molecular pharmaceutics

articles

Molecular Basis of the Interaction for an Essential Subunit PA−PB1 in Influenza Virus RNA Polymerase: Insights from Molecular Dynamics Simulation and Free Energy Calculation

Huanxiang Liu[*,†,‡] **and Xiaojun Yao**[‡,§]

School of Pharmacy, State Key Laboratory of Applied Organic Chemistry, and Department of Chemistry, Lanzhou University, Lanzhou 730000, China

Received May 13, 2009; Revised Manuscript Received October 5, 2009; Accepted November 2, 2009

Abstract: The emergence of the extremely aggressive influenza recently has highlighted the urgent need for new effective treatments. The influenza RNA-dependent RNA polymerase (RdRp) heterotrimer including PA, PB1 and PB2 has crucial roles in viral RNA replication and transcription. The highly conserved PB1 binding site on PA can be considered as a novel potential drug target site. The interaction between PB1 binding site and PA is crucial to many functions of the virus. In this study, to understand the detailed interaction profile and to characterize the binding hot spots in the interactions of the PA−PB1 complex, an 8 ns molecular dynamics simulation of the subunit PA−PB1 combined with MM-PBSA (molecular mechanics Poisson−Boltzmann surface area), MM-GBSA (molecular mechanics generalized Born surface area) computations and virtual alanine scanning were performed. The results from the free energy decomposition indicate that the intermolecular van der Waals interaction and the nonpolar solvation term provide the driving force for binding process. Through the pair interaction analysis and virtual alanine scanning, we identified the binding hot spots of PA and the basic binding motif of PB1. This information can provide some insights for the structure-based RNA-dependent RNA polymerase inhibitors design. The identified binding motif can be used as the starting point for the rational design of small molecules or peptide mimics. This study will also lead to new opportunities toward the development of new generation therapeutic agents exhibiting specificity and low resistance to influenza virus.

Keywords: Influenza virus RNA polymerase; protein−protein interaction; molecular dynamics simulation; free energy calculation; structure-based drug design

1. Introduction

The outbreaks of the extremely aggressive influenza A H5N1 and H1N1, and the human fatalities they have already caused, have heightened the awareness of both general population and governments to the threat of influenza virus.[1] Considerable efforts worldwide have been pursued to discover novel therapeutic agents against all types of influenza to control avian or human pandemic of influenza A. For the development of anti-influenza virus drugs, elucidation of the underlying mechanisms of viral replication is critical.[2] Influenza A virus carries eight negative-strand RNA genome segments, each bound by the virally encoded RNA-dependent RNA polymerase (RdRp) complex, which plays numerous essential roles in viral replication and pathogenesis.[3,4] The precise roles of each subunit are still being actively investigated.[5−14]

* Author to whom correspondence should be addressed. Current address: School of Pharmacy, Lanzhou University, Lanzhou 730000, China. E-mail: hxliu@lzu.edu.cn. Tel: +86-931-8915686. Fax: +86-931-8915686.

† School of Pharmacy.

‡ State Key Laboratory of Applied Organic Chemistry.

§ Department of Chemistry.

(1) Taubenberger, J. K.; Reid, A. H.; Lourens, R. M.; Wang, R.; Jin, G.; Fanning, T. G. Characterization of the 1918 influenza virus polymerase genes. *Nature* **2005**, *437*, 889–893.

10.1021/mp900131p
Published on Web 11/02/2009

The subunit interactions in RNA polymerase have been characterized by extensive mutagenesis, showing that the N-terminal tip of PB1 binds to the C terminus of PA,[2,5−8] and the loss of PA abolishes RNA polymerase activity and viral replication.[9] PA is involved in the assembly of functional polymerase complex, cap binding and virion RNA (vRNA) promoter binding.[9,10] The carboxyl-terminal region of PA is thought to bind to PB1 for complex formation and nuclear transport.[6,11] Previous reports have also shown that the PB1 N-terminal 25 residues (PB1N) specifically bind to the C-terminal region of PA.[2,6,12] Therefore, the PA−PB1 interaction surface is a potential drug target not being exploited currently.

Most of the current influenza drugs target either hemeagglutinin (HA) or neuraminidase (NA), two major antigens present at the virion surface.[15] Sixteen different HA subtypes and nine different NA subtypes have been identified.[16] Oseltamivir (sold as Tamiflu) and zanamivir (Relenza), for example, are NA inhibitors which prevent viral particles being released from infected cells.[17−20] Billions of pounds worth of oseltamivir were stockpiled worldwide in response to the last Asian and the recent worldwide epidemics, but resistant influenza is already emerging. The anti-influenza drug amantadine targets M2 protein, a viral proton channel.[21,22] However, a single residue change is sufficient to confer resistance, which has risen sufficiently to render the drug useless against many strains. Both oseltamivir and amantadine target proteins with a single known function and substantial sequence variation between viral strains. So, there exists ample scope to develop new lead molecules disrupting other processes in the viral life cycle. The highly conserved PB1 binding site on PA, whose conservation can be found by the sequence alignment of PA around PB1 binding site for different types of influenza (Figure 1) generated by the PRALINE sequence alignment toolbox,[23] may have considerable potential as a drug target site, given that the interaction is crucial to many viral functions.

(2) Ghanem, A.; Mayer, D.; Chase, G.; Tegge, W.; Frank, R.; Kochs, G.; García-Sastre, A.; Schwemmle, M. Peptide-mediated interference with influenza A virus polymerase. *J. Virol.* **2007**, *81*, 7801–7804.

(3) Deng, T.; Sharps, J. L.; Brownlee, G. G. Role of the influenza virus heterotrimeric RNA polymerase complex in the initiation of replication. *J. Gen. Virol.* **2006**, *87*, 3373–3377.

(4) Plotch, S. J.; Bouloy, M.; Ulmanen, I.; Krug, R. M. A unique cap (m7GpppXm)- dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* **1981**, *23*, 847–858.

(5) Zurcher, T.; de la Luna, S.; Sanz-Ezquerro, J. J.; Nieto, A.; Ortin, J. Mutational analysis of the influenza virus A/Victoria/3/75 PA protein: studies of interaction with PB1 protein and identification of a dominant negative mutant. *J. Gen. Virol.* **1996**, *77*, 1745–1749.

(6) Perez, D. R.; Donis, R. O. Functional analysis of PA binding by influenza A virus PB1: effects on polymerase activity and viral infectivity. *J. Virol.* **2001**, *75*, 8127–8136.

(7) Ohtsu, Y.; Honda, Y.; Sakata, Y.; Kato, H.; Toyoda, T. Fine mapping of the subunit binding sites of influenza virus RNA polymerase. *Microbiol. Immunol.* **2002**, *46*, 167–175.

(8) Liang, Y.; Huang, T.; Ly, H.; Parslow, T. G.; Liang, Y. Mutational Analyses of Packaging Signals in Influenza Virus PA, PB1,and PB2 Genomic RNA Segments. *J. Virol.* **2008**, *82*, 229–236.

(9) Kawaguchi, A.; Naito, T.; Nagata, K. Involvement of influenza virus PA subunit in assembly of functional RNA polymerase complexes. *J. Virol.* **2005**, *79*, 732–744.

(10) Hara, K.; Schmidt, F. I.; Crow, M.; Brownlee, G. G. Amino acid residues in the N-terminal region of the PA subunit of influenza A virus RNA polymerase play a critical role in protein stability, endonuclease activity, cap binding, and virion RNA promoter binding. *J. Virol.* **2006**, *80*, 7789–7798.

