

# Insight into Shared Properties and Differential Dynamics and Specificity of Secretory Phospholipase A<sub>2</sub> Family Members

Shan Zhang, Weikang Gong, Zhongjie Han, Yang Liu, and Chunhua Li\*



Cite This: *J. Phys. Chem. B* 2021, 125, 3353–3363



Read Online

ACCESS |

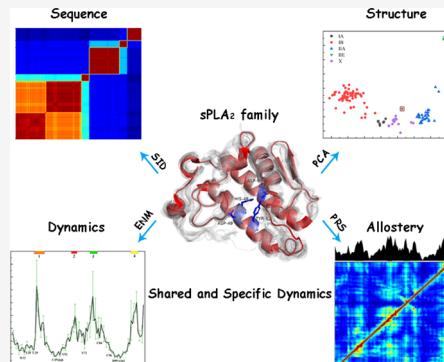
Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Understanding generic mechanisms of functions shared by the secretory phospholipase A<sub>2</sub> (sPLA<sub>2</sub>) family involved in the lipid metabolism and cell signaling and the molecular basis of function specificity for family members is an intriguing but challenging problem for biologists. Here, we explore the issue through extensive analyses using a combination of structure-based methods and bioinformatics tools on 130 sPLA<sub>2</sub> family members. The principal component analysis of the structure ensemble reveals that the enzyme has an open–close motion which helps widen the substrate binding channel, facilitating its binding to phospholipid. Performing elastic network model and sequence analyses found that the residues critical for family functions, such as cysteine and catalytic residues, are highly conserved and undergo minimal movements, which is evolutionarily essential as their perturbation would impact the function, while the four residue regions involved in the association with the calcium ion/membrane are lowly conserved and of high mobility and large variations in low-to-intermediate frequency modes, which reflects the specificity of members.

The analyses from perturbation response scanning also reveal that the above four regions with high sensitivity to an external perturbation are member-specific, suggesting their different roles in allosteric modulation, while the minimal sensitive residues are the shared characteristics across family members, which play an important role in maintaining structural stability as the folding core. This study is helpful for understanding how sequences, structures, and dynamics of sPLA<sub>2</sub> family members evolve to ensure their common and specific functions and can provide a guide for accurate design of proteins with finely tuned activities.



## INTRODUCTION

Secretory phospholipases A<sub>2</sub> (sPLA<sub>2</sub>s), a subfamily of lipolytic enzymes, accounting for more than one-third of phospholipases A<sub>2</sub> (PLA<sub>2</sub>s), play a central role in the cellular lipid metabolism and signaling.<sup>1</sup> Interestingly, different sPLA<sub>2</sub> isoforms have their own individual hydrolysis preferences for phospholipids, while they share the capacity to catalyze hydrolysis of the sn-2 ester bond of phospholipids, thereby producing different lipid mediators—signaling molecules involved in many inflammatory diseases including asthma, arthritis, and cancer.<sup>2</sup> Thus, an intriguing problem is how sequences, structures, and dynamics of sPLA<sub>2</sub> family members evolve to ensure their common and specific functions. The insight into the issue is not only helpful for the understanding of the related mechanisms but also helpful for protein design and the design of sPLA<sub>2</sub> enzyme inhibitors for inflammatory diseases.

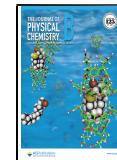
To date, the crystal structures have been solved for many sPLA<sub>2</sub> isoforms purified from snake venom, humans, porcine and bovine pancreas, plants, and microbes. A total of 16 sPLA<sub>2</sub> groups (IA, IB, IIA, IIB, IIC, IID, IIE, IIF, III, V, IX, X, XIA, XIB, XII, and XIV) have been identified, which are characterized by at least 6 highly conserved disulfide bonds, a His-Asp catalytic dyad, and a calcium-binding loop (Xxx-Cys-Gly-Xxx-Gly-Gly).<sup>3</sup> The enzymes are classified into different

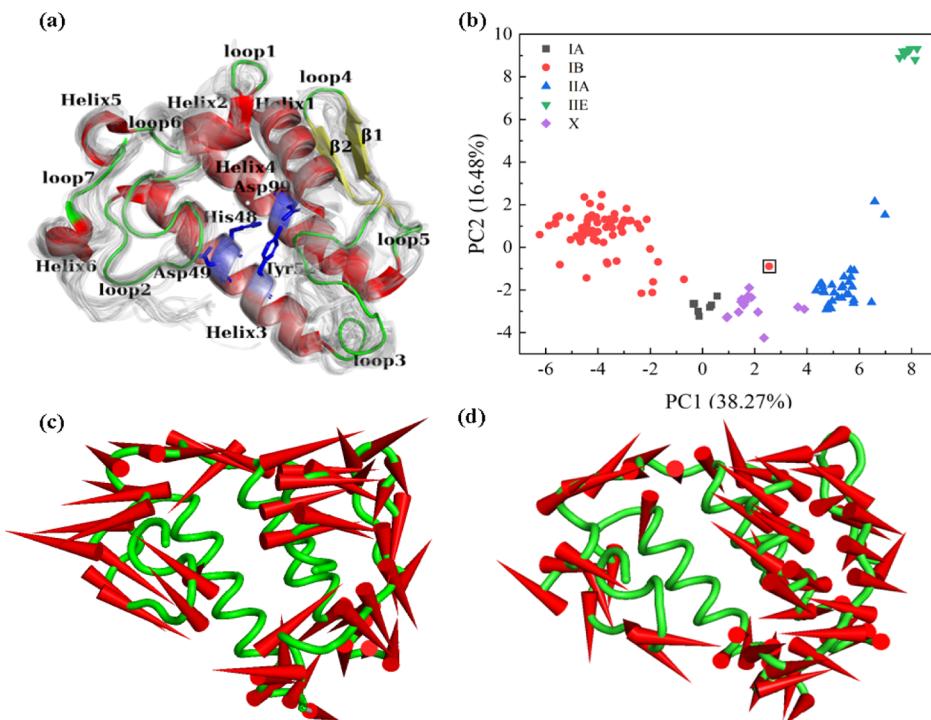
groups based on the amino acid sequence, molecular weight, and disulfide bonding pattern. Despite pronounced differences in amino acid sequences, sPLA<sub>2</sub> enzymes preserve a similar folding configuration, especially for those from groups I, II, V, and X, which contain three long  $\alpha$ -helices, two  $\beta$ -strands, and a conserved Ca<sup>2+</sup>-binding loop (loop2) (Figure 1a). The sPLA<sub>2</sub> catalytic cycle includes four steps: association of the enzyme with the membrane through its substrate binding region, a single phospholipid molecule extraction from the membrane and binding to the enzyme's binding pocket, enzymatic hydrolysis of the sn-2 ester bond of the phospholipid, and release of the hydrolysis product.<sup>4</sup> In the cycle, the most important event is the enzyme's association with the membrane through its substrate binding region, during which the membrane, acting as an allosteric ligand, binds at the allosteric site of the enzyme's interfacial surface, shifting the enzyme's conformation from the closed state to an open state,

Received: February 12, 2021

Revised: March 12, 2021

Published: March 29, 2021





**Figure 1.** (a) Superposition of 130 sPLA<sub>2</sub> family member structures from groups IA, IB, IIA, IIE, and X. The member with PDB ID 1vkq is the reference structure with helices colored in red, loops in green, and β-strands in yellow. The catalytic residues His48, Asp49, Tyr52, and Asp99 (blue) are shown in the stick model. (b) Distribution of 130 sPLA<sub>2</sub> family members in the subspace spanned by PC1 and PC2 capturing more than 54% of the total variation. (c,d) porcupine plots generated from PC1 and PC2, respectively. The orientation of the arrows pointing from backbone atoms denotes the direction of the eigenvector (PC), and the size of each arrow indicates the magnitude of the corresponding eigenvalue.

which is critical for the subsequent enzyme's binding with its preferred phospholipid.<sup>4</sup> Experiments have shown that different sPLA<sub>2</sub> isoforms have their own most suitable substrates including phosphatidylglycerol (PG), phosphatidylethanolamine (PE), phosphatidylcholine, and phosphatidylserine (PS).<sup>5</sup> These phospholipid substrates are different in volume size and electrostatic properties. Thus, the structure and dynamics of the enzyme's interfacial binding surface and catalytic sites have been the hot spots of study.

