Chem. Res. Chin. Univ., 2015, 31(6), 1029—1038 doi: 10.1007/s40242-015-5071-9

How Mutations Affecting the Ligand-receptor Interactions: a Combined MD and QM/MM Calculation on CYP2E1 and Its Two Mutants

WANG Yan, ZHENG Qingchuan*, ZHANG Jilong, XIE Mo, ZHAN Jiuyu and ZHANG Hongxing
Institute of Theoretical Chemistry, Jilin University, Changchun 130021, P. R. China

Abstract Cytochrome P450(CYP) 2E1 is a dual function monoxygenase with a crucial role in the metabolism of 6% of drugs on the market at present. The enzyme is of tremendous interest for its association with alcohol consumption, diabetes, obesity and fasting. Despite the abundant experimental mutagenesis data, the molecular origin and the structural motifs for the enzymatic activity deficiencies have not been rationalized at the atomic level. In this regard, we have investigated the effects of mutation on the structural and energetic characteristics upon single point mutations in CYP2E1, N219D and S366C. The molecular dynamics(MD) simulation combined with quantum mechanics/molecular mechanics(QM/MM) and noncovalent interaction(NCI) analysis was carried out on CYP2E1 and its two mutants. The results highlight the critical role of Phe207, which is responsible for both structural flexibility and energetic variation, shortening the gap between the theory and the experimentally observed results of enzymatic activity decrease. The underlying molecular mechanism of the enzymatic activity deficiencies for mutants may be attributed to the changes of spatial position of Phe207 in the two mutants. This work provides particular explanations to how mutations affect ligand-receptor interactions based on combined MD and QM/MM calculations. Furthermore, the mutational effects on the activity of CYP2E1 obtained in the present study are beneficial to both the experimental and the computational works of CYPs and may allow researchers to achieve desirable changes in enzymatic activity.

Keywords Cytochrome P450(CYP) 2E1; Molecular dynamics(MD) simulation; Quantum mechanics/molecular mechanics(QM/MM, ONIOM) calculation; Noncovalent interaction(NCI) analysis

1 Introduction

Cytochrome P450s(CYPs) are widespread heme-containing mixed function oxygenases that catalyze the hydroxylation of nonactived hydrocarbon reactions, dealkylation epoxidation and dehydrogenation reactions involved in the oxidative metabolism[1,2]. They play an indispensable role in the metabolism of endogenous and exogenous substrates like drugs and environmental chemicals[3]. These enzymes are primarily responsible for the phase I drug metabolism. During this phase, the drug molecules are made more water-soluble for clearance. Members of this family, including CYP3A4, CYP2D6, CYP2C9, CYP2E1 and CYP1A2 are medically significant and account for the metabolism of approximately 90% of drugs currently on the market[4—6].

In human, CYP2E1 is a member of various isoforms of CYPs. It only accounts for about 2% of total CYPs in the liver cells, while it participates the metabolism of 6% of drugs on the market at present[7]. CYP2E1 is of extensive interest to the pharmaceutical industry for its involvement in the metabolic processes of many low molecular weight compounds and in liver toxicity[8—10]. The enzyme activity deficiency, which is associated with alcohol consumption[10], may lead to diabetes[11], obesity[12] and fasting[13]. Therefore, there is tremendous interest in CYP2E1 as it is closely related to human health. The size of the active site in CYP2E1 is the smallest(18.2 nm^3) among human CYP crystal structures available[14—16]. This feature might give an explanation to CYP2E1's preference to metabolize small molecular weight compounds.

N219D and S366C[17] are two of the six CYP2E1 natural mutants. Among these six mutants, R76H[18], V179I[19], V389I[18] and H457L[20] have all been studied. How N219D and S366C affect the activity of CYP2E1 has not been investigated. The variant N219D has one single altered amino acid mutation on the F helix. This mutation which changes the residue from Asn(medium size and polar) to Asp(medium size and acidic) is located *ca.* 1.7 nm away from the heme iron. The S366C mutation which changes residue from Ser(small size and polar) to Cys(medium size and polar) is on the loop region near K helix. The mutation site is *ca.* 1.2 nm from the heme iron. These two mutations may affect the ligand-receptor interactions between indazole molecule and its surrounding residues on the active site. Previous studies[15,21,22] have pointed out that ligand-receptor interactions such as π-π interaction on the active site of

*Corresponding author. E-mail: zhengqc@jlu.edu.cn
Received March 2, 2015; accepted March 27, 2015.
Supported by the National Natural Science Foundation of China(No.21273095).

wild type CYP2E1 are crucial for the regulation of enzyme's activity. Such interactions might vary in mutational systems. The changes of ligand-receptor interactions on the active site may finally influence the activity of CYP2E1 enzyme.

In the present study, MD simulations were carried out for wild type and two mutants(N219D and S366C) of CYP2E1. In order to better describe ligand-receptor interactions on CYP2E1's active site and visualize such interactions, ONIOM calculations and NCI analysis were performed on these three systems. Our work provides detailed atomistic insights into the structure-function relationships of CYP2E1 and its two mutants under dynamics conditions. Furthermore, the mutational effects on the activity of CYP2E1 obtained in the present study are beneficial to both experimental and computational work of CYPs and may allow researchers to achieve desirable changes in enzymatic activity.

2 Computational Details

2.1 Preparation of Structures

The crystal structure of CYP2E1 bound with 1H-indazole(LZ1) was retraced from the RCSB Brookhaven Protein Data Bank(PDB ID: 3E6I)[23], which served as the initial structure of the wild type. The starting structures of the two mutants were modeled *via* Discovery Studio 2.5 software package[24] based on the X-ray structure of wt2E1. The missed residue, Met138 in the crystal structure was added with the help of Discovery Studio 2.5. The protonation states of the ionizable residues were determined by PROPAK[25] according to the previous research[22]. Hydrogen atoms were subsequently added to the complex structures with *t*-Leap module of AMBER11[26]. The initial model of 1H-indazole was extracted from the crystal structure(PDB ID: 3E6I). The force filed parameters for 1H-indazole complex with heme were supplied by AMBER11 with electrostatic potential fitting procedure and Gaussian 09 software package[27]. Structural optimization of 1H-indazole complex with heme was conducted *via* B3LYP combined with an LACVP basis set on Fe element and 6-31+G* basis set on the others based on the Gaussian 09. RESP fitting procedure was employed for charge derivation based on the optimal conformation. The force filed parameters developed by Rydberg *et al*.[28] were assigned for heme. The overall structures of wt CYP2E1 and LZ1 molecule were manifested in Fig.1.

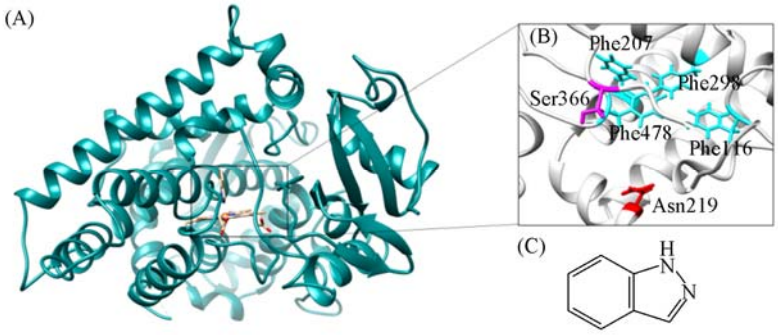

Fig.1 Overall structure of wt2E1(A), a close view for the four important Phe residues and the two mutational residues(B) and chemical structure of 1H-indazole(LZ1) molecule(C)

(A) wt2E1 structure presented in a cyan colored ribbon sytle. 1H-Indazole and heme molecule are in stick model.