(11) Toyoda, T.; Adyshev, D. M.; Kobayashi, M.; Iwata, A.; Ishihama, A. Molecular assembly of the influenza virus RNA polymerase: determination of the subunit-subunit contact sites. *J. Gen. Virol.* **1996**, *77*, 2149–2157.

(12) Gonzalez, S.; Zurcher, T.; Ortin, J. Identification of two separate domains in the influenza virus PB1 protein involved in the interaction with the PB2 and PA subunits: a model for the viral RNA polymerase structure. *Nucleic Acids Res.* **1996**, *24*, 4456–4463.

(13) Yuan, P.; Bartlam, M.; Lou, Z.; Chen, S.; Zhou, J.; He, X.; Lv, Z.; Ge, R.; Li, X.; Deng, T.; Fodor, E.; Rao, Z.; Liu, Y. Crystal structure of an avian influenza polymerase PA_N reveals an endonuclease active site. *Nature* **2009**, *458*, 909–914.

(14) Dias, A.; Bouvier, D.; Crépin, T.; McCarthy, A. A.; Hart, D.; Baudin, J. F.; Cusack, S.; Ruigrok, R. W. H. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **2009**, *458*, 914–918.

(15) Hsieh, H. P.; Hsu, J. T. Strategies of development of antiviral agents directed against influenza virus replication. *Curr. Pharm. Des.* **2007**, *13*, 3531–3542.

(16) World Health Organization. A revision of the system of nomenclature for influenza viruses: a WHO memorandum. *Bull. W.H.O.* **1980**, *58*, 585−591.

(17) Kim, C. U.; Lew, W.; Williams, M. A.; Liu, H.; Zhang, L.; Swaminathan, S.; Bischofberger, N.; Chen, M. S.; Mendel, D. B.; Tai, C. Y.; Laver, W. G.; Stevens, R. C. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: Design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J. Am. Chem. Soc.* **1997**, *119*, 681–690.

(18) von Itzstein, M.; Wu, W. Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Van Phan, T.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Ryan, D. M.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* **1993**, *363*, 418–423.

(19) Russell, R. J.; Haire, L. F.; Stevens, D. J.; Collins, P. J.; Lin, Y. P.; Blackburn, G. Michael.; Hay, A. J.; Gamblin, S. J.; Skehel, J. J. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. *Nature* **2006**, *443*, 45–49.

(20) Liu, Y.; Zhang, J.; Xu, W. Recent progress in rational drug design of neuraminidase inhibitors. *Curr. Med. Chem.* **2007**, *14*, 2872–2891.

(21) Wang, C.; Takeuchi, K.; Pinto, L. H.; Lamb, R. A. Ion channel activity of influenza virus M2 protein: characterization of the amantadine block. *J. Virol.* **1993**, *67*, 5585–5594.

(22) Stouffer, A. L.; Acharya, R.; Salom, D.; Levine, A. S.; Costanzo, L. D.; Soto, C. S.; Tereshko, V.; Nanda, V.; Stayrook, S.; DeGrado, W. F. Structural basis for the function and inhibition of an influenza virus proton channel. *Nature* **2008**, *451*, 596–599.

Figure 1. The sequence alignment of PA around PB1 binding site for different types of influenza generated by the Praline sequence alignment toolbox.

Recently, two groups from China and Japan reported the crystal structure of a principal part of the influenza polymerase PA protein in complex with an inhibitory peptide of PB1 (PB1 amino terminus, $PB1_N$) successively.[24,25] The availability of the PA−$PB1_N$ structure offers a starting point for further investigation into the structure and function of the influenza virus polymerase. Although these high-resolution crystal structures provide plentiful atomic-level structural information, the understanding of the detailed interaction and binding energy profile in the PA−PB1 complex is still very limited. An increasing comprehension of these elements can provide useful information for modulating the stability of the PA−PB1 complex. This can lead to new opportunities toward the development of a new generation of therapeutic agents exhibiting specificity and low resistance for H5N1 and H1N1 to overcome a serious drawback of the current influenza virus inhibitors.

Structural and energetic analysis of protein−protein interactions is a challenging endeavor both experimentally and theoretically. Experimental probing of protein−protein binding affinity entails the development of sophisticated yet accurate assays. Despite an increasing number of crystallographically determined protein−protein complexes, thermodynamic experiments, and mutational studies, the understanding of the energetics and dynamics for the binding of two proteins at the atomic level is still limited.[26,27] In recent years, molecular dynamics (MD) simulations have been used

(23) Simossis, V. A.; Heringa, J. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic Acids Res.* **2005**, *33*, W289−W294.

(24) He, X.; Zhou, J.; Bartlam, M.; Zhang, R.; Ma, J.; Lou, Z.; Li, X.; Li, J.; Joachimiak, A.; Zeng, Z.; Ge, R.; Rao, Z.; Liu, Y. Crystal structure of the polymerase PA_C-$PB1_N$ complex from an avian influenza H5N1 virus. *Nature* **2008**, *454*, 1123–1126.

(25) Obayashi, E.; Yoshida, H.; Kawai, F.; Shibayama, N.; Kawaguchi, A.; Nagata, K.; Tame, J. R. H.; Park, S.-Y. The structural basis for an essential subunit interaction in influenza virus RNA polymerase. *Nature* **2008**, *454*, 1127–1131.

(26) Stites, W. E. Protein−Protein Interactions: Interface Structure, Binding Thermodynamics, and Mutational Analysis. *Chem. Rev.* **1997**, *97*, 1233–1250.

widely to study protein−protein interactions.[27−30] MD simulations can provide not only plentiful dynamic structural information on protein complex structures in solution but also a wealth of energetic information, including binding free energy between protein partners. The most rigorous MD-based approaches to estimate binding free energy include free energy perturbation and thermodynamic integration methods. Because of long convergence time, they are computationally intensive and not often used on large systems such as protein−protein complexes.

A more commonly used and tractable approach is molecular mechanics Poisson−Boltzmann surface area (MM-PBSA) method.[31,32] In this first-principle based method, gas-phase energy calculated using conventional molecular mechanics force fields such as AMBER[33,34] is combined with a continuum model of solvation that includes a surface-area based nonpolar contribution and a polar solvation free energy calculated with the Poisson−Boltzmann (PB) model. Solute entropy can be incorporated from statistical thermodynamics with normal-mode analysis. Recently, there has been an increased interest in faster molecular mechanics generalized Born surface area (MM-GBSA), a variant of MM-PBSA, which replaces the PB electrostatics with the generalized Born (GB) approximate model of electrostatics in water.[31,35,36] Another advantage of MM-GBSA is that it utilizes a fully pairwise potential useful for decomposing the total binding free energy into atomic/group contributions, in a structurally nonperturbing formalism.[29] By decomposing total binding free energy to individual residue, we can obtain detailed binding energy such as the contribution of individual residue on the binding surface.