Many experimental and theoretical research studies focus on the binding key sites and binding behavior of the sPLA<sub>2</sub> enzyme to the membrane. On the experimental side, Huang et al. observed that sPLA<sub>2</sub> is located on the membrane surface rather than penetrating the membrane bilayer using the monolayer technique.<sup>6</sup> Burke et al. used hydrogen–deuterium exchange mass spectrometry to detect group IA phospholipase A<sub>2</sub> from cobra venom and found the rigid characteristic of the two extensively disulfide-bonded helices which bear the catalytic function.<sup>7</sup> On the theoretical side, all-atomic molecular dynamics (MD) simulation has been a widely used computational tool to explore protein dynamic features in atomic details.<sup>8</sup> With MD simulations, Manukyan studied the effects of the solvent and membrane environments on the conformation of human sPLA<sub>2</sub> IIA.<sup>9</sup> Ramakrishnan et al. compared the dynamics of sPLA<sub>2</sub> from russell's viper at inhibitor-bound and free states and explored the mechanism of the activity abolishment due to the inhibitor binding.<sup>10</sup> To identify potent inhibitors against sPLA<sub>2</sub>, Chinnasamy et al. used virtual screening and molecular docking methods to select its potent inhibitors from the Traditional Chinese Medicine Database (TCM) and additionally provided a deeper structural insight into the binding key residues of Arg30, Gly31, and

Tyr63.<sup>11</sup> Although the research studies above help to reveal some key sites and interactions between sPLA<sub>2</sub> enzymes from a certain group and membrane, they only focus on one single group of sPLA<sub>2</sub> and have not explored the differences in structure dynamics across different groups, which is critical for the understanding of the subtle mechanisms of different groups of sPLA<sub>2</sub> enzymes in specific substrate recognition, despite their similar folding conformations.

MD simulation is a time-consuming method and has difficulty in investigating the large-scale functional motions of proteins. To address the issue, several coarse-grained models have been recently proposed. Among them, the elastic network model (ENM) has been proven to be a particularly effective computational technique for investigating the function-relevant motions of proteins.<sup>12</sup> Two mostly used ENM methods, the Gaussian network model (GNM)<sup>13</sup> and the anisotropic network model (ANM),<sup>14</sup> are simple yet effective methods to explore the intrinsic dynamics of proteins. It has been proven in numerous application studies that the low-frequency motion modes calculated by GNM and ANM represent the large-scale collective motions usually relevant to protein functions.<sup>15</sup> Utilizing low-frequency modes, our group developed ENM-based methods to study structure flexibility<sup>16,17</sup> and unfolding<sup>18,19</sup> of protein/RNA molecules and protein allostery.<sup>20</sup> Combined with the ENM model, perturbation response scanning (PRS) analysis<sup>21</sup> was proposed to obtain protein dynamics and allosteric properties, and currently, the method has been widely applied to identify the key residues in allosteric control and long-range communications for large protein assemblies.<sup>22,23</sup> Additionally, a systematic method of approach was developed by Bahar et al. for characterizing the shared (signature) as well as unique

structural and dynamic properties of a protein family using a combination of structure-based models and methods.<sup>24</sup> Still aiming at a protein family, Tobi developed an ENM-based method to make a function classification for globin family members based on their dynamics' similarity.<sup>25</sup> Besides, in 2020, an insightful review by Bahar et al. revealed the evolutionary constraints on structure dynamics of a protein family to achieve the required functions: the slowest motion modes are conserved by a protein family for the common function; the low-to-intermediate frequency (LTIF) modes reflect the function specificity across family members; the fastest frequency modes ensure the core stability, which is the basics of functional specialization.<sup>26</sup>

Besides the ENM model, currently, the protein structure network (PSN) has been widely used in the studies of protein folding, allosteric transition, and key site prediction.<sup>27</sup> Once the network for the protein structure is constructed, various network parameters can be calculated to decipher both local and overall structural characteristics, including degree, clustering coefficient, and betweenness. Combined with some dynamics methods, the PSN shows a strong power in the study of protein allosteric communication and dynamic coupling.<sup>28</sup>

In this work, inspired by Bahar's and Dob's works, we examine the signature dynamics of the sPLA<sub>2</sub> family as well as the distinctive features of family members required to achieve their specific catalytic activities. The key residues are identified which play an important role in the dynamics related to specific recognition and interactions of sPLA<sub>2</sub> members with their substrates.

## METHODS

### Database of Secretory Phospholipase A<sub>2</sub> (sPLA<sub>2</sub>) Structures.

For the sPLA<sub>2</sub> family, considering the availability of the X-ray structures and explicit grouping, we collected and downloaded 130 member structures (Table S1) from the Protein Data Bank (PDB),<sup>29</sup> which belong to groups IA, IB, IIA, IIE, and X. Based on the root-mean-square deviation (RMSD) between two structures, the representative member (with PDB code 1a3f for group IA, 1vkq for group IB, 1kvo for group IIA, 5wzu for group IIE, and 6g5j for group X) of each group was selected which has the minimum average RMSD from the others in the group.

**Principal Component Analysis.** The slow and functional motions of biomolecules can be extracted by principal component analysis (PCA), which is a statistical method based on covariance analysis.<sup>30</sup> The orthogonal eigenvectors and the corresponding eigenvalues can be obtained from the covariance matrix. The eigenvectors, also called principal components (PCs), indicate the directions of concerted motions. The corresponding eigenvalue describes the magnitude of the motion along the direction. In our study, PCA is performed using the R package Bio3D<sup>31</sup> in combination with Dali Sever<sup>32</sup> for structure alignment.

**Elastic Network Model.** In GNM (a special case of ENM), a protein structure is modeled as an elastic network where a residue is replaced with several nodes (here, the C<sub>α</sub> atom) and the interactions between the nodes within a given cutoff distance  $r_c$  (7 Å adopted here) are represented as springs with a uniform force constant. By this simplification, the total internal potential energy of the network of  $N$  nodes can be written as

$$V_{\text{GNM}} = \frac{1}{2} \gamma [\Delta \mathbf{R}^T (\Gamma \otimes \mathbf{E}) \Delta \mathbf{R}] \quad (1)$$

where the column vector  $\Delta \mathbf{R}$  represents the fluctuation of the  $N$  nodes,  $\mathbf{E}$  is the unitary matrix,  $\otimes$  is the matrix direct product,  $\gamma$  is the force constant of the springs, and  $\Gamma$  is the  $N \times N$  symmetric Kirchhoff matrix, the elements of which are defined as follows

$$\Gamma_{ij} = \begin{cases} -1 & \text{if } i \neq j, R_{ij} \leq r_c, \\ 0 & \text{if } i \neq j, R_{ij} > r_c, \\ -\sum_{j,j \neq i}^N \Gamma_{ij} & \text{if } i = j \end{cases} \quad (2)$$

where  $R_{ij}$  is the distance between the  $i$ th and  $j$ th nodes. Then, the mean square fluctuation (MSF) of each node and the fluctuation cross-correlation between different nodes are expressed as

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i \rangle = \frac{3k_B T}{\gamma} [\Gamma^{-1}]_{ii} \quad (3)$$

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = \frac{3k_B T}{\gamma} [\Gamma^{-1}]_{ij} \quad (4)$$

where  $T$  is the absolute temperature and  $k_B$  is the Boltzmann constant.

