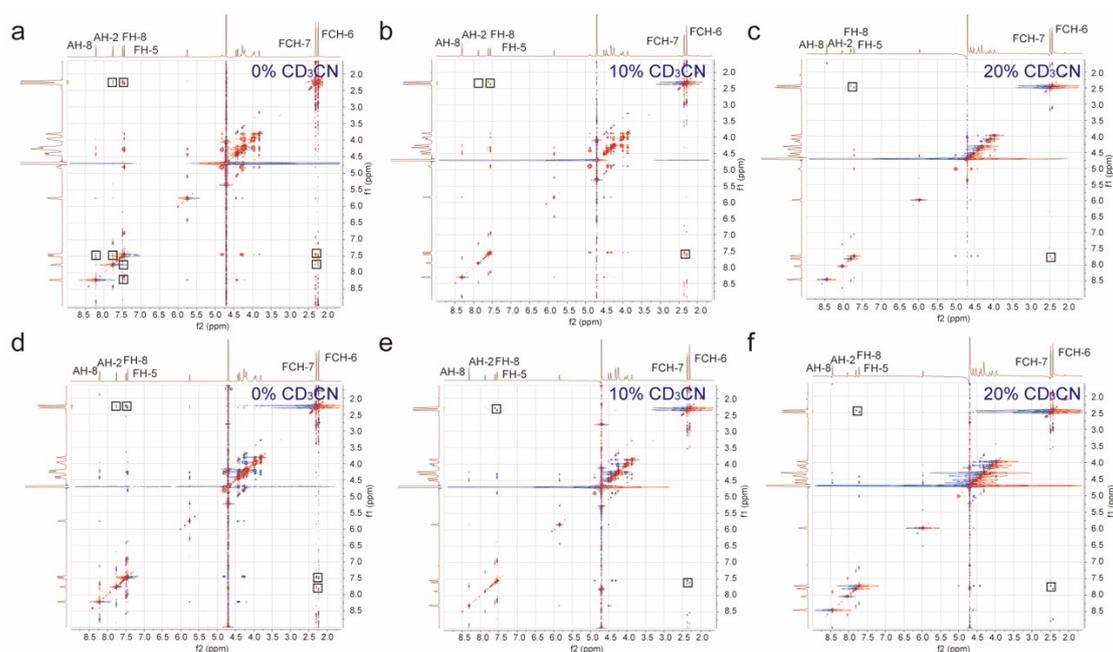
Figure S18 ^{1}H-^{1}H nuclear overhauser effect spectroscopy (NOESY) and rotating frame overhauser effect spectroscopy (ROESY) of FAD. (a)-(c) The NOESY of FAD in D$_2$O with CD$_3$CN ratio of 0% (a), 10% (b) and 20% (c). The cross-peaks were marked with black boxes. (d)-(f) The ROESY of FAD in D$_2$O with CD$_3$CN ratio of 0% (d), 10% (e) and 20% (f). The experiments were performed in 600 MHz spectrometer with the FAD concentration of 10 mM. As shown in Fig. S18a, there are several cross-peaks between AH-8 and FH-8, AH-8 and FH-5, AH-2 and FH-8, AH-2 and FH-5, AH-2 and FCH-7, AH-2 and

FCH-6, FH-8 and FCH-7, FH-5 and FCH-6, representing the NOE transfer between each pair of position and indicating the short distance between them. This result indicates the stacking of isoalloxazine and adenine in FAD molecule in water condition. However, with the increasing of CD$_3$CN ratio, the cross-peaks between AH-8 and FH-8, AH-8 and FH-5, AH-2 and FH-8, AH-2 and FH-5, AH-2 and FCH-7, AH-2 and FCH-6 gradually disappear (Fig. S18b and 18c), while the cross-peaks between FH-8 and FCH-7, FH-5 and FCH-6 still exist. These results suggest a longer distance between the isoalloxazine and adenine induced by the addition of CD$_3$CN, which further supports the extending of FAD backbone angle observed in nanopore experiments. Similar to NOESY, for ROESY, the cross-peaks between hydrogen in isoalloxazine and adenine disappear with the CD$_3$CN ratio increasing, while the NOE transfer between FH-8 and FCH-7, FH-5 and FCH-6 in isoalloxazine still exist. These results further support the expanding of FAD backbone in CD$_3$CN conditions. Therefore, the acetonitrile would indeed induce the conformational change of FAD.
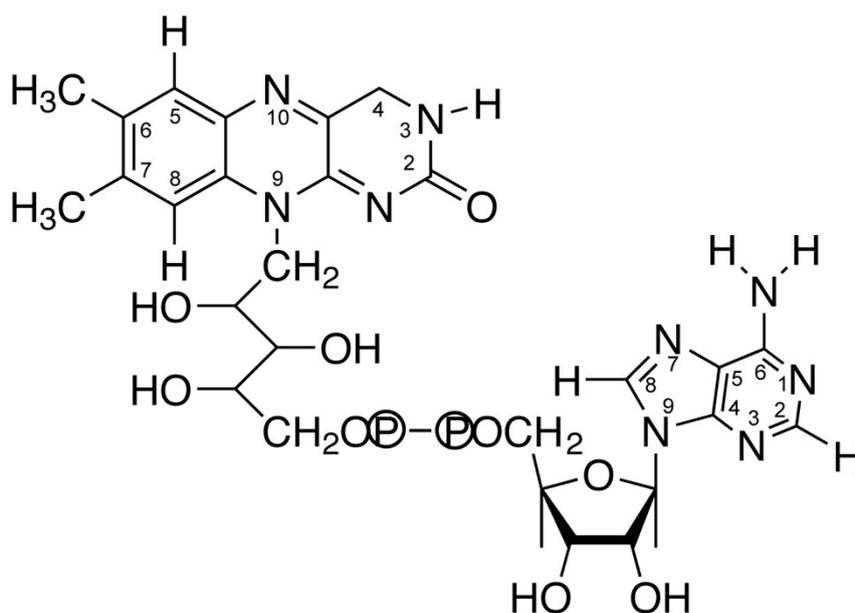
Figure S19. The structure of flavin adenine dinucleotide.