As an extension of the MM-PBSA approach, computational alanine scanning was also developed to estimate the contribution of individual residue to overall protein−protein binding free energies. Computational alanine scanning can be used to measure the effect after deleting an amino acid side chain beyond the Cβ carbon atom on the affinity of a protein−protein complex. Individual substitution of an amino acid residue with alanine yields a map about the critical interactions in the interface. This kind of computational alanine scanning uses a simple free energy function to calculate the effect of alanine mutation on the binding free energy of a protein−protein complex. The advantage of this method compared with the experimental alanine scanning is its high speed, which makes large-scale applications to all protein−protein interfaces with known structures feasible. Identification of hot spots provides a starting point for the design of the inhibitors targeting protein−protein interactions. At present, computational alanine scanning has been proved to be an effective and reliable method. It can be used as an alternative of experimental mutagenesis toward identification of binding affinity hot spots at protein−protein interfaces.[37−40] Despite its advantages, computational alanine scanning also has one obvious limitation. It cannot be used to add a larger group than those simulated because of highly unfavorable van der Waals interactions caused by atomic position overlap. In detail, glycine cannot be mutated to alanine during computational alanine scanning process.

In this study, molecular dynamics (MD) simulations were used to obtain dynamic structural information about the PA−PB1 complex. We aimed to describe the interaction features of the protein−protein complex much better than just the static crystal structure. Based on the obtained MD trajectory, MM-PBSA and MM-GBSA were applied to

(27) Gohlke, H.; Case, D. A. Converging Free Energy Estimates: MM-PB(GB)SA Studies on the Protein-Protein Complex Ras-Raf. *J. Comput. Chem.* **2004**, *25*, 238–250.
(28) Zoete, V.; Meuwly, M.; Karplus, M. Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 79–93.
(29) Gohlke, H.; Kiel, C.; Case, D. A. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol.* **2003**, *330*, 891–913.
(30) Cui, Q.; Sulea, T.; Schrag, J. D.; Munger, C.; Hung, M.-N.; Naïm, M.; Cygler, M.; Purisima, E. O. Molecular Dynamics-Solvated Interaction Energy Studies of Protein-Protein Interactions: The MP1-p14 Scaffolding Complex. *J. Mol. Biol.* **2008**, *379*, 787–802.
(31) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* **1998**, *120*, 9401–9409.
(32) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: combining molecular Molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897.
(33) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
(34) Lee, M. C.; Duan, Y. Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized Born solvent model. *Proteins: Struct., Funct., Bioinform.* **2004**, *55*, 620–634.
(35) Tsui, V.; Case, D. A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **2000**, *56*, 275–291.
(36) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the generalized Born model suitable for macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712–3720.
(37) Kortemme, T.; Kim, D. E.; Baker, D. Computational alanine scanning of protein-protein interfaces. Sci. *STKE* **2004**, *2004*, p12.
(38) Kortemme, T.; Baker, D. A simple physical model for binding energy hot spots in protein-protein complexes. *Proc. Natl Acad. Sci. U.S.A.* **2002**, *99*, 14116–14121.
(39) Massova, I.; Kollman, P. A. Computational alanine scanning to probe protein-protein interactions: a novel approach to evaluate binding free energies. *J. Am. Chem. Soc.* **1999**, *121*, 8133–8143.
(40) Huo, S.; Massova, I.; Kollman, P. A. Computational alanine scanning of the 1:1 human growth hormone-receptor complex. *J. Comput. Chem.* **2002**, *23*, 15–27.

calculate the binding free energy and understand the detailed interaction profile and to characterize the binding hot spots by examining various contributions to binding energy of the PA−PB1 complex. Furthermore, the virtual alanine scanning was also used to verify binding hot spots found by pair interaction analysis based on the results of MM-GBSA. The detailed interaction profile from MD simulation and MM-PBSA/MM-GBSA calculation will help us understand the critical features determining the PA−PB1 binding process and give some insights for structure-based design of inhibitors.

At present, there are a variety of computational methods to identify and evaluate the surface binding pockets, such as LIGSITE,[41] LIGSITE^{CSC},[42] FINDSITE,[43] PocketPicker[44] and SITEMAP.[45] Among these methods, LIGSITE^{CSC} is widely acceptable, and FINDSITE is a new promising binding pocket prediction method.[43] These pocket prediction approaches provide useful tools for identification of potential protein binding-pockets and provide a starting point for structure-based function identification and structure-based drug discovery. The advantages of these methods are not dependent on the accurate crystal structure and are very fast. Compared with these approaches, the method MD combined with MM-PBSA and MM-GBSA needs a crystal or homology structure as the starting point of molecular dynamics. But it can provide more accurate and detailed information about the binding site such as the contribution of each residue, which is very important for structure-based drug design and impossible to be obtained from other binding site prediction methods.

2. Experimental Section

2.1. Molecular System. There are currently only two resolved crystal structures available for the PA−PB1 complex, with PDB codes 3CM8 at 2.9 Å and 2ZNL at 2.3 Å resolution.[24,25] When we began this study, only the crystal structure 3CM8 was released, and this structure was then used as the initial structure of our simulation. By aligning two crystal structures, we found that they were almost the same except for the conformation of some loops in PA section. These loops are far from the PA−PB1 interface, and there is no proof that they have an impact on the function of the PA and PA−PB1 binding surface. The crystal structure

Figure 2. Ribbon representation of the crystal structure of PA_C−$PB1_N$ complex. $PB1_N$ is represented in magenta, and PA_C is in green. This figure was created by using PyMol.

of the PA−PB1 complex (Figure 2 generated by Pymol[46]) includes an N-terminally truncated PA domain covering the residues 257−716 (termed PA_C) and the N-terminal of PB1 domain (the first 15 residues from Met1 to Gln15, called as $PB1_N$). For PA_C, the missing residues were not added during the molecular dynamics simulation. For $PB1_N$, to remove the unnatural negative charge of C terminal, the *C*-termini were capped with an *N*-methyl group (NME). All the crystal water molecules were kept during the following molecular dynamics simulation.

2.2. Molecular Dynamics Simulation. All the calculations of energy minimization and system equilibration protocols were performed with the AMBER10 suite of the program[47] using the AMBER99 force field.[48] To keep the whole system neutral, one Cl^- counterion was added. Each system was then solvated using atomistic TIP3P water[49] in a cubic box with at least 10 Å distance around the complex. In summary, the system was constructed using periodic boundary conditions with a rectangular box of dimension 102 Å × 109 Å × 79 Å, consisting of a PA_C−$PB1_N$ complex,

(41) Hendlich, M.; Rippmann, F.; Barnickel, G. LIGSITE: Automatic and efficient detection of potential small molecule-binding sites in proteins. *J. Mol. Graphics Modell.* **1997**, *15*, 359–363.
(42) Huang, B.; Schröer, M. LIGSITE^{csc}: predicting ligand binding sites using the Connolly surface and degree of conservation. *BMC Struct. Biol.* **2006**, *6*, 19–29.
(43) Brylinski, M.; Skolnick, J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. *Proc. Natl Acad. Sci. U.S.A.* **2008**, *105*, 129–134.
(44) Weisel, M.; Proschak, E.; Schneider, G. PocketPicker: analysis of ligand binding-sites with shape descriptors. *Chem. Cent. J.* **2007**, *1*, 1–17.
(45) Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug. Des.* **2007**, *69*, 146–148.
(46) DeLano, W. L. The PyMOL Molecular Graphics System; DeLano Scientific: Palo Alto, CA, 2002. http://www.pymol.org.
(47) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
(48) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–25.
(49) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

49 crystallographic water molecules, 1 counterion and 22000 TIP3P water molecules.