GNM can provide the amplitudes of residue fluctuations but no information about their directions, and this information is considered in ANM. In ANM, the total potential energy of the network can be written as

$$V_{\text{ANM}} = \frac{1}{2} \gamma \sum_{i,j}^N (\mathbf{R}_{ij} - \mathbf{R}_{ij}^0)^2 \quad (5)$$

where  $\mathbf{R}_{ij}$  and  $\mathbf{R}_{ij}^0$  refer to the instantaneous and equilibrium distances between nodes  $i$  and  $j$ , respectively. The protein dynamical properties are determined by a Hessian matrix  $\mathbf{H}$  whose element is a submatrix with a size of  $3 \times 3$ . The submatrix  $\mathbf{h}_{ij}$  is calculated as the matrix of second-order derivatives of the potential with respect to the Cartesian coordinates of the nodes. When  $i \neq j$ , the corresponding  $\mathbf{h}_{ij}$  is

$$\mathbf{h}_{ij} = \begin{pmatrix} \frac{\partial^2 V}{\partial X_i \partial X_j} & \frac{\partial^2 V}{\partial X_i \partial Y_j} & \frac{\partial^2 V}{\partial X_i \partial Z_j} \\ \frac{\partial^2 V}{\partial Y_i \partial X_j} & \frac{\partial^2 V}{\partial Y_i \partial Y_j} & \frac{\partial^2 V}{\partial Y_i \partial Z_j} \\ \frac{\partial^2 V}{\partial Z_i \partial X_j} & \frac{\partial^2 V}{\partial Z_i \partial Y_j} & \frac{\partial^2 V}{\partial Z_i \partial Z_j} \end{pmatrix} \quad (6)$$

When  $i = j$ , the submatrix is

$$\mathbf{h}_{ii} = - \sum_{i \neq j} \mathbf{h}_{ij} \quad (7)$$

In ANM, the two corresponding dynamical properties can be written as

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i \rangle = \frac{k_B T}{\gamma} (\mathbf{H}_{3i-2,3i-2}^{-1} + \mathbf{H}_{3i-1,3i-1}^{-1} + \mathbf{H}_{3i,3i}^{-1}) \quad (8)$$

$$\langle \Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j \rangle = \frac{k_B T}{\gamma} (\mathbf{H}_{3i-2,3j-2}^{-1} + \mathbf{H}_{3i-1,3j-1}^{-1} + \mathbf{H}_{3i,3j}^{-1}) \quad (9)$$

The cross-correlation is normalized as

$$C_{ij} = \frac{\langle \Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j \rangle}{[\langle (\Delta\mathbf{R}_i^2) \rangle \times \langle (\Delta\mathbf{R}_j^2) \rangle]^{1/2}} \quad (10)$$

Cross-correlation values range from  $-1$  to  $1$ . The positive ones indicate that the residues move in the same direction, and the negative ones indicate that they move in the opposite directions. The higher the absolute value is, the more the two residues are correlated. The zero value implies that the motions of residues are completely uncorrelated.

**Perturbation Response Scanning Approach.** The PRS<sup>21</sup> approach, which is based on the linear response theory (LRT),<sup>33</sup> was designed to deduce protein allosteric properties. PRS allows for the calculation of the response of residue  $k$  to perturbation at residue  $i$ . The  $3N$ -dimensional vector  $\Delta\mathbf{R}$  of node displacements in response to the application of a perturbation ( $3N$ -dimensional force vector  $\mathbf{F}$ ) obeys Hooke's law  $\mathbf{F} = \mathbf{H} \cdot \Delta\mathbf{R}$ , where  $\mathbf{H}$  is the  $3N \times 3N$  Hessian matrix in ANM theory. The idea in PRS is to exert a force of a given magnitude on one residue in the network at a time and observe the response of the overall network. The force exerted on residue  $i$  is expressed as

$$\mathbf{F}^i = (000 \cdots \Delta\mathbf{F}_x^i \Delta\mathbf{F}_y^i \Delta\mathbf{F}_z^i \cdots 000)^T \quad (11)$$

and the resulting response is

$$\Delta\mathbf{R}^i = \mathbf{H}^{-1} \mathbf{F}^i \quad (12)$$

where  $\Delta\mathbf{R}^i$  is a  $3N$ -dimensional vector that describes the displacements of all the residues away from their equilibrium positions (in  $N$  blocks of dimension 3, each) in response to the exerted force  $\mathbf{F}^i$ , which are nonzero only for the three terms related to residue  $i$ .

Here, the average value of the squared residue  $k$  displacements  $\langle \|\Delta\mathbf{R}_k^{(i)}\|^2 \rangle$  in response to multiple exerted perturbations on residue  $i$  is taken as the sensitivity of residue  $k$  to the perturbation at residue  $i$ .<sup>34</sup> The multiple forces are along seven directions, that is,  $x$ -;  $y$ -;  $z$ -; both  $x$ - and  $y$ -; both  $x$ - and  $z$ -; both  $y$ - and  $z$ -; and all  $x$ -,  $y$ -, and  $z$ -directions. Repeating (scanning) this procedure for all sites yields a response matrix  $\mathbf{P}$  with a size of  $N \times N$ , each column of which provides a measure of the sensitivities of all residues to the perturbation at the residue corresponding to the column. The average over the columns of the normalized  $\mathbf{P}$  yields the sensitivity profile, with peaks therein designated as sensors often corresponding to the functional residues involved in the execution of allosteric structural changes.<sup>34</sup>

**Multiple Sequence Alignment and Coevolution Analysis.** A set of sPLA<sub>2</sub> sequences were collected through blastp search<sup>35</sup> in the nonredundant protein sequence database for residue coevolution analysis. Before the performance of a multiple sequence alignment (MSA), they were refined by removing the sequences belonging to other proteins or with more than 80% sequence identity,<sup>36</sup> finally resulting in 789 sequences. The MSA is carried out using ClustalX<sup>37</sup> with default parameters. The analysis of residue coevolution is performed based on the mutual information (MI) between their positions in the MSA using the MISTIC approach.<sup>38</sup>

## Construction of the Protein Structure Network.

Protein structure formation and function exertion rely on the complex network of inter-residue interactions to some extent.<sup>27</sup> In PSN, a protein is converted into an indirect and unweighted graph which looks like a network of nodes, where nodes are amino acid residues and edges are inter-residue interactions (here, a distance cutoff of  $8.0 \text{ \AA}$  is adopted). The degree of a node is the number of other nodes to which it is connected.

## RESULTS AND DISCUSSION

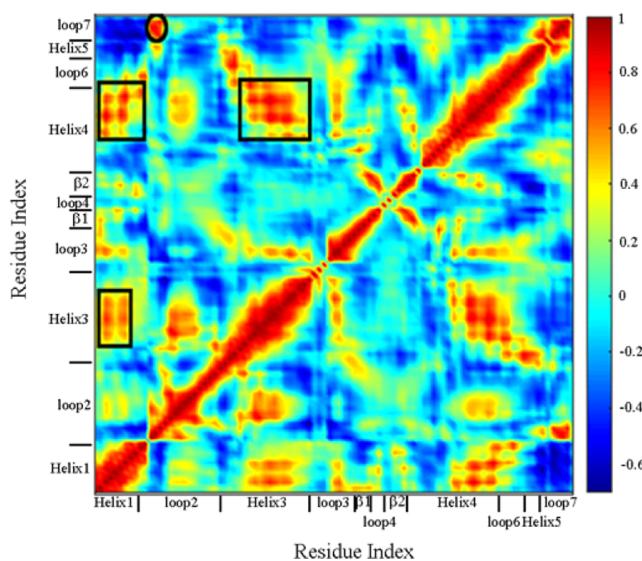
**Statistical Analyses of Sequence Identity and Structural Similarity for the 130 sPLA<sub>2</sub> Family Members.** We first selected the reference structure from the 130 sPLA<sub>2</sub> family members based on the RMSD between two structures, which is 1vkq (PDB ID) from group IB having the minimum average RMSD from the others in the family. Figure S1a displays the distribution of percent sequence identity (SID) of the 130 members with respect to the reference (with the heat map of SID between sequences shown in Figure S2a). The percent SID varies in the range of 35–99. Despite the difference in SID, family members share the same fold (Figure 1a), as evidenced by their RMSDs less than  $\sim 2.0 \text{ \AA}$  (Figure S1b). Note that the largest structural variations occur at loop2, loop3, loop4, and loop6 (Figure 1a). In the following, we will characterize the signature dynamics of the family as well as distinctive features of members required to achieve their specific catalytic and allosteric activities.