Figure S20 I/I_0 histogram of (a) poly(dA)$_2$ and (b) FMN by WT aerolysin nanopore at + 80 mV. The data were acquired in 1.0 M KCl, 10 mM Tris, and 1.0 mM EDTA at pH 8.0 in the presence of 2.0 µM FAD.

Table S1 I/I_0 values of three populations generated from FAD molecules by aerolysin nanopore.

I/I_0	0%	5%	10%	15%	20%
P1	0.42 ± 0.01	0.30 ± 0.01	0.31 ± 0.01	0.42 ± 0.03	0.53 ± 0.01
P2	0.76 ± 0.02	0.70 ± 0.01	0.69 ± 0.02	0.71 ± 0.01	0.66 ± 0.01
P3	0.71 ± 0.02	0.57 ± 0.01	0.52 ± 0.01	0.56 ± 0.02	0.59 ± 0.02

Table S2 Setup for the construction of mixture of acetonitrile and water with different ratio using Packmol program.

Acetonitrile percentage (%)	molecule number		Box size (Å)
	water	acetonitrile	
5	15240	264	80x80x80
10	14872	544	80x80x80
15	14047	816	80x80x80
20	13457	1107	80x80x80

Table S3 I/I_0 values, area percentage of each peak and the R-Square value of Gaussian mixture model fitting to the distribution of I/I_0 for FAD in WT aerolysin nanopore experiments.

ACN ratio	P1		P3		P2		R-square
	I/I_0	Percentage	I/I_0	Percentage	I/I_0	Percentage	
0%	0.42	36.2%	0.71	20.1%	0.78	43.7%	0.927
5%	0.30	32.5%	0.57	24.8%	0.70	42.7%	0.947
10%	0.31	33.2%	0.52	26.2%	0.69	40.6%	0.963
15%	0.42	31.9%	0.56	24.4%	0.71	43.7%	0.964
20%	0.53	43.9%	0.59	6.4%	0.66	49.7%	0.980

Table S4 ψ_{N-O-N} values, area percentage of each peak and the R-Square value of Gaussian mixture model fitting to the distribution of ψ_{N-O-N} for FAD in MD simulations.

ACN ratio	P1		P3		P2		R-square
	ψ_{N-O-N}	Percentage	ψ_{N-O-N}	Percentage	ψ_{N-O-N}	Percentage	

0%	55°	39.6%	102°	51.7%	161°	8.7%	0.951
5%	78°	15.1%	106°	48.0%	160°	36.9%	0.954
10%	105°	24.7%	123°	50.2%	162°	25.1%	0.985
15%	110°	60.5%	139°	14.4%	162°	25.1%	0.971
20%	116°	53.7%	148°	3.9%	164°	42.3%	0.974

Reference

1 A. Balijepalli, J. Ettedgui, A. T. Cornio, J. W. F. Robertson, K. P. Cheung, J. J. Kasianowicz, C. Vaz, *ACS Nano* **2014**, *8*, 1547-1553.

2 A. Swami, R. Jain, *J. Comput. Sci. Tech.* **2011**, *12*, 2825-2830.

3 J. Nakane, M. Wiggin, A. Marziali, *Biophys. J.* **2004**, *87*, 615-621.

4 G. Kuppuraj, D. Kruise, K. Yura, *J. Phys. Chem. B* **2014**, *118*, 13486-13497.

5 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. MacKerell Jr., *J. Comput. Chem.* **2010**, *31*, 671-690.

6 W. Yu, X. He, K. Vanommeslaeghe, A. D. MacKerell Jr., *J. Comput. Chem.* **2012**, *33*, 2451-2468.

7 W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, M. Klein, *J. Chem. Phys.* **1983**, *79*, 926-935.

8 W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33-38.

9 L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez. *J. Comput. Chem.* **2009**, *30*, 2157-2164.

10 Y.-Q. Wang, M.-Y. Li, H. Qiu, C. Cao, M.-B. Wang, X.-Y. Wu, J. Huang, Y.-L. Ying, Y.-T. Long. *Anal. Chem.* **2018**, *90*, 7790-7794.

11 J. Phillips, *et al. J. Comput. Chem.* **2005**, *26*, 1781-1802.

12 Z.-L. Hu, *et al. Anal. Chem.* **2018**, *90*, 4268-4272.

13 P. Batcho, D. Case, T. Schlick, *J. Chem. Phys.* **2001**, *115*, 4003-4018.