The energy minimization was first conducted using the steepest descent method in AMBER10 for 5000 iterations with a force constant of 10 kcal/mol·Å^2 applied to all atoms. This was followed by a further energy minimization with steepest descent method switched to conjugate gradient every 500 steps totally for 2500 steps without any constraints. Long-range Coulombic interactions were handled using the particle mesh Ewald (PME) summation.[50] For the equilibration and subsequent production run, the SHAKE algorithm[51] was employed on all atoms covalently bonded to a hydrogen atom, allowing for an integration time step of 2 fs. The system was gently annealed from 0 to 300 K over a period of 50 ps and equilibrated for 50 ps at a temperature of 300 K using a Langevin thermostat with a coupling coefficient of 1.0/ps with a gradual reduction restraint from 5.0 to 2.0 kcal/mol·Å^2 on the complex. All subsequent stages were carried out in the isothermal isobaric (NPT) ensemble using a Berendsen barostat[52] with a target pressure of 1 bar and a pressure coupling constant of 2.0 ps. The systems were again equilibrated for 500 ps while maintaining the force constants on the restrained atoms. The production phase of the simulation was run under the same conditions for a total of 8 ns. Here, the same condition refers to the residues 257−388 far from binding surface being restrained as above with a fixed harmonic potential of 2.0 kcal/mol·Å^2 during the production simulations. Coordinate trajectories were recorded every 1 ps throughout all the equilibration and production runs.

2.3. Binding Free Energy Calculation. The first step of MM-PBSA and MM-GBSA methods is the generation of multiple snapshots from an MD trajectory of the protein−protein complex, stripped of water molecules and counterions. Snapshots, equally spaced at 10 ps intervals, were extracted from the MD production runs, giving 800 snapshots for 8 ns. For each snapshot, the free energy is calculated for each molecular species (complex, protein 1, and protein 2), and the binding free energy is computed as the difference

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{protein1}} - G_{\text{protein2}}$$

The free energy, G, for each species can be calculated by the following scheme using the MM-PBSA and MM-GBSA methods:[31,32]

$$G = E_{\text{gas}} + G_{\text{sol}} - TS$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{ele}} + E_{\text{vdw}}$$

$$E_{\text{int}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}}$$

$$G_{\text{sol}} = G_{\text{PB(GB)}} + G_{\text{nonpolar}}$$

$$G_{\text{nonpolar}} = \gamma\,\text{SAS}$$

Here, E_{gas} is the gas-phase energy; E_{int} is the internal energy; E_{bond}, E_{angle}, and E_{torsion} are the bond, angle, and torsion energies, respectively; and E_{ele} and E_{vdw} are the Coulomb and van der Waals energies, respectively. E_{gas} was calculated using the AMBER99 molecular mechanics force field.[48] G_{sol} is the solvation free energy and can be decomposed into polar and nonpolar contributions. $G_{\text{PB(GB)}}$ is the polar solvation contribution calculated by solving the PB and GB equations.[31,32] Dielectric constants for solute and solvent were set to 1 and 80, respectively. G_{nonpolar} is the nonpolar solvation contribution and was estimated by the solvent accessible surface area (SAS) determined using a water probe radius of 1.4 Å. The surface tension constant γ was set to 0.0072 kcal/mol/Å^2.[53] T and S are the temperature and the total solute entropy, respectively. Vibrational entropy contributions can be estimated by many methods such as normal-mode analysis. However, we did not calculate the vibrational entropy contributions here since our aim is not to obtain the absolute Gibbs energy but to identify the hotspot residues of binding surface and the detailed interaction features.

To obtain a detailed view of the protein−protein binding, MM-GBSA was used to decompose the interaction energies to each residue involved in the interaction by considering molecular mechanics and solvation energies without consideration the contribution of entropies.

2.4. Virtual Alanine Scanning Mutagenesis. In order to identify binding affinity hot spots on the PA_C−$PB1_N$ complex, we performed "single-structure" computational alanine scanning of the complex, based on 800 snapshots extracted from MD simulation. The MM-GBSA and MM-PBSA methods[31,32] were used to estimate the relative binding free energies. The basic process includes two steps, the generation of mutated snapshots and the calculation of the binding free energy difference between the wild type and mutant complex. First, we need to generate the snapshots from the wild type trajectory every 10 ps for alanine scanning. Then, mutations to alanine are performed only on selected residues in the active site. Alanine mutations were generated by truncation of residues after the Cβ and adding hydrogen in the same direction as the Cγ. Partial charges for the mutated residue were then changed to those of alanine. Among the residues, glycine cannot be mutated to alanine. In addition, the underlying approximation of the single trajectory mutation protocol is that the mutant and wild type undergo similar conformational changes, and that local side chain reorganizations are small perturbations relative to the

(50) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–9593.

(51) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(52) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–690.

(53) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.

alanine mutation itself. Generally, proline has a severe impact on the secondary structure. Thus, in our investigation, alanine scanning was not performed for proline. After obtaining the wild type snapshots and mutated snapshots, the MM-PBSA method was used to estimate the relative binding free energies of the mutant complex.

3. Results

3.1. MM-PBSA and MM-GBSA. The components of the binding free energies for the PA_C-PB1_N complexes were calculated with the MM-PBSA and MM-GBSA methods. The snapshot structures used for these calculations were extracted from the equilibrated portions of 8 ns molecular dynamics (MD) trajectory carried out in the presence of explicit solvent molecules. It should be noted that long time simulations are rarely used to evaluate the binding free energy using surface area continuum solvent methods. However, several long time simulation studies showed that averaging over such a long time window was necessary because of the large fluctuations observed in the computed free energies.[30,54−56] Here, the equilibration of the MD trajectories was monitored from the convergence of the plots of the root-mean-square deviation of Cα carbon atoms (Cα-rmsd) for the binding interface from their initial configuration (Figure 3a). It can be seen from Figure 3a that the relative fluctuation for the rmsd of the interface is very small after the previous initial equilibration. To further test the convergence of the simulation, the time evolution of various free energy components for the complex was also monitored (Figure 3b). As can be seen from Figure 3b, the fluctuation of the calculated binding free energy by the MM-PBSA and MM-GBSA methods is very small along the whole trajectory. Although the enthalpy components for Coulomb contributions between different snapshots have a high fluctuation, they are compensated by corresponding fluctuations in the polar solvation contribution. As a result, the overall polar interaction (including Coulomb contribution and polar solvation contribution) is stable along the last 8 ns trajectory. From the structural and energetic analysis, we can see that the system is stable from both the rmsd and binding energies during the last 8 ns molecular dynamics simulation.