### Principal Component Analysis of sPLA<sub>2</sub> Structures.

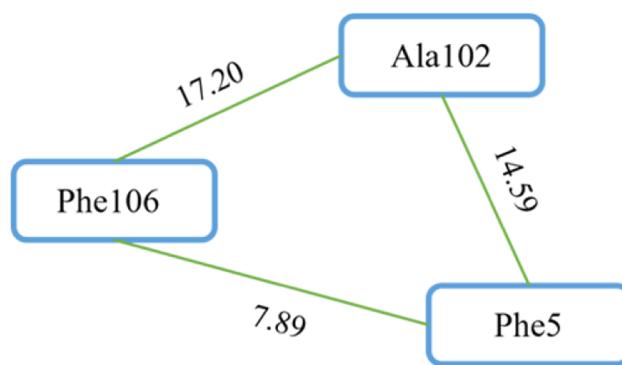
The conformational differences among sPLA<sub>2</sub> structures can be determined from PCA. The first PC (PC1) describes the direction of maximal variance of the structure distribution, succeeded by PC2, and so on. Of interest is to view the distribution of sPLA<sub>2</sub> structures in the subspace spanned by PC1 and PC2 (Figure 1b), which permits us to discriminate or cluster the conformations based on their most distinctive structural dissimilarities or similarities. From Figure 1b, PC1 and PC2 clearly divide the structure ensemble into five well-defined clusters. Interestingly, they just correspond to the five groups classified by SID (Figure S2a). In the PC1 direction, groups IA and IB and groups IIA and IIE are near to each other, respectively, while the latter two are separated evidently in the PC2 direction. From Figure S2a, the former and the latter have higher SID than other pairs of groups (in agreement with their biological naming), with the former's SID slightly higher than the latter's, consistent with the former's smaller structural differences than the latter's (Figure S2b). The higher similarity in sequence and structure between groups IA and IB than between groups IIA and IIE may suggest that the former two are more homologous to each other than the latter two.

In addition, note that PCA clearly divides all family members except for the one with PDB ID 3elo (Figure 1b) belonging to group IB but clustering into group X. All structures in group X are from humans, and those in group IB are from porcine or bovine except for 3elo belonging to humans, possibly implying certain similarity of 3elo to group X. For group IB, we also found that there is a slight difference between the structures from bovine and porcine (Figure S2b), which is well consistent with the evident difference in SID between them (Figure S2a). Thus, the accurate separation of different groups of sPLA<sub>2</sub> family members indicates that the structural features elucidated by PCA for them can be well traced back to their sequence similarities.

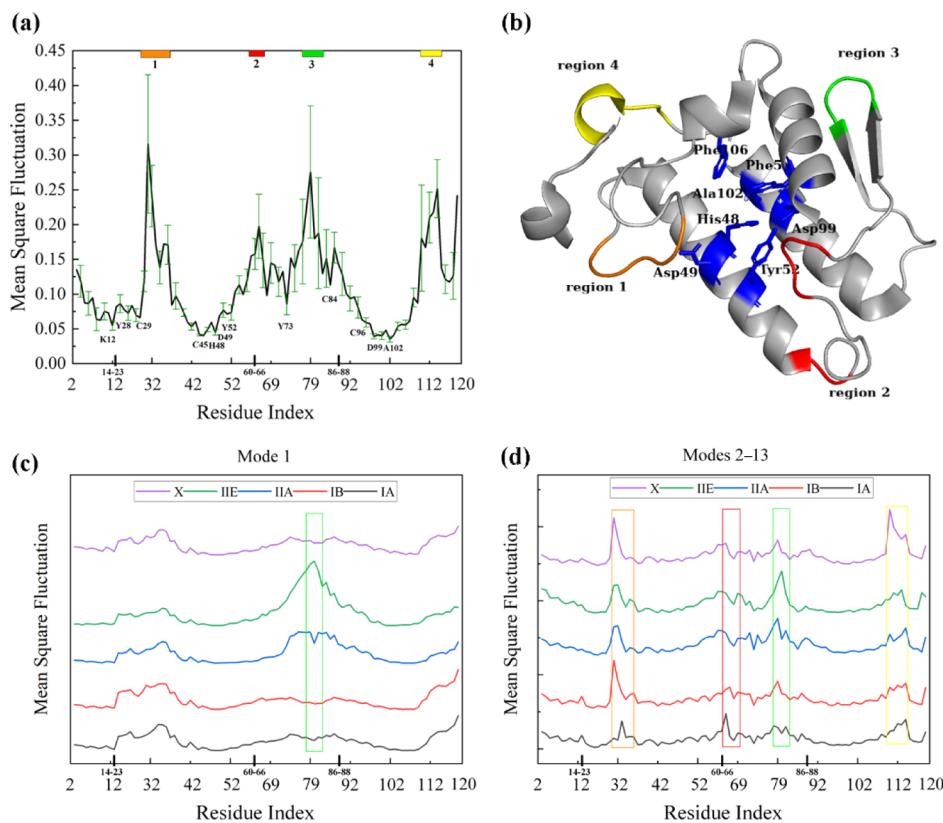
(a)



(b)



**Figure 2.** (a) Residue fluctuation cross-correlations calculated by ANM (with a distance cutoff of 17.0 Å and 29 slowest motion modes considered that contribute more than 50% to the residue fluctuation) on the reference structure 1vkq. (b) Coevolutionary residues with the numbers on lines connecting two residues being their MI values.



**Figure 3.** Residue fluctuations obtained by GNM for sPLA<sub>2</sub> family member structures. (a) Average of MSFs (black line) and standard deviations (green line) obtained by the GNM slowest 13 modes for 130 sPLA<sub>2</sub> family member structures. Residues along the abscissa refer to the reference one 1vkq (residues Trp3-Asn119 excluding Pro14-Asn23, Ser60-Asp66, and Ser86-Asn88 indicated by the black break points). The labeled residues indicate the sites with minimal fluctuations and minimal standard deviations. The regions labeled with 1–4 (colored bars along the upper abscissa) centered around Gly30-Thr36, Leu58-Thr70 (excluding Ser60-Asp66), Ser78-Glu81, and Pro110-His115, respectively, display the highest differences among family members. (b) Regions 1–4, catalytic residues (His48, Asp49, Tyr52, and Asp99) and the coevolutionary residue triplet (Phe5-Ala102-Phe106) mapped on the reference structure. (c,d) Residue fluctuation profiles contributed by the first slowest mode and the first 2–13 modes for the five groups' representative structures, respectively.