2.2 Molecular Dynamics(MD) Simulations

MD simulations for wt2E1 and two mutants including energy minimization were carried out *via* AMBER11[26] software package and ff99SB force field[29]. To keep the whole system neutral, chloride ions(Cl^-) were added with the aid of *t*-Leap procedure of AMBER11 based on a Coulomb potential grid. Each system was then solvated with the TIP3P water model[30] in a truncated octahedron box with a distance of 10.0 nm around the solute. The whole systems were submitted to 2000 steps of steepest decent minimization followed by 3000 steps of conjugate gradient minimization. Subsequently, the systems were heated from 0 K to 300 K for 1000 ps in the canonical ensemble(NVT ensemble), applying harmonic restraints with a force constant of 4184.0 $kJ\cdot mol^{-1}\cdot nm^{-2}$ on the protein and small molecules. A Langvin thermostat was adopted. Subsequently, the two systems were equilibrated in an NPT ensemble under a constant pressure(1.0×10^5 Pa) for 2000 ps. The relaxation time for barostat bath was set to 2.0 ps. Finally, a total of 100 ns was simulated for each system under NPT ensemble condition and periodic boundary conditions and particle-mesh Ewald[31](PME) for long range electrostatics. Short range interactions were cut off at 1.0 nm, and bonds involving hydrogen were held fixed with the help of SHAKE. The time step was set to 2 fs.

2.3 MD Trajectories Analysis

Trajectories generated from MD simulations were analyzed *via* several orders of AmberTools1.5. The root-mean-square deviation(RMSD) value is a useful estimation for quantifying conformational changes of the same protein. In this study, C_α-RMSD was calculated for all the systems. To evaluate local structure fluctuation of helixes and β-sheets, the hydrogen bond population was calculated. Hydrogen bonds were determined *via* the distance between the heavy atoms in a cutoff of 0.3 nm and the angle between the acceptor and donor atoms in a cutoff of 120°. 100 ns MD simulations were performed for wt2E1 and two mutants, which can generate 50000 snapshots. Cluster analysis is the task of grouping a set of snapshots in such a way that snapshots in the same group(named a cluster) are more similar to each other than to those in other groups(clusters). The cluster analysis of protein conformations

was carried out with average linkage as the clustering algorithm, and backbone atom RMSD as the distance metric. This average-gel inkage algorithm is recommended in all cluster algorithms[32]. Clustering analysis can be summarized as follows. In the average linkage algorithm, the distance from one cluster to another is defined as the average to all distances between individual points of the two clusters. At each iteration step, the two closest clusters are merged. This merging continues until the desired number of cluster is obtained(here is 5). VMD[33], Chimera[34] and PyMOL[35] softwares were employed to visualize the trajectories and to depict structural representations.

2.4 Principal Component Analysis

Principal component analysis(PCA) is used to space a protein's conformational space into essential subspace containing several degrees of freedom that describe the motions of the proteins[36]. PCA uses a covariance matrix constructed with structures sampled from MD simulations. By diagonalization of a covariance matrix of coordinates of the systems, the motions of a structure along the trajectory are determinated. To obtain the dominant motion over a MD simulation, trajectory would be projected along the direction described by a selected eigenvector. Through calculating two largest projections on the average structure, the primary moving direction of the protein can be obtained. PCA was performed *via* ProDy[37] software. Three-dimensional structural snapshots were visualized with VMD and its plugin NWWiz[33,37].

2.5 MM-GB/SA Calculations

The MM-GB/SA methods[38,39] implemented in AMBER 11[26] were applied to calculating the binding free energy between the ligand and the enzyme[40]. The binding free energy(ΔG_{bind}) in MM-GB/SA between a ligand(L) and a receptor(R) to form a complex RL was calculated as

$$\Delta G_{bind}=G_{complex}-(G_{receptor}+G_{ligand}) \quad (1)$$

$$G=E_{MM}+G_{sol}-TS \quad (2)$$

$$E_{MM}=E_{int}+E_{ele}+E_{vdW} \quad (3)$$

$$G_{sol}=G_{GB}+G_{SA} \quad (4)$$

In Eq.(2), E_{MM}, G_{sol}, and TS represent molecular mechanics component in gas phase, the stabilization energy due to salvation, and a vibrational entropy term, respectively. E_{MM} is given as a sum of E_{int}, E_{ele} and E_{vdW} which are internal, Coulomb and van der Waals interaction terms, respectively. Solvation energy, G_{sol}, is separated into an electrostatic solvation free energy(G_{GB}) and a nonpolar solvation free energy(G_{SA}). The former can be obtained by the Generalized Born(GB) method. The latter is considered to be proportional to the molecular solvent accessible surface area(SASA)[41]. The binding free energies were obtained by averaging over the values calculated for 3000 snapshots from the last 30 ns of the trajectories at 5 ps intervals for the complex structure.

2.6 QM/MM Calculations

QM/MM calculations were performed *via* a two-layer ONIOM scheme in the Gaussian 09 software package[27]. The ONIOM method is a hybrid quantum chemical approach developed by Morokuma and his co-workers[42—44], which allows different levels of theory to be applied to different parts of a molecular system. In a two-layer ONIOM calculation[42,45,46], the molecular system is divided into two parts. The 'model' system is composed of the most crucial elements of the system and is handled with a high-level, accurate and expensive computational method, which can describe chemical bond breaking and formation. The 'real' system contains the complete system and a low-level, inexpensive computational method, often molecular mechanics, is applied to this system. By using this method, the environmental effects of the molecular surroundings on the model system are well described. In this study, the high-level method was the functional B3LYP[47,48], and the low-level method adopted the Amber99 force field[49]. The feasibility of the ONIOM(B3LYP:Amber) method has previously been demonstrated for several enzymatic systems[50—56]. The QM region is composed of four Phe residues(Phe116, Phe207, Phe298 and Phe478), the inhibitor LZ1 and the heme. Link hydrogen atoms were employed to saturate the dangling covalent bonds. The QM region containing 144 atoms for each of the three systems was described by the functional B3LYP. For geometries optimizations, we used LanL2DZ basis set[57] on iron and 6-31G basis set on the remaining atoms. The rest of the system(MM region) was treated by AMBER parm99 force field. The total numbers of atoms for wt2E1 and the two mutants are 7660, 7658 and 7660, respectively. During the ONIOM calculations, the electrostatic embedding method was employed. Based on the optimized structures for these three systems, a model approximation method was employed here. Since we only focused on the core region for these three systems, in the energy calculation only co-factor heme, LZ1 molecule and four Phe residues were taken into consideration. To better describe weak intermolecular interactions, the D3 correction proposed by Grimme and co-workers[58—60] was applied to B3LYP energies *in vacuo*.