The detailed contribution of various energy components based on the MM-PBSA and MM-GBSA methods is given in Table 1. One of the advantages of the MM-PBSA and MM-GBSA approaches is that it enables a decomposition of the free energy into identifiable contributions. Thus, polar

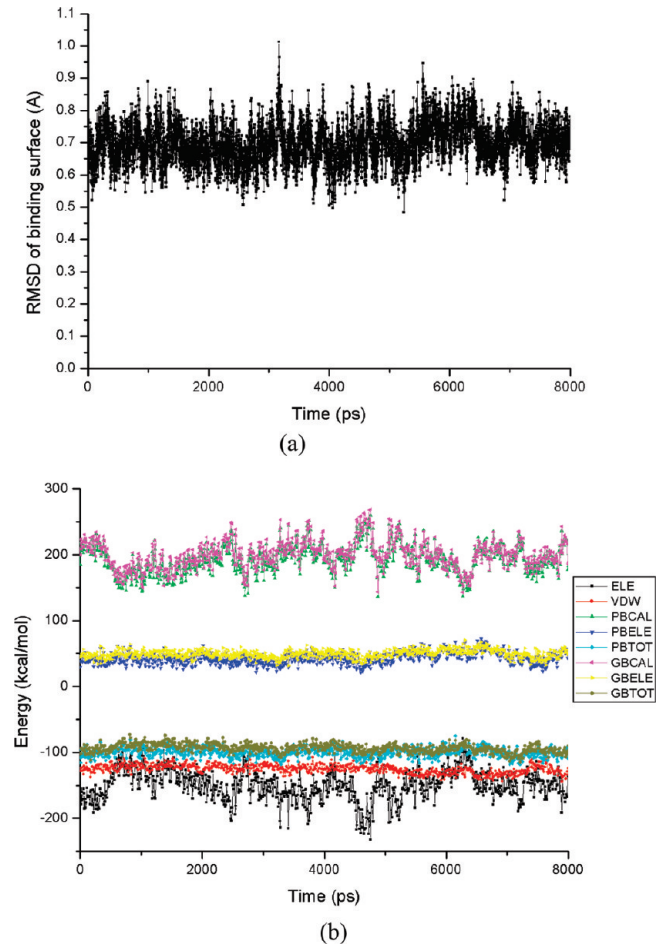

Figure 3. The monitoring for the equilibration of the MD trajectories: (a) time series of the rms deviation from the initial structure for the complex interface; (b) time evolution of binding free energy components for the complex.

($\Delta G_{ele,PB} = \Delta E_{ele} + \Delta G_{sol-pol,PB}$, $\Delta G_{ele,GB} = \Delta E_{ele} + \Delta G_{sol-pol,GB}$), nonpolar or hydrophobic energies ($\Delta G_{np} = \Delta E_{vdw} + \Delta G_{sol-np}$) were analyzed separately. As can be seen in Table 1, the intermolecular van der Waals interaction and the nonpolar solvation term provide the driving force for binding. The highly favorable nonpolar binding free energy probably reflects the great number of favorable interactions from the various hydrophobic pockets that cover the surface of the binding site between PA_C and $PB1_N$. In contrast to the nonpolar components, the polar interactions were found to make unfavorable contribution to the binding energy with $\Delta G_{ele,PB}$ of 43.27 kcal/mol and $\Delta G_{ele,GB}$ of 49.54 kcal/mol based on the MM-PBSA and MM-GBSA methods, respectively. It is surprising that the polar interactions make a weak contribution to binding energy since there are several hydrogen bonds between the peptide ligand of $PB1_N$ and PA_C domain. However, decomposition of the polar interaction contribution into its Coulombic and solvation components showed that indeed the direct intermolecular electrostatic interactions were always favorable to the binding but their contributions could not compensate the large desolvation penalties associated with the binding event, thereby always leading to an unfavorable contribution. This compensation

(54) Pearlman, D. A. Evaluating the molecular mechanics poisson-boltzmann surface area free energy method using a congeneric series of ligands to p38 MAP kinase. *J. Med. Chem.* **2005**, *48*, 7796–7807.

(55) Stoica, I.; Kashif Sadiq, S.; Coveney, P. V. Rapid and Accurate Prediction of Binding Free Energies for Saquinavir-Bound HIV-1 Proteases. *J. Am. Chem. Soc.* **2008**, *130*, 2639–2648.

(56) Fratev, F.; Jónsdóttir, S. Ó.; Mihaylova, E.; Pajeva, I. Molecular Basis of Inactive B-RAF WT and B-RAF V600E Ligand Inhibition, Selectivity and Conformational Stability: An *in Silico* Study. *Mol. Pharmaceutics* **2009**, *6*, 144–157.

Table 1. Binding Free Energy Components of the PA_C-PB1_N Complex[a]

contribution[b]	PA_C-PB1_N	PA_C	$PB1_N$	delta[c]
E_{ele}	−13000.6	−12485.7	−363.95	−150.95
E_{vdw}	−1935.05	−1805.78	−3.75	−125.52
E_{int}	9907.91	9571.08	336.83	0
E_{gas}	−5027.76	−4720.42	−30.87	−276.47
G_{sol-np}	171.81	175	14.63	−17.81
$G_{sol-pol,PB}$	−6089.96	−6050.37	−233.81	194.22
$G_{sol,PB}$	−5918.15	−5875.38	−219.18	176.41
$G_{ele,PB}$	−19090.6	−18536.1	−597.75	43.27
$H_{total,PB}$	−10945.9	−10595.8	−250.05	−100.06
$G_{sol-pol,GB}$	−6148.16	−6112.42	−236.24	200.5
$G_{sol,GB}$	−5976.34	−5937.42	−221.61	182.69
$G_{ele,GB}$	−19148.8	−18598.1	−600.19	49.54
$H_{total,GB}$	−11004.1	−10657.8	−252.48	−93.78

[a] All values are given in kcal/mol. [b] E_{ele}: Coulombic energy. E_{vdw}: van der Waals energy. E_{int}: internal energy. $E_{gas} = E_{ele} + E_{vdw} + E_{int}$. G_{sol-np}: nonpolar solvation free energy. $G_{sol-pol,PB}$: polar solvation free energy based on the PBSA method. $G_{sol,PB} = G_{sol-np} + G_{sol-pol,PB}$. $G_{ele,PB} = E_{ele} + G_{sol-pol,PB}$. $H_{total,PB} = E_{gas} + G_{sol,PB}$. $G_{sol-pol,GB}$: polar solvation free energy based on the GBSA method. $G_{sol,GB} = G_{sol-np} + G_{sol-pol,GB}$. $G_{ele,GB} = E_{ele} + G_{sol-pol,GB}$. $H_{total,GB} = E_{gas} + G_{sol,GB}$. [c] Contribution($PA_C-PB1_N$) − contribution($PA_C$) − contribution($PB1_N$).

phenomenon has been previously observed in several studies of protein−ligand interactions in solution.[57−60]

Since our purpose here is to use MM-PBSA calculations to reveal the sources governing binding through a decomposition of binding free energies, rather than to calculate the absolute binding energy, we did not calculate the solute entropy contribution in this study and the binding energy shown here (−100.06 kcal/mol for MM-PBSA and −93.78 kcal/mol for MM-GBSA) is not the total free energy. In addition, it is well-known that the binding free energies calculated with MM-PBSA and MM-GBSA do not reproduce the absolute experimental values accurately; they have been shown to correlate with the experiment values well.[61,62]

3.2. Pair Interaction Energy Analysis. Pair interaction energy analysis and virtual alanine scanning were performed to detect the important interactions of the PA−PB1 binding process and identify the key residues of the protein−protein interaction interface. Here and in the whole context, the important interactions are the interactions primarily responsible for the binding process. The pair interaction decomposition results for PB1 and PA_C are given in Figure 4. From Figure 4a, it can be seen that several residues of the PA_C surface give a large contribution to binding energy with more than 2 kcal/mol free energy contribution. They include Trp706, Phe710, Asn412, Ile621, Gln623, Gln408, Gln670, Leu666, Pro620 and Gly622. In addition, some other residues also have an obvious contribution to the binding process of PA and PB1 with more than 1 kcal/mol free energy contribution. These residues are Trp619, Leu667, Met595, Phe411, Thr639 and Leu640. Residue energy decomposition results for PB1 are given in Figure 4b. From this figure, nine residues including Val3, Asn4, Pro5, Leu7, Leu8, Phe9, Leu10, Val12 and Ala14 have more than 2 kcal/mol free energy contribution.