As is known, the structure ensemble of protein family members solved at different states can reflect the dynamics of the protein at the equilibrium state to some extent.<sup>39</sup> Thus, PCA can give insights into the functional dynamics of the sPLA<sub>2</sub> enzyme. To quantitatively understand the movements captured by PC1 and PC2, porcupine plots were generated using the extreme projections of all the structures on PC1 and PC2, as shown in Figure 1c,d, respectively. Among all the secondary structures, helix3 and helix4 (core region), where the catalytic sites (His48, Asp49, Tyr52, and Asp99 shown in Figure 1a) and disulfide bonds (involving Cys44, Cys45, Cys51, Cys91, Cys96, and Cys98) important for stability are located, are of almost minimum mobility in the first two principal movement directions, which is the dynamic requirement for the enzyme to exert catalytic functions (see detailed analyses in the section of GNM analysis of sPLA<sub>2</sub> family members). Additionally from Figure 1c, the enzyme has an open–close motion, consistent with the experimental observation,<sup>40</sup> which helps widen the substrate binding channel and meanwhile makes helix3 and helix4 exposed to the environment (membrane), thereof facilitating the catalytic sites' substrate recognition.

**Residue Fluctuation Cross-Correlations of the Reference Structure.** In order to detect the functional movement coupling between residues in sPLA<sub>2</sub>, we calculated the residue fluctuation cross-correlations according to eq 10 with ANM on the reference structure 1vkq, with the results shown in Figure 2a.

From Figure 2a, helix5 and loop7 are found to be strongly negatively correlated with helix1 and positively correlated with the beginning part of loop2, which is also observed in the motions along PC1 and PC2 (Figure 1c,d). These regions are located around the catalytic sites, whose motions contribute partially to the enzyme's open–close movement. In addition, for helix3 and helix4 bearing the catalytic role, a strong positive correlation is observed between helix3 residues Asp42-Tyr52 and helix4 residues Cys96-Ser107, of which His48, Asp49, Tyr52, and Asp99 make the catalytic network,<sup>41</sup> which ensures the stable exertion of catalytic functions. Additionally, helix1 Trp3-Lys12 are found to be positively correlated with helix3 and helix4, with the latter partially due to the formation of a hydrophobic triad (Phe5-Ala102-Phe106), which is supported by our coevolutionary analysis (Figure 2b) and a previous study mentioned in the following.<sup>36</sup>

As a whole, the strong correlated motions around the catalytic sites help the enzyme bind with the membrane and phospholipid, facilitating the catalytic reaction.

**GNM Analysis of sPLA<sub>2</sub> Family Members.** In ENM, the slowest or global modes provide robust information on the large amplitude (often allosteric) and collective motions encoded in the protein structure.<sup>42</sup> In order to detect sPLA<sub>2</sub> family members' shared and specific flexibility related to their functions, we constructed the GNMs for the 130 sPLA<sub>2</sub> family members (see Methods), calculated the MSFs of residues based on the first 13 slowest motional modes contributing more than 50% to residue fluctuations, and also obtained their averages and standard deviations, with the results shown in Figure 3a. The curve shown in Figure 3a is usually called the signature profile.

**Shared Dynamics of sPLA<sub>2</sub> Family Members.** From Figure 3a, we first focus on the residues that represent similar dynamic fluctuations (signature dynamics), that is, have small standard deviations, and interestingly, these residues including

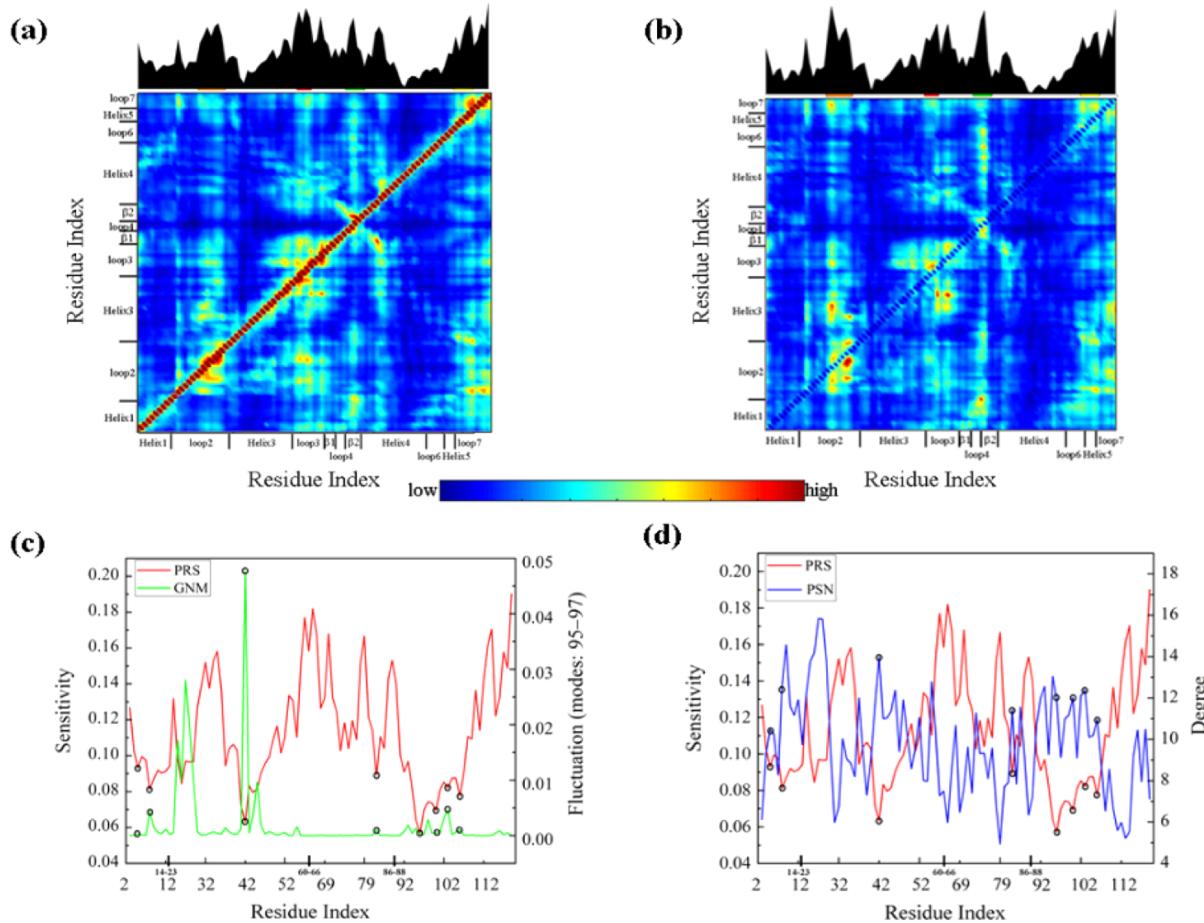
Lys12, Tyr28, Cys29, Cys45, His48, Asp49, Tyr52, Tyr73, Cys84, Cys96, Asp99, and Ala102 consistently occupy minima in the signature profile.

Among them, the four residues His48, Asp49, Tyr52, and Asp99 make up the catalytic network by coordinating the nucleophilic water molecule and stabilizing the oxyanion intermediate (essential for the enzyme's catalytic reaction) in the sPLA<sub>2</sub> family.<sup>10</sup> The conserved His48-Asp99 catalytic dyad occupies minima (with minimal mobility), which is in accord with the precise and tight positioning of catalytic residues—a requirement for mechanochemical activity of enzymes.<sup>43</sup> Additionally, the four cysteine residues (Cys29, Cys45, Cys84, and Cys96) highly conserved (Figure S4a) in the sPLA<sub>2</sub> family, involved in the formation of disulfide bridges (Cys29-Cys45, Cys84-Cys96), mainly serve to maintain the structural stability of the core helical region where the catalytic sites are located. Finally, Tyr28 and Tyr73 are also conserved in the family, of which Tyr73 is involved in the aromatic–aromatic interaction with the position 75 Tyr/Phe/Try residue, very important for the enzyme's conformational stability.<sup>44</sup> As a whole, these shared residue positions either making up the common catalytic network or playing an important role in stabilizing enzyme structures of the sPLA<sub>2</sub> family therefore consistently occupy minima (with minimal mobility) in the signature profile, which can be traced back to their high sequence conservation and structural similarities.