2.7 Visualization of Noncovalent Interactions

To identify the interactions between the 1H-indazole and the surrounding protein residues(especially Phe residues), the NCI(non-covalent interactions) index[61] was employed, which provides a visual way to detect weak interactions in chemical systems. This method is on the basis of the analysis of the electron density, ρ, and its reduced gradient, s, where

$$s=\frac{1}{2(3\pi^2)^{1/3}}\frac{|\nabla\rho|}{\rho^{4/3}} \quad (5)$$

and permits highlighting interactions characterized by a low density regime. The analysis of the second derivatives of the density along the main axis of variation allows distinguishing between different interaction types. In particular, nonbonded interactions, such as steric repulsion are identified by the positive sign($\lambda_2>0$) of the second eigenvalue, λ_2, of the electron density Hessian matrix; and bonding interactions, such as hydrogen bonds are characterized by a negative sign($\lambda_2<0$). And van der Waals interactions present a negligible density overlap, which gives $\lambda_2\cong 0$. The data of electron densities were

obtained from afore mentioned ONIOM calculations. Multiwfn software package[62] was employed for the NCI analysis. The results were depicted by VMD[33]. This method has been successfully applied in many cases[63,64].

3 Results and Discussions

3.1 Overall Structural Features for Cyp2E1 and Its Two Mutants

Root-mean-square deviations(RMSDs) provide insights into the conformational changes of the structure. RMSDs were calculated with respect to the initial frame which was generated by MD simulations for the wild type(wt), N219D and S366C systems. The RMSD plot is demonstrated in Fig.2(A). From Fig.2(A), the RMSD values of the protein backbone for wt and S366C systems raised up slowly ranging from 0.1 nm to 0.2 nm and at last attained equilibrium at about 70 ns. Whereas, for N219D system, the RMSD value of the protein backbone increased suddenly at 40 ns to about 0.35 nm and then fluctuated around 0.3 nm of the rest of the simulation time. This may indicate that the structure of the N219D system undergoes certain conformational changes. The RMSDs for Phe116, Phe207, Phe298 and Phe478 were also monitored over the simulation time with the results manifested in Fig.2(B). The RMSD values for Phe residues of N219D system were higher than those of S366C system, which implies the four Phe residues of N219D system may undergo certain conformational changes.

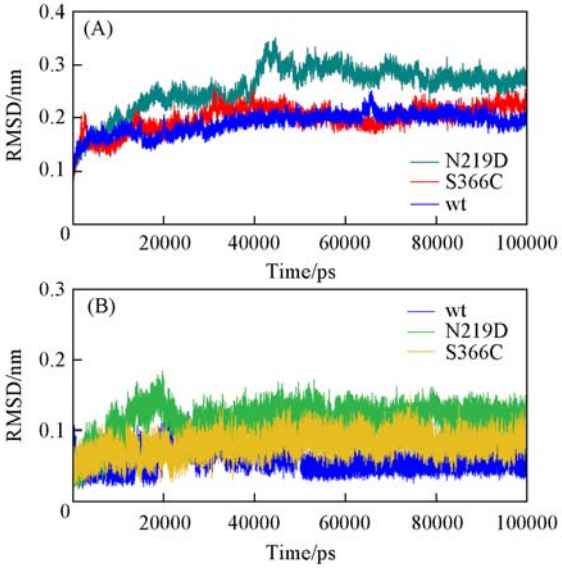

Fig.2 Calculated root-mean-square deviations(RMSDs) of the backbone atoms referenced to the corresponding starting structure(A) and RMSD of four Phe residues(Phe116, Phe207, Phe298 and Phe478)(B)

Root-mean-square fluctuation(RMSF) offers details on fluctuations of each residue over simulation time. High values indicate large flexibility of certain residues. The RMSF plot is manifested in Fig.3 and the residues with higher RMSF values are labelled. In order to easily find out the region where the residues with higher RMSF value locate, the result of secondary structure analysis is attached on the top panel of Fig.3 based on the DSSP[65] calculation. The overall flexibility of wild type system changed slightly during the whole MD simulation. Compared to those of N219D and S366C systems, the most flexible residue is Met138. This residue locates at the C-D loop of CYP2E1, the terminal side of C-helix. In the two mutants, some hydrogen bonds may form in the region where Met138 located and make Met138 less flexible. For N219D system, the most flexible residue compared to those of the other systems is Asp219. Asp219 locates near the F-G loop region. The mutation at 219 position changes the residue from neutral to negatively charged residue. Such mutation may cause the formation of salt bridge between this residue and the surrounding residues. This would also make Asp219 interact actively with its adjacent residues and thus drag the residue away from its origin position. In S366C system, Pro120 and His278 possessed higher RMSF values compared to those of wt and N219D systems. These residues all come from loop region and are far away from active pocket.

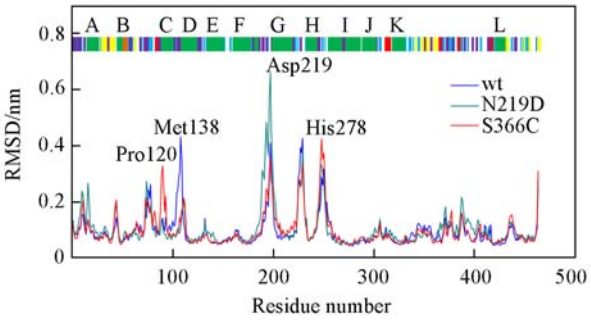

Fig.3 Root-mean-square fluctuations(RMSF) of wt, N219D and S366C systems

wt system is represented by blue lines, N219D system cyan lines and S366C system red lines. The results from DSSP are shown in the top panel of this figure. Helices are indicated in green(and labeled at the top of figure), β-sheets in yellow and turns in purple. A, B, etc. stand for the well-known helix of the CYPs family.

3.2 Identification of Crucial Residues and Hydrogen Bond Analysis

To characterize the interactions between active site residues and substrates, the close contacts were calculated along the MD simulation trajectories for all the simulated systems. Any residues around the LZ1 molecule, which are within a range of 0.45 nm are considered to have close contact with LZ1. The results of close contact analysis are demonstrated in Table 1 and Fig.S1(see the Electronic Supplementary Material of this paper). Three Phe residues(Phe207 in F helix, Phe298 in I helix and Phe478 in β4-1/β4-2 turn) have significant close contact with the LZ1 molecule. There might be π-π interaction between these three Phe residues and LZ1 molecule.