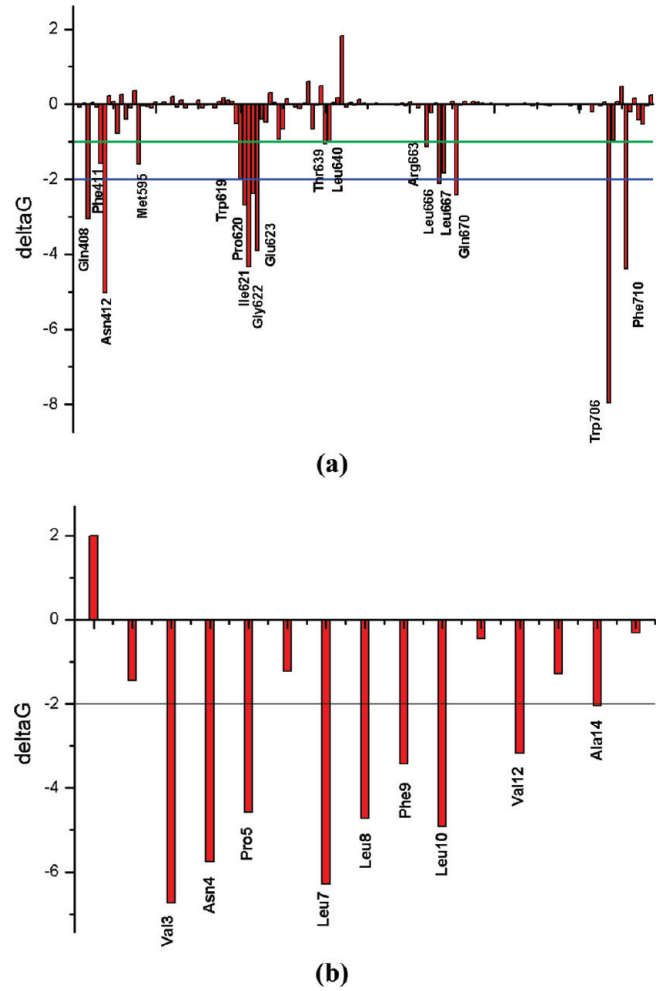

Figure 4. Pair interaction energy analysis between PA_C and PB1: (a) PA_C residues; (b) PB1 residues.

(57) Miyamoto, S.; Kollman, P. A. What determines the strength of noncovalent association of ligands to proteins in aqueous solution. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8402–8406.
(58) Wang, W.; Kollman, P. A. Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. *J. Mol. Biol.* **2000**, *303*, 567–582.
(59) Masukawa, K. M.; Kollman, P. A.; Kuntz, I. D. Investigation of neuraminidase-substrate recognition using molecular dynamics and free energy calculations. *J. Med. Chem.* **2003**, *46*, 5628–5637.
(60) Basdevant, N.; Weinstein, H.; Ceruso, M. Thermodynamic Basis for Promiscuity and Selectivity in Protein-Protein Interactions: PDZ Domains, a Case Study. *J. Am. Chem. Soc.* **2006**, *128*, 12766–12777.
(61) Kuhn, B.; Kollman, P. A. Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. *J. Med. Chem.* **2000**, *43*, 3786–3791.
(62) Wang, J.; Morin, P.; Wang, W.; Kollman, P. A. Use of MM-PBSA in Reproducing the Binding Free Energies to HIV-1 RT of TIBO Derivatives and Predicting the Binding Mode to HIV-1 RT of Efavirenz by Docking and MM-PBSA. *J. Am. Chem. Soc.* **2001**, *123*, 5221–5230.

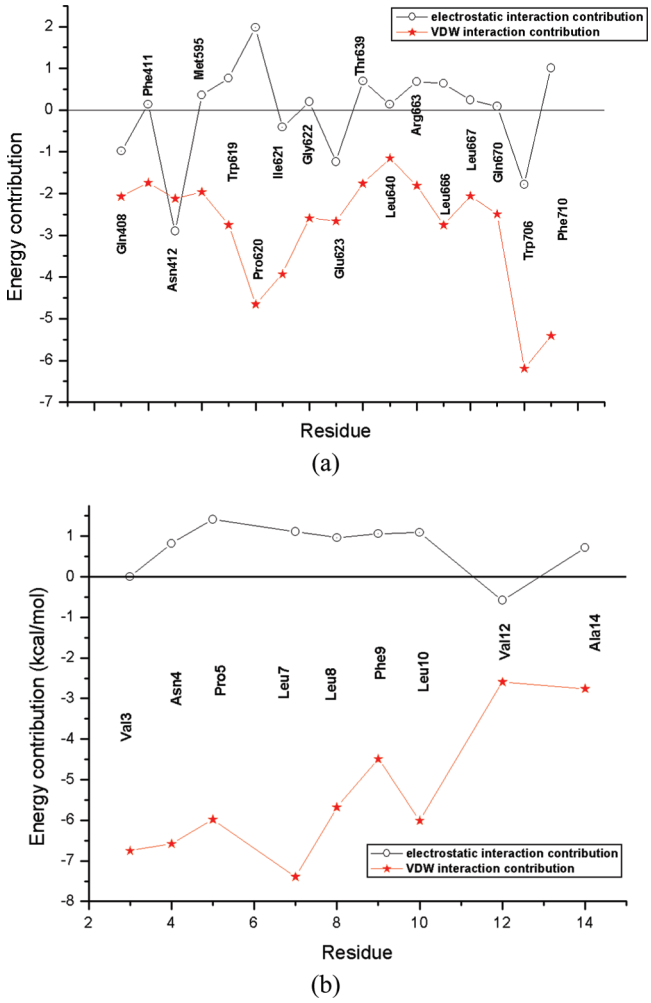

Figure 5. The electrostatic and VDW interaction contribution for the important residues of PA (a) and PB1 (b).

In order to identify the detailed contribution of each important residue, the energy contribution of each surface important residue is further decomposed into polar and van der Waals interaction. The polar section is the sum of ΔE_{ele} and $\Delta G_{sol-pol,GB}$. From Figure 5a and Figure 5b, it can be seen that most of the important residues both for PA and PB1 have a large van der Waals interaction contribution while the polar interaction has a relatively small influence especially for residues from PB1. But for several residues, the polar interactions also contribute obviously, for example, Gln408, Asn412, Glu623 and Trp706. Most of these polar interactions originate from the formation of hydrogen bonds. The detailed interaction network between PA and PB1 is given in Figure 6 produced by the Ligplot program.[63] From this figure, Gln408, Asn412, Glu623 and Trp706 form the important hydrogen bonds with PB1, whose importance can be proved by monitoring the occupation probability of these hydrogen bonds during the MD process.

(63) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **1995**, *8*, 127–134.