Then, we focus on the peak regions in the signature profile as they usually correspond to ligand recognition sites, and their variations among members may reflect member-specific ligand binding.<sup>45</sup> Indeed, it is the case that these peak regions consistently have a large standard deviation and are involved in membrane/ligand recognition. We label them with 1–4, as shown in Figure 3a (also see Figure 3b for them mapped on the reference structure). Region 1 located in loop2 is a Ca<sup>2+</sup>-binding loop, region 2 situated in loop3 is involved in membrane binding, region 3 belongs to loop4 whose function is not clear now, and region 4 belongs to loop6 and helix5 that are also involved in membrane binding.<sup>40</sup> In the following part, we will further analyze the member-specific properties of these four regions.

**Specific Dynamics of sPLA<sub>2</sub> Family Members.** Shared and specific fluctuations among protein family members can be highlighted by the slowest motion modes and LTIF modes.<sup>26</sup> The former typically relate to functional (or allosteric) changes in the structure, robustly shared among family members, while the latter reflect differentiated motions (functional specificity) across protein family members.<sup>46</sup> In order to detect the shared and specific properties, for the representative structures for five sPLA<sub>2</sub> groups, we dissected the GNM-calculated mobility profiles in two frequency regimes: the first slowest motion mode (Figure 3c) and the LTIF modes (modes 2–13) (Figure 3d). From Figure 3c, the curves from all representatives maintain the same generic shape for different regions except for region 3 (connecting two  $\beta$ -strands) that exhibits some variations among the representatives, especially for group IIE's region 3 obviously different from the others'. The preceding PCA also indicates the structurally large difference of group IIE from others. In sequence, region 3's high variability can also be found from the sequence alignment (Figure S3). The possible function of region 3 is deduced in the following.

From Figure 3d obtained from the LTIF modes (2–13 modes), we focus in particular on the four regions 1–4 depicted in Figure 3a,b, which are involved in membrane/



**Figure 4.** Evaluation of the role of sPLA<sub>2</sub> residues as sensors of allosteric signals and the variations among family members. (a) Average PRS heat map and (b) its standard deviation among sPLA<sub>2</sub> family members. The bar plot on the upper abscissa (sensitivity) describes the propensity of residues to serve as sensors. (c) Residue sensitivity profile from the average PRS and the average fluctuation profile of 130 sPLA<sub>2</sub> family members obtained by the GNM fastest 3 motion modes (94–96). (d) Residue sensitivity profile from the average PRS and the average degree profile of 130 sPLA<sub>2</sub> family members obtained by the PSN method. The residues marked with black circles occupy minima in the sensitivity profile and maxima in the average fluctuation profile (also in the average degree profile).

ligand binding, reflecting the member-specific dynamic features. Region 1 participates in calcium ion binding with six to eight oxygen atoms from its backbone carbonyls, and the bound calcium ion is further constrained near the catalytic sites partially by the catalytic residue Asp49's side chain, which is indispensable for the enzyme catalysis.<sup>9</sup> Region 1's remarkable difference in flexibility among different representatives (Figure 3d) may cause their different binding efficacies/affinities with the calcium ion, which may suggest the differentiated allosteric modulations of region 1 in the enzyme's activities among different sPLA<sub>2</sub> family members.

Both region 2 and region 4 participate in the interactions of the sPLA<sub>2</sub> enzyme with the membrane. Compared with region 1, generally, both regions have a relatively low flexibility which can help the enzyme easily bind with the membrane by hydrophobic and hydrogen bonding interactions.<sup>40</sup> The two regions of some groups present high variability in sequence length and secondary structure type. Group IB has an extra pancreatic loop (residues 62–66) in region 2 that corresponds to the inserted gray part in the reference structure (Figure 3b).<sup>47</sup> Region 4 folds in loop in groups IA and IIA, with a small helix in groups IB, IIE and X.<sup>48</sup> Both regions' variability in sequence, structure, and dynamics makes different family members present specific preferences for membrane types and

phospholipid types accommodated in the catalytic sites. For example, group IIA has more basic amino acids in the membrane binding region, making it have a considerable high affinity and hydrolysis activity for PG.<sup>49</sup> Although there is high SID of group IIE with group IIA, group IIE has no preference for PG due to it having more uncharged amino acids in this region, which partially determines its hydrolysis preference for PE and PS.<sup>50</sup> In summary, the above analyses suggest that regions 2 and 4 play a significant functional role, perhaps via an allosteric modulation of the shape or size of the catalytic channel.

For region 3, the obvious difference in residue fluctuation profiles contributed by LTIF modes presents across different groups. A previous MD simulation study observed the high fluctuation of region 3 in structure 1poe (group IIA) and its evidently different fluctuations in membrane and water environments.<sup>9</sup> This hints that region 3 may participate in the modulation of member-specific membrane recognition.

As a whole, the results from the analyses on the dynamics of sPLA<sub>2</sub> family members that the slowest modes are shared by family members and the LTIF modes can distinguish subfamilies and confer specificity provide new insights into the achievement of the family members' balance between evolutionary adaptability and functional specificity.

**Sequence Conservation and Coevolution of sPLA<sub>2</sub> Family Members.** The sequence evolution analysis adds another dimension for probing the structure and function of proteins. Through sequence analyses, we examine to what extent the shared key residues that play an important role in the structure and function of the sPLA<sub>2</sub> family are sequentially conserved and, conversely, to what extent those responsible for differentiation among members are sequentially variable.

Figure S4a displays the residue conservation profile of the sPLA<sub>2</sub> family obtained from an MSA of 789 sequences. The highest peaks therein indicate the most conserved residues, of which 10 cysteine residues are all involved in the formation of disulfide bonds, Tyr73 participates in the aromatic–aromatic interactions, and the 3 residues (His48, Tyr52, and Asp99) are catalytic residues. The presence of these residues is the shared characteristic of the sPLA<sub>2</sub> family. Dynamically, they generally exhibit minimal mobility in the signature profile (Figure 3a). Thus, these shared residues responsible for stabilizing the structure or catalyzing the hydrolysis reaction are highly conserved and dynamically constrained, which confirms the evolutionary requirement. In contrast, the residues responsible for member differentiation like regions 1–4 distinguished by their large-amplitude member-specific motions exhibit low conservation (Figure S4a). The low mobility of conserved residues and high mobility of variable residues are consistent with the concept of coupling among sequence, structure, and dynamics variations, which has been validated in several earlier studies.<sup>51,52</sup> Such juxtaposition of sequentially conserved (dynamically constrained) and sequentially variable (dynamically flexible) residues appears to be a design feature to mutually support the respective generic and specific properties of sPLA<sub>2</sub> family members.

Next, we focus on the sequence coevolution. Generally, the most coevolutionary residue pairs make close tertiary contacts within domains, which can suggest the information on intradomain couplings. We estimated the degree of coevolution between residues with their MI mentioned in the **Methods and Materials** section. Table S2 gives the 10 pairs of residues with the highest MI values. Among them, there are seven pairs of residues, each located in the same secondary structure element. For the remaining three pairs, Cys11-Cys77 (ranked first) forms a specific disulfide bond connecting helix1 and  $\beta$ 1 in group I.<sup>53</sup> For pairs Ala102-Phe106 (third) and Ala102-Phe5 (fifth) (Figure 2b), the involved three residues are all nearby the catalytic sites with Ala102 and Phe106 located at helix4 and Phe5 at helix1 (Figure 3b). Spatially, they are in close proximity to each other and have positively correlated movements (Figure 2a). Researchers have found that residues Phe5, Ala102, and Phe106 participate in the formation of the highly conserved hydrophobic substrate binding channel, facilitating the phospholipid binding to catalytic sites.<sup>36,54</sup>

**PRS Analyses of Family Members.** The PRS approach combines the ENM with LRT to assess the allosteric influence on each protein residue upon an external force exerted on the protein. We calculated the average PRS heat map over 130 sPLA<sub>2</sub> family members (see the **Methods and Materials** section) with the result presented in Figure 4a. The average over all elements in the  $j$ th column provides a measure of the ability of residue  $j$  to serve as a sensor, as shown in the bar plot along the higher abscissa with the peaks indicating the residues having the high potential to serve as sensors.