Table 1 Results of close contact for wt, N219D and S366C systems

System	Percentage of close contact(%)			
	Phe116	Phe207	Phe298	Phe478
wt	13.89	64.49	31.52	97.48
N219D	NA	77.93	1.57	93.90
S366C	0.01	86.85	81.18	57.77

Based on the results of close contact analysis, hydrogen bonds for LZ1, crucial Phe residues and mutational residues were also monitored over the MD simulation. The results of hydrogen bonds analysis are demonstrated in Fig.4 and Table S1(see the Electronic Supplementary Material of this paper). In wt system, two hydrogen bonds were detected. One formed between Asn367 and LZ1, the other occurred between Phe207 and Ser211. The occupancies of these two hydrogen bonds are 79.44% and 94.06%, respectively. The hydrogen bond between Asn367 and LZ1 kept the LZ1 molecule in the active pocket. More hydrogen bonds were identified for N219D and S366C systems. In N219D system, five hydrogen bonds formed between D219(the mutation residue) and its adjacent residues. For S366C system, after Ser366 mutated to Cys366, a couple of hydrogen bonds occurred in the Cys366 located region. The hydrogen bond between Phe207 and Ser211 is arrestive. For this hydrogen bond, there is a decrease in the percentage of occupancy for N219D(70.34%) and S366C(74.28%) compared to that of wt system(94.06%). Such a reduced percentage may influence the spatial position of Phe207. Aforementioned RMSF calculations revealed Met138 possessed a higher RMSF value in wt system. The results of hydrogen bonds analysis for certain residues can give an explanation. In N219D system, a hydrogen bond of 47.36% was formed between Gly139 and Asn134, while this hydrogen bond was not detected in wt system. In S366C system, the percentage of hydrogen bond between Ser129 and Ile125(98.46%) was increased compared to that of wt system(85.24%). Ser129 located near Met138. These two hydrogen bonds in mutational systems help to stabilize Met138. The changes of hydrogen bond network for the two mutants may bring some diverse properties to the structures of mutational systems.

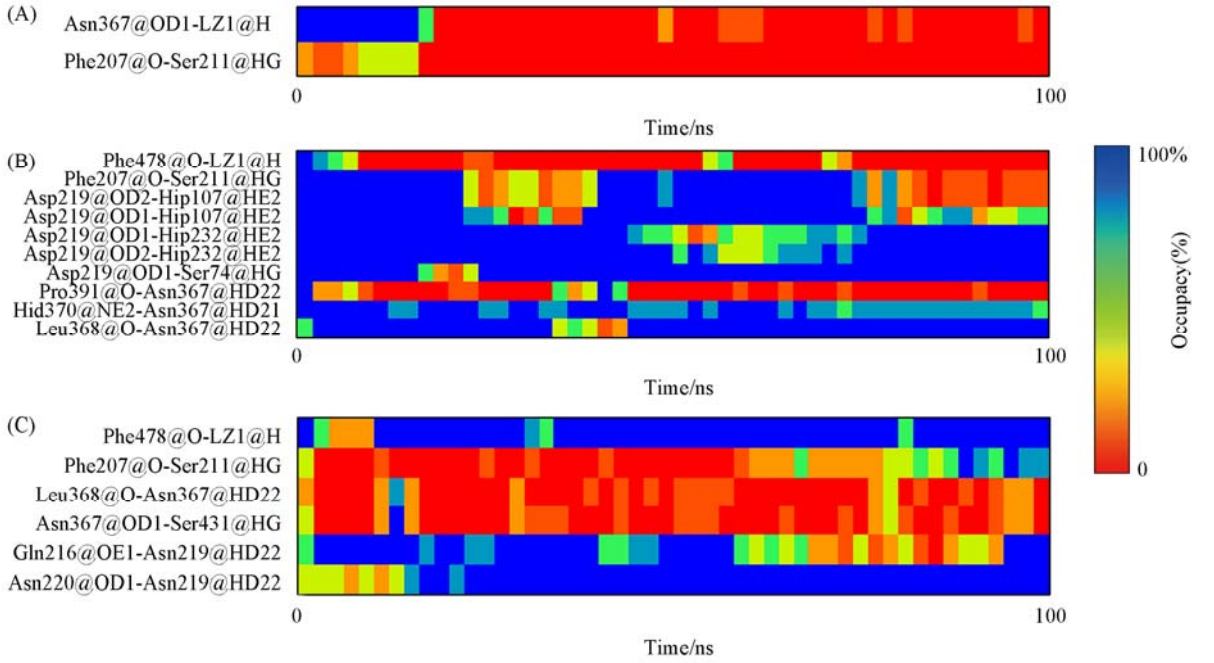

Fig.4 Percentages of occupancies of hydrogen bonds in wt(A), N219D(B) and S366C(C) systems

All these hydrogen bonds are monitored during the whole simulation time of each system.

3.3 Clustering Analysis for wt and Mutational Systems

In order to get a deeper insight into the structural details of wt and mutational systems, clustering analysis of the structures was carried out. Each trajectory was divided into five clusters by the average linkage algorithm. In these five clusters, five snapshots were chosen to be the representative structures of clusters and named them as follows, C0, C1, C2, C3 and C4. C0 represents the initial structure.

For wt system, the most popular cluster is C0 accounting for 49.3% of the total population. For N219D system, the most popular cluster is C3 accounting for 41.1% of total population. The most popular cluster for S366C system is C2, which possessed 34.6% of total population. In order to verify the structural variance between each representative structure and crystal structure, C0 of wt system, C3 of N219D system and C2 of S366C system were respectively superimposed on the crystal structure. The results are demonstrated in Table 2 and Fig.5. As can be seen in Table 2, the RMSD values for the three systems are resemblance which implies the overall structural variances among the representative structures are unconspicuous. However, there are notable structural variations for the B-C and F-G loop regions of N219D and S366C as manifested in Fig.5. Previous close contact analysis revealed a couple of Phe residues (Phe207, Phe298 and Phe478) possessed close contact with the LZ1 molecule. Phe207 is in the F-G loop region. The behavior of these three Phe residues are also worthy to note. As demonstrated in Fig.S2(A)(see the Electronic Supplementary Material

Table 2 Populations of representative structures and their RMSD values having been superimposed on crystal structure

Structure	Population(%)	RMSD of crystal structure/nm
C0 of wt	49.3	0.11
C3 of N219D	41.1	0.13
C2 of S366C	34.6	0.12

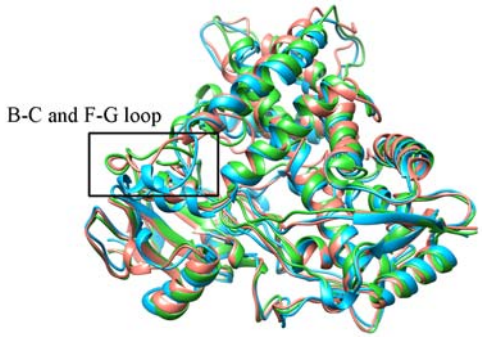

Fig.5 Alignment of three representative structures for wt, N219D and S366C systems

All the structures are in ribbon style. wt is colored in salmon, N219D in deep sky blue and S366C in lime green.

of this paper), the centroid distance of Phe207 and Phe298 is close in wt system, while in N219D and S366C systems, these two Phe residues are away from each other. It should be noted that the LZ1 molecule is far away from Phe207 in wt system as shown in Fig.S2(B)(see the Electronic Supplementary Material of this paper), whereas the distances of these two groups in mutational systems are quite close. Such orientation of Phe207 and LZ1 for mutational systems was similar to that of the π-π interaction pointed out by McGaughey *et al.*[66]. This method found stacked and staggered π-π interactions *via* performing the following tests. (1) The distance between the centroids of each pair of aromatic rings was determined to find those which fall within the center distance cutoff(default 0.8 nm). (2) For those distances, an atom from each ring should be within the closest atom distance cutoff(0.45 nm). (3) The angle θ between the normal of one ring and the centroid-centroid vector must fall between 0° and ± the angle θ cutoff(default 60°), and the angle γ between the normal and each ring must fall between 0° and ± the angle γ cutoff(default 30°). The criterion for π-π interaction was once employed in our previous work[67]. Based on the criterion for π-π interaction, this interaction between Phe207 and LZ1 was monitored. The results show that no π-π interaction was detected between Phe207 and LZ1 in wt system, while 14.5% and 29.5% π-π interactions formed for N219D and S366C systems, respectively.