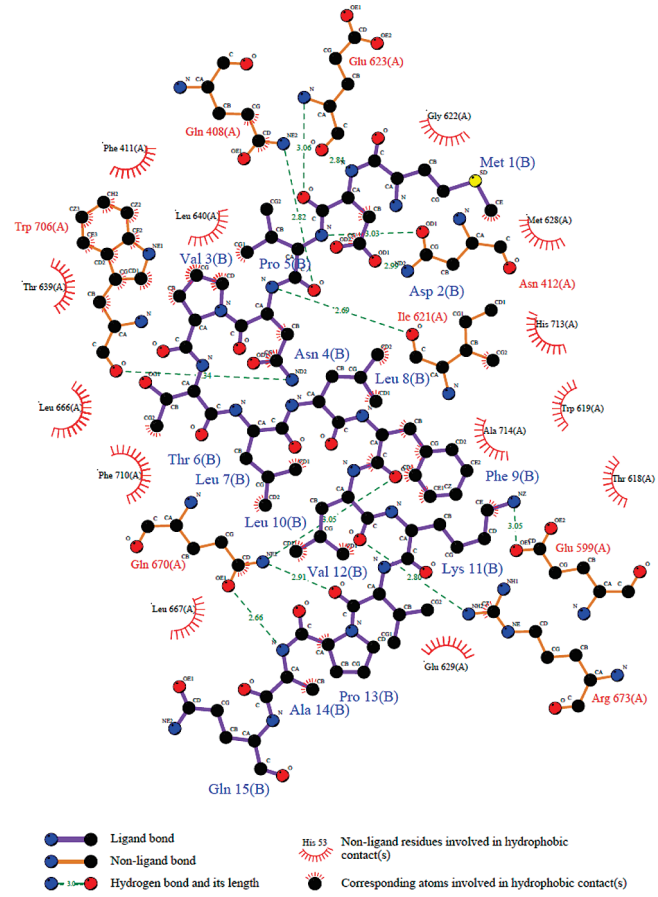

Figure 6. The detailed interaction network between the PA and PB1 produced by the Ligplot program.

3.3. Virtual Alanine Scanning Results. In order to verify the binding hot spots identified by pair interaction analysis, virtual alanine scanning was also performed for the residues with a contribution more than 1.0 kcal/mol from residue energy decomposition. The virtual alanine scanning results for PA and PB1 residues are given in Figure 7. By comparing Figure 7 and Figure 4, we can see that the pair interaction analysis and virtual alanine scanning give consistent results, which also indicates the reliability of our analysis results. By comparing the results of pair interaction energy analysis and virtual alanine scanning analysis, we can see that the residues Val3, Asn4, Pro5, Leu7, Leu8, Phe9 and Leu10 have 2 kcal/mol contributions in both methods. It also proves these residues are essential for PB1 binding to PA and can be considered as the binding motif of PB1.

4. Discussion

4.1. Key Features Determining the Binding of PA and PB1. *4.1.1. Nonpolar Interactions Dominate PA−PB1 Binding.* The results from free energy decomposition indicate that the intermolecular van der Waals interaction and the nonpolar solvation term provide the driving force for binding. The highly favorable nonpolar binding free energy probably reflects the great number of favorable interactions from the various hydrophobic pockets that cover the surface of the binding site between PA_C and $PB1_N$ (Figure 8). By drawing

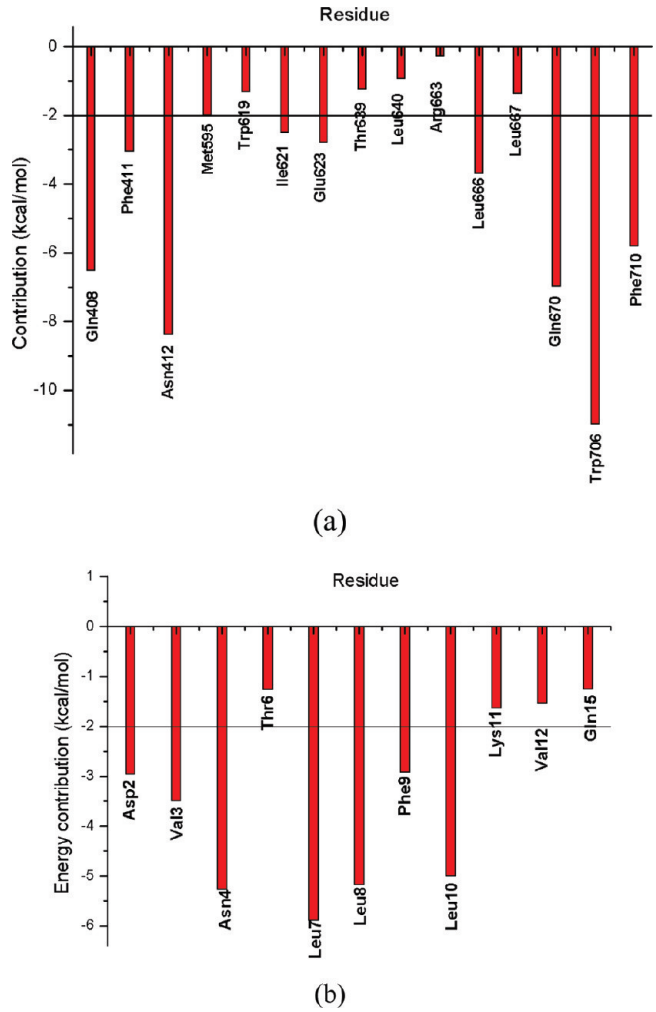

Figure 7. Virtual alanine scanning results for PA (a) and PB1 residues (b).

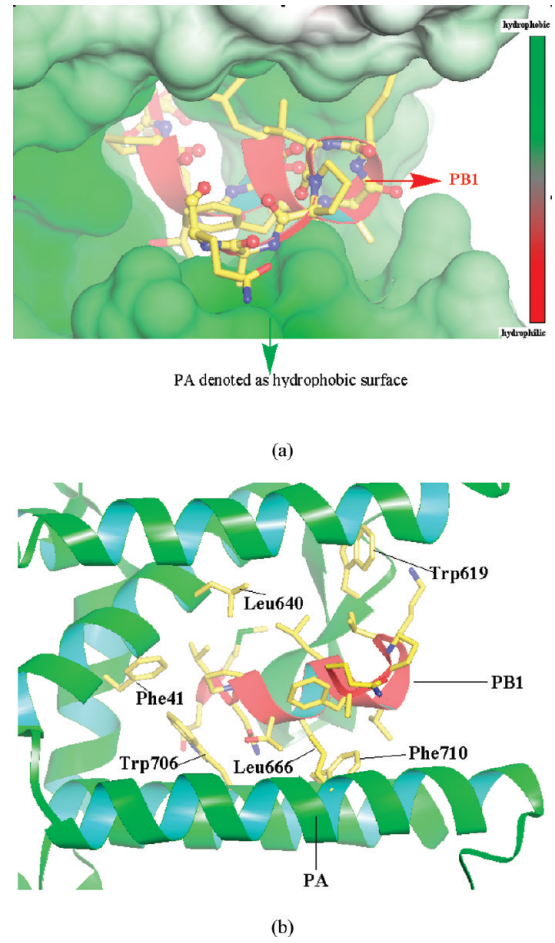

Figure 8. The interaction interface of PA−PB1: (a) the surface representation of PA with PB1 in the red cartoon; (b) the key hydrophobic residues of the PA surface.

the hydrophobic surface of PA (Figure 8a), we can see the whole binding interface of PA is full of hydrophobic features. From the detailed structural analysis of the PA−PB1 complex, there are three hydrophobic pockets. One hydrophobic pocket is composed of Trp706 and Phe411 from PA, which is responsible for the binding of Pro5 from PB1. The second hydrophobic pocket, formed by Phe710 and Leu666, plays an important role for the binding of Phe9 from PB1. Leu640, Val636, Met595 and Trp619 constitute the third hydrophobic pocket and have a close contact with Leu8 of PB1.