The PRS map describes the propensities of residues to sense perturbations and thus elicit cooperative responses, such as an

allosteric conformational change induced upon ligand binding to a highly “sensitive” sensor.<sup>34</sup> In order to detect the variations of residue sensitivities among family members, Figure 4b gives the standard deviation of PRS maps. From Figure 4a,b, many residues that show high signals in panel a also exhibit peaks in panel b, suggesting that the sites distinguished by their strong roles in allosteric communication have member-specific roles. Interestingly, the strongest sensors are consistent well with the identified four regions, according to the signature profile. Region 1 responsible for Ca<sup>2+</sup> binding is necessary for the enzyme’s catalytic activity. Regions 2 and 4 are involved in the recognition with the membrane which acts as an allosteric ligand. That is to say, these regions play a key role in the enzyme allostery. In addition, their member specificity indicated by the high standard deviations is consistent well with their differentiated functions across family members mentioned in the section of specific dynamics of sPLA<sub>2</sub> family members. In summary, PRS is an effective approach to identify the key sensor residues that function in protein allosteric processes.

Also from Figure 4a,b, it can be seen that many residues with low signals in panel a also exhibit dips in panel b. These residues with minimal sensitivity are Phe5; Met8 in helix1; Asp42 in helix3; Ile82 in  $\beta$ 2; and Cys96, Arg100, Ala103, and Phe106 in helix4 (see Figure 4c for the residue sensitivity profile), which are also characterized by the lowest variances, hinting that these insusceptible residues are shared among members. They are tightly packed in the structure (Figure S5) and occupy maxima in the average fluctuation profile over 130 sPLA<sub>2</sub> family members obtained by the GNM fastest 3 motion modes (94–96) (Figure 4c). The fast modes reflect the geometric irregularity in protein structure,<sup>55</sup> and the fluctuations associated with fast modes are accompanied by a decrease in entropy larger than that for slow modes.<sup>56</sup> Our previous studies show that the residues acting in the fastest modes are often associated with the protein fold core playing an important role in structural stability.<sup>39,56</sup> Additionally, researchers found that besides the lowest modes, the fastest modes are also highly conserved among members, which is required for chemical precision or structural stability of proteins.<sup>57</sup> In addition, note that they also have a high average degree (Figure 4d) over family member structure networks, reflecting their local high connectivity attitude within the networks. Node degree is also an index for indicating a residue’s important role in structure stability.<sup>58</sup> Thus, these residues’ tight packing (high connectivity) in structure and high activeness in the fastest modes indicate their important role in stabilizing the structure. In summary, the residues with high sensitivity (high mobility) are member-specific, which play a key role as sensors in sPLA<sub>2</sub> enzyme allostery, while the residues with low sensitivity (fold core) are the shared characteristics, which contribute largely to sPLA<sub>2</sub> enzyme structural stability.

Finally, for the dynamics methods GNM, ANM, and PRS used in this work, they are all effective for providing protein dynamical information, but they have different suitability or focus on different aspects.<sup>34,59</sup> GNM and ANM have the advantages in analyzing the large-amplitude collective motions of biomacromolecules in their native states, both of which can reproduce the molecular local structural flexibility and functionally motional correlations between residues, which are important for understanding molecular thermal motions and protein interdomain movement couplings, respectively.

Comparatively speaking, GNM is more suitable to analyze the molecular flexibility,<sup>59</sup> and ANM, due to its considering the direction information of residue movements, is of better ability in reproducing residue motional correlations.<sup>60</sup> PRS is specifically designed to quantify the residues' responses to an external perturbation on a protein residue, which can be used to find potential allosteric sites.<sup>34</sup> By utilizing the combination of these approaches, the dynamical properties of sPLA<sub>2</sub> family members can be analyzed more comprehensively.

## CONCLUSIONS

For the secretory phospholipase A<sub>2</sub> (sPLA<sub>2</sub>) family, playing a central role in the cellular lipid metabolism and signaling, we explore the shared and differentiated mechanisms of functions for sPLA<sub>2</sub> family members using a series of structure-based models and bioinformatics tools.

With PCA, we obtain the distribution of sPLA<sub>2</sub> family members along the first two PCs which also shed light on the functional motion of the enzyme. Five sPLA<sub>2</sub> groups (IA, IB, IIA, IIE, and X) are accurately separated by PC1 and PC2, where groups IA and IB and groups IIA and IIE have an evidently higher similarity in structure than other pairs of groups with the former higher than the latter, and an obvious structural difference exists between group IB members from bovine and porcine, which can be well traced back to their sequence similarities. The porcupine figures describing the two PCs indicate a shared open–close motion of the enzyme, which helps widen the hydrophobic phospholipid binding channel, consistent with the experimental observation.

The ANM analysis validates again the functional open–close movement where particularly the core helical regions (helix3 and helix4) bearing the catalytic role are dynamically strongly coupled, ensuring the stable exertion of catalytic functions. Additionally, the hydrophobic triad (Phe5-Ala102-Phe106), found by our coevolutionary analysis, has a strongly positively correlated movement, facilitating the formation of the highly conserved hydrophobic substrate binding channel.

Utilizing GNM, we explore the functionally shared and specific fluctuations among sPLA<sub>2</sub> family members based on the slowest motion modes and LTIF modes, respectively. The results show that for the shared important residues for the structure and function, they are highly conserved and consistently dynamically strongly constrained, which is evolutionarily essential as their perturbation would impact the function. These residues include catalytic residues (His48, Asp49, Tyr52, and Asp99), cysteine residues (Cys29, Cys45, Cys84, and Cys96 involved in the formation of disulfide bridges), and aromatic residues Tyr28 and Tyr73, which are critical for enzymatic hydrolysis of phospholipid or for structural stability. Conversely, those residue regions that present high mobility and large variations in LTIF modes are member-specific. These features presumably underlie the specificity of family members and determine their specific binding of ligands. The identified such four regions in the sPLA<sub>2</sub> family are involved in Ca<sup>2+</sup> binding necessary for the enzyme's catalytic activity and the recognition and binding with the membrane which acts as an allosteric ligand, respectively.

The PRS analysis reveals that the highly sensitive residues (sensors) to an external perturbation signal are member-specific, which in the sPLA<sub>2</sub> family just correspond to the four regions mentioned above, suggesting that they play different roles in allosteric communication as highly sensitive sensors.

Conversely, the minimal sensitive residues are the shared characteristics across sPLA<sub>2</sub> family members, which are highly active in the fastest modes and have high connectivity in PSN, revealing their large contribution to sPLA<sub>2</sub> enzyme structural stability likely as the folding core.

This study sheds light on the functional dynamics underlying the shared and differentiated functions of sPLA<sub>2</sub> family members as well as the key residue sites that enable the family's adaptation to the common catalytic reaction and specific substrate/ligand binding and allosteric activity. The results can help assist in the design, evaluation, and alterations of the specific functionalities of structural homologues.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.1c01315>.