3.4 Relative Motions for wt, N219D and S366C Structures

To quantitatively identify the relative motions of the structures, PCA was carried out for each system, and the structural variations along the first principal component are demonstrated(Fig.6). It is easy to distinguish the remarkable motions in the three systems projected on different regions of the structures. wt system possessed less relative motions among these three structures. The collective motions for N219D and S366C systems are similar. A conspicuous collective motion occurred in the F-G loop region for both the mutational systems. The trendy of this collective motion is semblable.

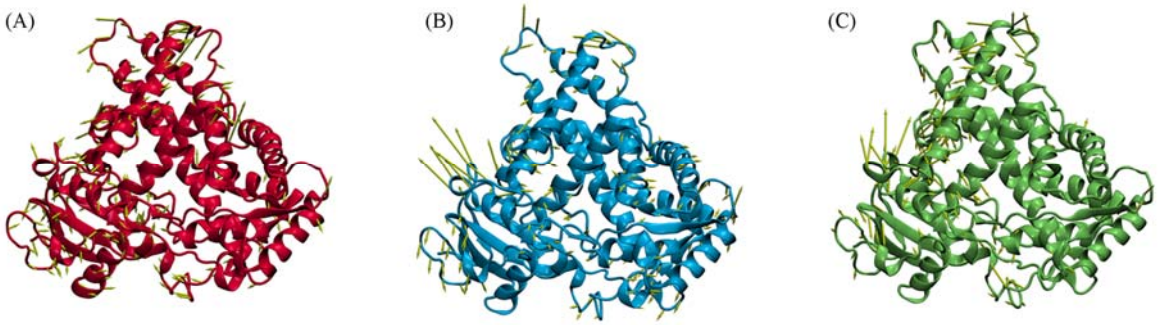

Fig.6 Porcupine plots in stereo showing wt(A), N219D(B) and S366C(C) systems with cones signifying the first eigenvectors movements

wt system is colored in red, N219D in blue and S366C in lime.

As manifested in Fig.6, the amplitude of collective motions for the F-G loop of N219D system is rather large. In order to characterize the motion of this loop region in N219D system, the centroid distance between co-factor heme and F-G loop region was monitored with the results demonstrated in Fig.7. There is a remarkable distance shift for F-G loop region in N219D system. In N219D system, position 219 is changed to a negative residue. There are some positive residues around the 219 position. This indicates that in N219D system there may be some electrostatic interactions between Asp219 and its adjacent residues. To this end, salt bridge analysis was employed with the results shown in Table S2(see the Electronic Supplementary Material of this paper). As shown in Table S2, in N219D system, Asp219 indeed formed a couple of salt bridges with its adjacent residues. His232 should be paid special attention. In wt and S366C systems, His232 formed hydrogen bond with Ala228 as shown in Table S1. However, in N219D system, this hydrogen bond disappeared. The formation of salt bridge between His232 and Asp219 in N219D system broke this hydrogen bond. The spatial positions for Ala228, His232 and Asp219 are worthy to note. Ala228 and His232 belong to the

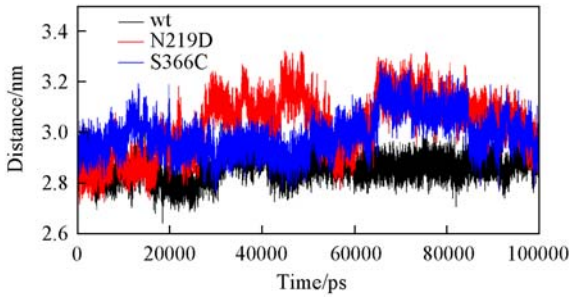

Fig.7 Centroid distance between the heme co-factor and the F-G loop monitored over the whole simulation time

F-G loop region while Asp219 is near F-G loop region. After the mutation occurred at the 219 position, the electrostatic interaction between Asp219 and His232 will drag His232 away from its own position. Such a position shift may disturb the overall structural stability of the F-G loop region, which will in turn make the F-G loop region more flexible.

3.5 Revealing the Ligand-receptor Interactions Between Crucial Phe Residue, LZ1 Molecule and Heme co-Factor

To explore the inhibitation ability of LZ1, MM-GB/SA calculations and energy decomposition were performed. A total of 3000 snapshots were extracted from the last 30 ns trajectory for binding free energy analysis. The results for wt, N219D and S366C systems are list in Table 3. The K_i value of inhibitor LZ1 to CYP2E1 was determined by Collom *et al.*[68]. Here, this value was employed to convert to its corresponding ΔG value (ΔG_{exp}).

As shown in Table 3, calculated binding free energy for wt system is very close to the ΔG_{exp} value, while the values for N219D and S366C are higher than the ΔG_{exp} value. This implies that the activity of mutational CYP2E1 is decreased. The results of binding free energy decomposition calculations are shown in Fig.8. It is worthy to note that there is a remarkable energy variance at Phe207 position. The decomposition binding energy of Phe207 for wt system is much lower than those of the two mutational systems. Asn367 has a higher decomposition free energy than mutational systems. Previous hydrogen analysis results can explain such energetic variances. There is a high percentage hydrogen bond between Asn367 and LZ1 in wt system, whereas this hydrogen bond is weaken in N219D and S366C systems.

Table 3 Binding free energies and its components for wild type and two mutants

System	ΔE_{ele}/(kJ·mol^{-1})	ΔE_{vdW}/(kJ·mol^{-1})	ΔG_{GB}/(kJ·mol^{-1})	ΔG_{SASA}/(kJ·mol^{-1})	$\Delta G_{MM\text{-}GB/SA}{}^{a}$/(kJ·mol^{-1})	$-T\Delta S$/(kJ·mol^{-1})	$\Delta G_{tot}{}^{b}$/(kJ·mol^{-1})
wt	–35.57±0.84	–89.54±0.63	48.24±1.24	–10.71±0.092	–88.95±2.42	49.04±0.59	–39.92±1.90
N219D	–37.82±0.42	–86.52±1.75	40.79±0.32	–10.12±0.18	–93.76±1.52	74.14±4.58	–19.04±3.07
S366C	–23.14±0.12	–79.41±0.028	40.50±1.42	–10.83±0.17	–72.89±1.40	55.22±1.17	–17.65±0.23

a. $\Delta G_{MM\text{-}GB/SA} = \Delta E_{ele} + \Delta E_{vdW} + \Delta G_{GB} + \Delta G_{SASA}$; *b*. $\Delta G_{tot} = \Delta G_{MM\text{-}GB/SA} - T\Delta S$; $\Delta G_{exp} = RT\ln K_i = -41.04$ kJ/mol; K_i=0.12 μmol/L.