Since the PA−PB1 binding pocket is full of hydrophobic characteristics and is composed of many hydrophobic residues (Figure 8b), the dominance of nonpolar interactions in PA−PB1 binding is not surprising. Especially for PB1, the experimental binding motif consists of all hydrophobic residues PLLFL. At the same time, nonpolar interactions have been shown to dominate the thermodynamics of protein−ligand recognition for many ligand classes including small ligands,[57,64] protein−protein interaction domains such as PDZ domain,[60] and even highly charged ligands such as DNA molecules.[65]

4.1.2. Polar Interactions Do Not Contribute Significantly to PA−PB1 Interactions. In contrast, the results indicated a very small and unfavorable contribution of the polar interactions to the binding free energy. Generally, this relatively small or unfavorable contribution of polar effects occurs because the direct intermolecular electrostatic interactions, usually favorable, cannot always compensate for the large desolvation penalties associated with ligand binding.[57,58] It should be noted, however, that in some cases electrostatic interactions drive complex formation.[66−68]

4.2. The Binding Hotspots for PA. Based on the consistent information of pair interaction analysis and virtual alanine

(64) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9997–10002.
(65) Jayaram, B.; McConnell, K.; Dixit, S. B.; Das, A.; Beveridge, D. L. Free Energy Componenet Analysis of 40 Protein-DNA complexes: A Consensus View on the Thermodynamics of Binding at the Molecular Level. *J. Comput. Chem.* **2002**, *23*, 1–14.
(66) Lee, L. P.; Tidor, B. Barstar is electrostatically optimized for tight binding to barnase. *Nat. Struct. Biol.* **2001**, *8*, 73–6.
(67) Sheinerman, F. B.; Honig, B. On the Role of Electrostatic Interactions in the Design of Protein-Protein Interfaces. *J. Mol. Biol.* **2002**, *318*, 161–177.
(68) Ahmad, M.; Gu, W.; Helms, V. Mechanism of Fast Peptide Recognition by SH3 Domains. *Angew. Chem., Int. Ed.* **2008**, *47*, 7626–7630.

scanning, several residues of PA such as Trp706, Phe710, Asn412, Gln408, Phe411, Leu666, Trp619-Gln623, Gln670, Leu667, Met595, Thr639, Leu640 were found to have a large contribution to the binding process. The importance of these residues can be proved partially by binding assay of various PA mutants and PB1 terminal section. For example, *in vitro* binding assay for $PB1_N$ to PA_C double mutants W706A/Q670A, L666G/F710E, L666G/F710G and W706A/F710Q indicates these mutations largely disrupt the binding of $PB1_N$ to PA_C.[24] In addition, the point mutations in the C terminal domain of PA including L640D, L666D, W706A from ref 25 approve that these point mutations greatly weaken or abolish the binding of PB. Our energy decomposition and virtual alanine scanning results indicate all these residues have an obvious contribution to the binding process. Additionally, the residues from 619−623 are very important for PA binding, which are in good agreement with *in vitro* binding assay results that the deletion of residues 619−630 for PA weakens largely the binding ability of PA to PB1.[25] According to the structural analysis, the section from the residues 619−623 anchors the orientation of PB1 and has interactions with PB1 terminal. From the pair interaction energy and virtual alanine scanning analysis, Gln408 and Asn412 also should play an important role for PB1 binding. By monitoring the MD trajectory, it can be found that the Gln408 and Asn412 always form the hydrogen bonds with Asn4 and Val3 from PB1 respectively (Figure 6). The formation of these hydrogen bonds makes the residues Gln408 and Asn412 have an obvious electrostatic interaction contribution to the binding process (Figure 5a). In addition, although there is no direct experimental evidence for the importance of other involved residues, their contribution can be verified from the structural analysis. For example, Phe411 together with Trp706 constitutes one important hydrophobic pocket, which is responsible for the binding of Pro5 from PB1. Leu640, Met595 and Trp619 constitute another hydrophobic pocket and have a close contact with Leu8.

From the above discussion, it can be seen that the obtained information of our pair interaction analysis and virtual alanine scanning analysis is reliable. Such information is very important for the discovery of new anti-influenza inhibitors targeting PA−PB1 interaction. Whether the discovery of small molecular inhibitors or the structure-based design of peptidomimetics, we need to consider and target these important interactions.

4.3. The Binding Motif of PB1 Peptide. Combining the pair interaction analysis with virtual alanine scanning, seven residues contributing more than 2 kcal/mol for the PA−PB1 binding process can be identified as the basic "binding motif" of PB1. All essential residues Pro5, Leu7, Leu8, Phe9 and Leu10 for PB1 binding to PA found by aspartate scanning mutagenic studies[6,24,25] were included in our hotspot residues, indicating our calculation results are very reliable. All these five residues belong to hydrophobic residues, indicating further the importance of hydrophobic interaction. Another two residues, Val3 and Asn4, also contribute obviously to PB1 peptide binding, which form strong interaction with residue 621−623 by hydrophobic stacking and hydrogen bonds. In addition, the side chain of Asn4 also stacks well with Trp706 of PA. The deletion of 619−630 results in the greatly weakened binding of PB1 to PA, indicating the importance of Val3 and Asn4 indirectly. At the same time, the residues 2−4 of PB1 tend to form β sheet and form the regular hydrogen bonds between β sheets. These interactions are important to the orientation of PB1 during the recognition process between PA and PB1. The residues Val12 and Ala14 with 2 kcal/mol contribution in pair interaction energy analysis may be caused by the unnatural and more free state since the terminal of short peptide will have a different conformation with it in the whole protein.

The identified binding motif consisting of seven residues will be a good starting point to design the peptidomimetics inhibitors. Although we cannot guarantee this short peptide can form the secondary structure the same as that in complex, the formation of the important interactions with PA can provide it a good opportunity to be a successful inhibitor.

5. Conclusions

We performed an 8 ns molecular dynamics simulation of the subunit PA_C−$PB1_N$ in avian influenza H5N1 virus RNA polymerase. Based on the obtained trajectory, MM-PBSA and MM-GBSA were used to analyze the detailed interaction and binding free energy in the PA_C−$PB1_N$ complex. The results from free energy decomposition indicate that the intermolecular van der Waals interaction and the nonpolar solvation term provide the driving force for the binding of PA−PB1, while polar interactions (including the Coulomb contribution and polar solvation term) do not contribute significantly to PA−PB1 interaction. Through the pair interaction analysis and virtual alanine scanning, we identified the hot spots for PA and PB1 respectively. For PA binding to PB1, residues Trp706, Phe710, Phe411, Asn412, Gln408, Gln670, Leu666, Leu640 and residues 619−623 contribute obviously. For PB1 binding to PA, the basic binding motif identified includes the residues of Val3, Asn4, Pro5, Leu7, Leu8, Phe9, Leu10. The information can provide some insights for the structure-based design or discovery of anti-influenza virus drugs and will be useful for human beings to fight against all types of influenza including the extremely aggressive influenza A H5N1 and H1N1 since the PA−PB1 target is highly conserved. The identified binding motif can be used as the starting point for rational design of small molecule or peptidomimetics.

Acknowledgment. This work was supported by the Program for New Century Excellent Talents in University (No. NCET-07-0399).

MP900131P