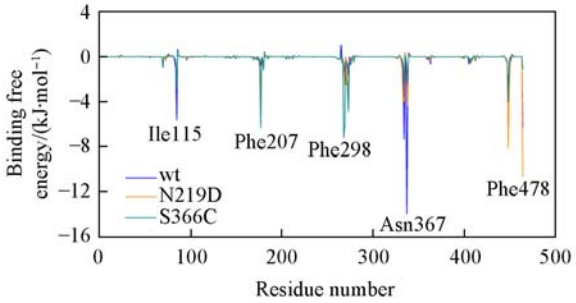

Fig.8 Decomposition binding free energy for wt, N219D and S366C systems

The residues with high decomposition binding free energies are labelled.

Previous π-π interaction analysis revealed that such interaction existed between Phe207 and the LZ1 molecule for N219D and S366C systems. Other Phe residues(Phe298 and Phe478) which possessed high occupancy close contact with LZ1 might also form π-π interaction with LZ1. In order to better describe such interactions, QM/MM calculation and B3LYP hybrid functional with Grimme's DFT-D3 dispersion correction were employed. For the representative structures obtained from the 100 ns MD simulation of each system, QM/MM(ONIOM) optimization was employed for these three structures. After optimization, the fully optimized structures were obtained. Since the main concern is to characterize the ligand-receptor interactions within the core region of each system, the simplified modes(the co-factor heme, the LZ1 molecule and the four Phe residues) were employed in the single point energy calculation. To better describe weak intermolecular interaction, the D3 correction was applied to B3LYP energies *in vacuo*. On the basis of the data of electron densities obtained from DFT-D3 calculation, the nonbonded interactions were visualization. The NCI analysis method[61] and Multiwfn software package[62] were employed for NCI analysis. The ONIOM optimized structures are shown in Fig.S3(see the Electronic Supplementary Material of this paper). The results of energy calculations and NCI analysis are shown in Table 4 and Fig.9.

As demonstrated in Fig.S3, the overall structural features of ONIOM optimized structures are similar to those of each representative structure obtained by MD simulations. The LZ1

Table 4 BSSE energy summary of key residues in wt, N219D and S366C systems

System	Pair	Interaction energy/(kJ·mol^{-1})	System	Pair	Interaction energy/(kJ·mol^{-1})
wt	LZ1 and Heme	–39.79	N219D	LZ1 and Phe116	–0.04
	LZ1 and Phe207	–2.34		Phe207 and Phe478	–6.74
	LZ1 and Phe478	–10.63		Phe298 and Phe478	–6.61
	LZ1 and Phe298	–2.68		Phe207 and Phe298	–3.81
	LZ1 and Phe116	–5.19	S366C	LZ1 and Heme	–32.55
	Phe116 and Phe478	–4.23		LZ1 and Phe207	–19.75
	Phe207 and Phe478	–5.86		LZ1 and Phe478	–0.54
	Phe298 and Phe478	–1.88		LZ1 and Phe298	–9.71
	Phe207 and Phe298	–7.66		Phe116 and Phe478	–4.06
N219D	LZ1 and Heme	–30.08		Phe207 and Phe478	–0.96
	LZ1 and Phe207	–14.90		Phe298 and Phe478	–0.13
	LZ1 and Phe478	–5.90		Phe207 and Phe298	–3.56
	LZ1 and Phe298	–0.21			

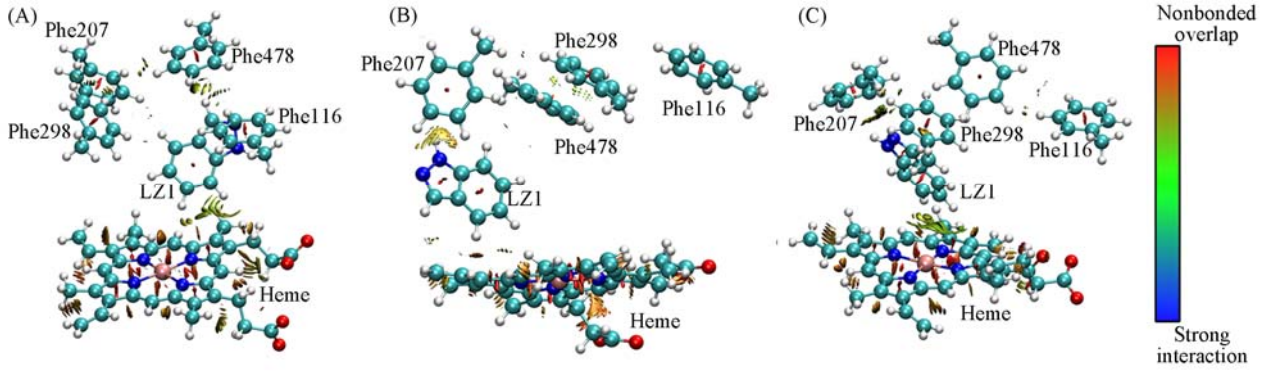

Fig.9 Visualization of nonbonding interaction based on the noncovalent interaction(NCI) analysis

(A) wt system; (B) N219D system; (C) S366C system. The surfaces are colored on a blue-green-red scale according to the values of sign($\lambda 2$)ρ, ranging from –0.04 a.u. to 0.02 a.u.

molecule in N219D and S366C systems is a little away from the center of the heme co-factor. As shown in Table 4, in N219D and S366C systems, the BSSE interaction energies for LZ1 and heme pair are lower than that for wt system(–39.79 kJ/mol for wt, –30.09 kJ/mol for N219D and –32.55 kJ/mol for S366C system). This result implies that the LZ1 molecule interacts with heme co-factor weakly in the mutational systems. Such weaken interactions between these two groups may affect the activity of CYP2E1. The BSSE interaction energy of Phe207 is also worthy to note. In wt system, the interaction energy between LZ1 and Phe207 is lower than those between them in the other two systems(–2.34 kJ/mol for wt, –14.90 kJ/mol for N219D and –19.75 kJ/mol for S366C). The interaction energy for LZ1 and Phe207 is enhanced in mutational systems. This result is in accordance to the previous π-π interaction analysis based on the MD simulation. On the basis of MD simulation and BSSE energy calculation, π-π interaction exists between Phe207 and the LZ1 molecule. As shown in Fig.S3, Phe298 inserts between Phe207 and the LZ1 molecule in wt system. Phe298 possesses a higher interaction energy with Phe207 in wt system than in N219D and S366C systems. Phe207 in the mutational system dragged the LZ1 molecule away from the heme co-factor. The high interaction energy between Phe207 and LZ1 molecule is the main reason for low interaction energy between LZ1 molecule and heme in the two mutational systems. Fig.9 exhibits an intuitionistic presentation of aforementioned nonbonded interactions. In N219D and S366C systems, an obvious golden surface existed between Phe207 and the LZ1 molecule. Such a kind of surface was not found in wt system. In wt system, Phe298 acted as a partition between Phe207 and the LZ1 molecule. In such an isolated mode, the LZ1 molecule can stay close to the center of heme(close to the iron atom). Such an isolated mode may facilitate the catalytic reactions to take place.

The tendency of the calculated binding free energies for wt, N219D and S366C systems is similar to that of the energies obtained by DFT-D3 calculations. Such an energetic accordance confirms the activity decrease for the mutational systems. The results of the decomposition free energy and BSSE interaction energy calculation highlight the role which Phe207 plays in the decrease of CYP2E1's activity for the two mutational systems.

4 Conclusions

The application of molecular dynamics(MD) simulation has become indispensable in computational domain of particular emphasis with regard to receptor-ligand interactions[67,69—76]. In order to figure out how N219D and S366C mutants affect the activity of CYP2E1, MD simulation combined with QM/MM(ONIOM) calculations were carried out. On the basis of the energy calculations, the activity of the two mutants is indeed decreased. The reduced activity for the two mutational systems was achieved by dragging the LZ1 molecule away from the heme iron center. Compared to wt system, N219D and S366C systems possessed π-π interaction between Phe207 and LZ1. The changes of the spatial position of Phe207 lead to the formation of such an interaction. There are distinguishing reasons for the change of the spatial position of Phe207 in N219D and S366C systems. The salt bridge interaction around D219 in N219D system changed the original conformation of the F-G loop region. The variation of the hydrogen bond occupancy of Phe207 and Ser211 in S366C system also affect the original conformation of this loop region. Several remarkable collective motions were observed for this region. Phe207 located near the F-G loop region. The spatial position of Phe207 changed along with the conformational change of the F-G loop region. Based on the energetic calculation and NCI analysis, Phe207 plays a crucial role in the decreased activity of CYP2E1 mutants. Our work provides detailed atomistic insights into the structure-function relationships of CYP2E1 and its two mutants under dynamics conditions. This work also provides particular explanations on how mutations affect weak interactions based on combined MD-QM/MM calculations. Furthermore, the mutational effects on the activity of CYP2E1 obtained in the present study are beneficial to both experimental and computational work of CYPs and may allow researchers to achieve desirable changes in enzymatic activities.

Electronic Supplementary Material

Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s40242-015-5071-9.

References

[1] Danielson P., *Current Drug Metabolism*, **2002**, *3*, 561

[2] Nelson D. R., *Archives of Biochemistry and Biophysics*, **1999**, *369*, 1

[3] Watanabe K. P., Kawai Y. K., Ikenaka Y., Kawata M., Ikushiro S. I., Sakaki T., Ishizuka M., *PloS One*, **2013**, *8*, e75689

[4] Scott E. E., Halpert J. R., *Trends in Biochemical Sciences*, **2005**, *30*, 5

[5] Park H., Lee S., Suh J., *J. Am. Chem. Soc*., **2005**, *127*, 13634

[6] Walsh A. A., Szklarz G. D., Scott E. E., *Journal of Biological Chemistry*, **2013**, *288*, 12932

[7] Miller G. P., *Expert Opin. Drug Met*., **2008**, *4*, 1053

[8] Trafalis D. T., Panteli E. S., Grivas A., Tsigris C., Karamanakos P. N., *Expert Opinion on Drug Metabolism & Toxicology*, **2010**, *6*, 307

[9] Ioannides C., Lewis V., David F., *Current Topics in Medicinal Chemistry*, **2004**, *4*, 1767

[10] Nanji A. A., Zhao S., Sadrzadeh S., Dannenberg A. J., Tahan S. R., Waxman D. J., *Alcoholism: Clinical and Experimental Research*, **1994**, *18*, 1280

[11] Barnett C., Rudd S., Flatt P., Ioannides C., *Biochemical Pharmacology*, **1993**, *45*, 313

[12] Raucy J. L., Lasker J. M., Kraner J. C., Salazar D. E., Lieber C. S., Corcoran G., *Molecular Pharmacology*, **1991**, *39*, 275

[13] Johansson I., Ekstroem G., Scholte B., Puzycki D., Joernvall H., Ingelman-Sundberg M., *Biochemistry*, **1988**, *27*, 1925

[14] DeVore N. M., Meneely K. M., Bart A. G., Stephens E. S., Battaile K. P., Scott E. E., *Febs Journal*, **2012**, *279*, 1621

[15] Li J., Wei D. Q., Wang J. F., Li Y. X., *Journal of Chemical Information and Modeling*, **2011**, *51*, 3217

[16] Porubsky P. R., Battaile K. P., Scott E. E., *Journal of Biological Chemistry*, **2010**, *285*, 22282

[17] *EMBL/GenBank/DDBJ Databases*, **2006**

[18] Hu Y., Oscarson M., Johansson I., Yue Q. Y., Dahl M. L., Tabone M., Arincò S., Albano E., Ingelman-Sundberg M., *Molecular Pharmacology*, **1997**, *51*, 370

[19] Fairbrother K. S., Grove J., de Waziers I., Steimel D. T., Day C. P., Crespi C. L., Daly A. K., *Pharmacogenetics and Genomics*, **1998**, *8*, 543

[20] Solus J. F., Arietta B. J., Harris J. R., Sexton D. P., Steward J. Q., McMunn C., Ihrie P., Mehall J. M., Edwards T. L., Dawson E. P., *Pharmacogenomics*, **2004**, *5*, 895

[21] Liu Y., Liu B. Y., Hao P., Li X., Li Y. X., Wang J. F., *Proteins: Structure, Function, and Bioinformatics*, **2013**, *81*, 945

[22] Shen Z., Cheng F., Xu Y., Fu J., Xiao W., Shen J., Liu G., Li W., Tang Y., *PloS One*, **2012**, *7*, e33500

[23] Porubsky P. R., Meneely K. M., Scott E. E., *Journal of Biological Chemistry*, **2008**, *283*, 33698

[24] *Discovery Studio*, Accelrys Inc., San Diego, CA, **2009**

[25] Olsson M. H., Søndergaard C. R., Rostkowski M., Jensen J. H., *Journal of Chemical Theory and Computation*, **2011**, *7*, 525

[26] Case D. E., Darden T., Cheatham Iii T., Simmerling C., Wang J., Duke R., Luo R., Walker R., Zhang W., Merz K., *AMBER 11*, University of California, San Francisco, **2010**

[27] Frisch M. J., Trucks G. W., Schlegel H. B., Scuseria G. E., Robb M. A., Cheeseman J. R., Scalmani G., Barone V., Mennucci B., Petersson G. A., Nakatsuji H., Caricato M., Li X., Hratchian H. P., Izmaylov A. F., Bloino J., Zheng G., Sonnenberg J. L., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Vreven T., Montgomery Jr. J. A., Peralta J. E., Ogliaro F., Bearpark M., Heyd J. J., Brothers E., Kudin K. N., Staroverov V. N., Kobayashi R., Normand J., Raghavachari K., Rendell A., Burant J. C., Iyengar S. S., Tomasi J., Cossi M., Rega N., Millam N. J., Klene M., Knox J. E., Cross J. B., Bakken V., Adamo C., Jaramillo J., Gomperts R., Stratmann R. E., Yazyev O., Austin A. J., Cammi R., Pomelli C., Ochterski J. W., Martin R. L., Morokuma K., Zakrzewski V. G., Voth G. A., Salvador P., Dannenberg J. J., Dapprich S., Daniels A. D., Farkas Ö., Foresman J. B., Ortiz J. V., Cioslowski J., Fox D. J.. *Gaussian 09. D01*, Gaussian Inc., Wallingford CT, **2009**

[28] Rydberg P., Olsen L., Norrby P. O., Ryde U., *Journal of Chemical Theory and Computation*, **2007**, *3*, 1765

[29] Hornak V., Abel R., Okur A., Strockbine B., Roitberg A., Simmerling C., *Proteins: Structure, Function, and Bioinformatics*, **2006**, *65*, 712

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

[31] Darden T., York D., Pedersen L., *Journal of Chemical Physics*, **1993**, *98*, 10089

[32] Shao J., Tanner S. W., Thompson N., Cheatham T. E., *Journal of Chemical Theory and Computation*, **2007**, *3*, 2312

[33] Humphrey W., Dalke A., Schulten K., *Journal of Molecular Graphics*, **1996**, *14*, 33

[34] Pettersen E. F., Goddard T. D., Huang C. C., Couch G. S., Greenblatt D. M., Meng E. C., Ferrin T. E., *Journal of Computational Chemistry*, **2004**, *25*, 1605

[35] DeLano W. L., *PyMOL*, Schroclinger, New York, **2002**

[36] Amadei A., Linssen A., Berendsen H. J., *Proteins: Structure, Function, and Bioinformatics*, **1993**, *17*, 412

[37] Bakan A., Meireles L. M., Bahar I., *Bioinformatics*, **2011**, *27*, 1575

[38] Xue Q., Zhang J. L., Zheng Q. C., Cui Y. L., Chen L., Chu W. T., Zhang H. X., *Langmuir*, **2013**, *29*, 11135

[39] Cui Y. L., Zheng Q. C., Zhang J. L., Xue Q., Wang Y., Zhang H. X., *Journal of Chemical Information and Modeling*, **2013**, *53*(12), 3308

[40] Swanson J. M. J., Henchman R. H., McCammon J. A., *Biophysical Journal*, **2004**, *86*, 67

[41] Hou T. J., Zhang W., Case D. A., Wang W., *J. Mol. Biol*., **2008**, *376*, 1201

[42] Dapprich S., Komáromi I., Byun K. S., Morokuma K., Frisch M. J., *Journal of Molecular Structure: Theochem*., **1999**, *461*, 1

[43] Vreven T., Mennucci B., da Silva C. O., Morokuma K., Tomasi J., *Journal of Chemical Physics*, **2001**, *115*, 62

[44] Vreven T., Morokuma K., Farkas Ö., Schlegel H. B., Frisch M. J., *Journal of Computational Chemistry*, **2003**, *24*, 760

[45] Maseras F., Morokuma K., *Journal of Computational Chemistry*, **1995**, *16*, 1170

[46] Vreven T., Byun K. S., Komáromi I., Dapprich S., Montgomery J. A., Morokuma K., Frisch M. J., *Journal of Chemical Theory and Computation*, **2006**, *2*, 815

[47] Becke A. D., *J. Chem. Phys*., **1993**, *98*, 5648

[48] Lee C., Yang W., Parr R. G., *Physical Review B*, **1988**, *37*, 785

[49] Cornell W. D., Cieplak P., Bayly C. I., Gould I. R., Merz K. M., Ferguson D. M., Spellmeyer D. C., Fox T., Caldwell J. W., Kollman P. A., *J. Am. Chem. Soc*., **1995**, *117*, 5179

[50] Torrent M., Vreven T., Musaev D. G., Morokuma K., Farkas Ö., Schlegel H. B., *J. Am. Chem. Soc*., **2002**, *124*, 192

[51] Vreven T., Morokuma K., *Theoretical Chemistry Accounts*, **2003**, *109*, 125

[52] Li J., Cross J. B., Vreven T., Meroueh S. O., Mobashery S., Schlegel H. B., *Proteins: Structure, Function, and Bioinformatics*, **2005**, *61*, 246

[53] Yoshizawa K., Shiota Y., *J. Am. Chem. Soc*., **2006**, *128*, 9873

[54] Godfrey E., Porro C. S., de Visser S. P., *Journal of Physical Chemistry A*, **2008**, *112*, 2464

[55] Lundberg M., Kawatsu T., Vreven T., Frisch M. J., Morokuma K., *Journal of Chemical Theory and Computation*, **2008**, *5*, 222

[56] Kong X., Ouyang S., Liang Z., Lu J., Chen L., Shen B., Li D., Zheng M., Li K. K., Luo C., *PLoS One*, **2011**, *6*, e25444

[57] Hay P. J., Wadt W. R., *Journal of Chemical Physics*, **1985**, *82*, 270

[58] Grimme S., *Journal of Computational Chemistry*, **2004**, *25*, 1463

[59] Grimme S., *Journal of Computational Chemistry*, **2006**, *27*, 1787

[60] Grimme S., Antony J., Ehrlich S., Krieg H., *Journal of Chemical Physics*, **2010**, *132*, 154104

[61] Johnson E. R., Keinan S., Mori-Sanchez P., Contreras-Garcia J., Cohen A. J., Yang W., *J. Am. Chem. Soc*., **2010**, *132*, 6498

[62] Lu T., Chen F., *Journal of Computational Chemistry*, **2012**, *33*, 580

[63] Raeburn J., Alston B., Kroeger J., McDonald T. O., Howse J. R., Cameron P. J., Adams D. J., *Mater. Horiz*., **2014**, *1*, 241

[64] Sun Z., Li Z., He Y., Shen R., Deng L., Yang M., Liang Y., Zhang Y., *J. Am. Chem. Soc*., **2013**, *135*, 13379

[65] Joosten R. P., Te Beek T. A., Krieger E., Hekkelman M. L., Hooft R. W., Schneider R., Sander C., Vriend G., *Nucleic Acids Research*, **2011**, *39*, D411

[66] McGaughey G. B., Gagné M., Rappé A. K., *Journal of Biological Chemistry*, **1998**, *273*, 15458

[67] Wang Y., Zheng Q. C., Zhang J. L., Cui Y. L., Xue Q., Zhang H. X., *Journal of Molecular Modeling*, **2013**, *19*, 5213

[68] Collom S. L., Laddusaw R. M., Burch A. M., Kuzmic P., Perry M. D., Miller G. P., *Journal of Biological Chemistry*, **2008**, *283*, 3487

[69] Wang Y., Zheng Q. C., Kong C. P., Tian Y., Zhan J., Zhang J. L., Zhang H. X., *Mol. Biosyst.*, **2015**, *11*(1) 252

[70] Xue W., Jin X., Ning L., Wang M., Liu H., Yao X., *Journal of Chemical Information and Modeling*, **2012**, *53*, 210

[71] Yang Y., Shen Y., Liu H., Yao X., *Journal of Chemical Information and Modeling*, **2011**, *51*, 3235

[72] Sun X., Feng Z., Zhang L., Hou T., Li Y., *PloS One*, **2014**, *9*, e107696

[73] Wu Y. J., Cui Y. L., Zheng Q. C., Zhang H. X., *Chem. J. Chinese Universities*, **2014**, *35*(12), 2605

[74] Wu Y., Zheng Q., Xu Y., Chu W., Cui Y., Wang Y., Zhang H., *Chem. Res. Chinese Universites*, **2014**, *30*(6), 1011

[75] Xu Y., Cui Y. L., Zheng Q. C., Zhang H. X., Sun C. C., *Chem. J. Chinese Universities*, **2013**, *34*(5), 1226

[76] Meng X. Y., Li Z., Niu R. J., Zhang H. X., Zheng Q. C., *Chem. Res. Chinese Universites*, **2012**, *28*(1), 137