Distributions of SID and RMSDs; heat maps of SID and RMSDs; sequence alignment; conservation of the sPLA<sub>2</sub> family; residues with minimal sensitivity; 130 sPLA<sub>2</sub> enzyme members used in this study; and residues with MI ([PDF](#))

## AUTHOR INFORMATION

### Corresponding Author

Chunhua Li – Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing 100124, China;  
Email: [chunhuali@bjut.edu.cn](mailto:chunhuali@bjut.edu.cn)

### Authors

Shan Zhang – Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing 100124, China;  
 [orcid.org/0000-0003-1283-8647](https://orcid.org/0000-0003-1283-8647)

Weikang Gong – Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing 100124, China;  
 [orcid.org/0000-0001-8797-784X](https://orcid.org/0000-0001-8797-784X)

Zhongjie Han – Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing 100124, China

Yang Liu – Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing 100124, China

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acs.jpcb.1c01315>

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (31971180 and 11474013).

## REFERENCES

- (1) Dennis, E. A. Introduction to Thematic Review Series: Phospholipases: Central Role in Lipid Signaling and Disease. *J. Lipid Res.* **2015**, *56*, 1245–1247.
- (2) Quach, N. D.; Arnold, R. D.; Cummings, B. S. Secretory phospholipase A2 enzymes as pharmacological targets for treatment of disease. *Biochem. Pharmacol.* **2014**, *90*, 338–348.
- (3) Filkin, S. Y.; Lipkin, A. V.; Fedorov, A. N. Phospholipase Superfamily: Structure, Functions, and Biotechnological Applications. *Biochemistry* **2020**, *85*, 177–195.
- (4) Mouchlis, V. D.; Bucher, D.; Mccammon, J. A.; Dennis, E. A. Membranes serve as allosteric activators of phospholipase A2,

- enabling it to extract, bind, and hydrolyze phospholipid substrates. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, E516–E525.
- (5) Murakami, M.; Sato, H.; Miki, Y.; Yamamoto, K.; Taketomi, Y. A new era of secreted phospholipase A2. *J. Lipid Res.* **2015**, *56*, 1248–1261.
- (6) Huang, W.-N.; Chen, Y.-H.; Chen, C.-L.; Wu, W. Surface pressure-dependent interactions of secretory phospholipase A2 with zwitterionic phospholipid membranes. *Langmuir* **2011**, *27*, 7034–7041.
- (7) Burke, J. E.; Karbarz, M. J.; Deems, R. A.; Li, S.; Woods, V. L., Jr.; Dennis, E. A. Interaction of group IA phospholipase A2 with metal ions and phospholipid vesicles probed with deuterium exchange mass spectrometry. *Biochemistry* **2008**, *47*, 6451–6459.
- (8) Karplus, M.; McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652.
- (9) Manukyan, A. K. Structural aspects and activation mechanism of human secretory group IIA phospholipase. *Eur. Biophys. J.* **2020**, *49*, 511–531.
- (10) Ramakrishnan, C.; Subramanian, V.; Velmurugan, D. Molecular dynamics study of secretory phospholipase A2 of Russell's viper and bovine pancreatic sources. *J. Phys. Chem. B* **2010**, *114*, 13463–13472.
- (11) Chinnasamy, S.; Selvaraj, G.; Selvaraj, C.; Kaushik, A. C.; Kaliamurthi, S.; Khan, A.; Singh, S. K.; Wei, D.-Q. Combining *in silico* and *in vitro* approaches to identification of potent inhibitor against phospholipase A2 (PLA2). *Int. J. Biol. Macromol.* **2020**, *144*, 53–66.
- (12) Wang, Y.; Rader, A. J.; Bahar, I.; Jernigan, R. L. Global ribosome motions revealed with elastic network model. *J. Struct. Biol.* **2004**, *147*, 302–314.
- (13) Bahar, I.; Atilgan, A. R.; Erman, B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Folding Des.* **1997**, *2*, 173–181.
- (14) Atilgan, A. R.; Durell, S. R.; Jernigan, R. L.; Demirel, M. C.; Keskin, O.; Bahar, I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **2001**, *80*, 505–515.
- (15) Kim, M. K.; Chirikjian, G. S.; Jernigan, R. L. Elastic models of conformational transitions in macromolecules. *J. Mol. Graphics Modell.* **2002**, *21*, 151–160.
- (16) Gong, W.; Liu, Y.; Zhao, Y.; Wang, S.; Li, C. Equally Weighted Multiscale Elastic Network Model and Its Comparison with Traditional and Parameter-Free Models. *J. Chem. Inf. Model.* **2021**, *61*, 921.
- (17) Wang, S.; Gong, W.; Deng, X.; Liu, Y.; Li, C. Exploring the dynamics of RNA molecules with multiscale Gaussian network model. *Chem. Phys.* **2020**, *538*, 110820.
- (18) Li, C.; Lv, D.; Zhang, L.; Yang, F.; Wang, C.; Su, J.; Zhang, Y. Approach to the unfolding and folding dynamics of add A-riboswitch upon adenine dissociation using a coarse-grained elastic network model. *J. Chem. Phys.* **2016**, *145*, 014104.
- (19) Xie, X. L.; Li, C. H.; Yang, Y. X.; Jin, L.; Tan, J. J.; Zhang, X. Y.; Su, J. G.; Wang, C. X. Allosteric transitions of ATP-binding cassette transporter MsbA studied by the adaptive anisotropic network model. *Proteins* **2015**, *83*, 1643–1653.
- (20) Han, Z.; Shao, Q.; Gong, W.; Wang, S.; Su, J.; Li, C.; Zhang, Y. Interpreting the Dynamics of Binding Interactions of snRNA and U1A Using a Coarse-Grained Model. *Biophys. J.* **2019**, *116*, 1625–1636.
- (21) Atilgan, C.; Atilgan, A. R. Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. *PLoS Comput. Biol.* **2009**, *5*, No. e1000544.
- (22) Di Paola, L.; Hadi-Alijanvand, H.; Song, X.; Hu, G.; Giuliani, A. The Discovery of a Putative Allosteric Site in the SARS-CoV-2 Spike Protein Using an Integrated Structural/Dynamic Approach. *J. Proteome Res.* **2020**, *19*, 4576–4586.
- (23) Liang, Z.; Zhu, Y.; Liu, X.; Hu, G. Role of protein-protein interactions in allosteric drug design for DNA methyltransferases. *Adv. Protein Chem. Struct. Biol.* **2020**, *121*, 49–84.
- (24) Mikulska-Ruminska, K.; Shrivastava, I.; Krieger, J.; Zhang, S.; Li, H.; Bayir, H.; Wenzel, S. E.; VanDemark, A. P.; Kagan, V. E.; Bahar, I. Characterization of Differential Dynamics, Specificity, and Allostery of Lipoxygenase Family Members. *J. Chem. Inf. Model.* **2019**, *59*, 2496–2508.
- (25) Tobi, D. Dynamics based clustering of globin family members. *PLoS One* **2018**, *13*, No. e0208465.
- (26) Wingert, B.; Krieger, J.; Li, H.; Bahar, I. Adaptability and specificity: how do proteins balance opposing needs to achieve function? *Curr. Opin. Struct. Biol.* **2021**, *67*, 25–32.
- (27) Greene, L. H. Protein structure networks. *Briefings Funct. Genomics* **2012**, *11*, 469–478.
- (28) Vanwart, A. T.; Eargle, J.; Luthey-Schulten, Z.; Amaro, R. E. Exploring residue component contributions to dynamical network models of allostery. *J. Chem. Theory Comput.* **2012**, *8*, 2949–2961.
- (29) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (30) Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **1933**, *24*, 417–441.
- (31) Grant, B. J.; Rodrigues, A. P. C.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. D. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **2006**, *22*, 2695–2696.
- (32) Holm, L.; Rosenstrom, P. DALI server: conservation mapping in 3D. *Nucleic Acids Res.* **2010**, *38*, W545–W549.
- (33) Ikeguchi, M.; Ueno, J.; Sato, M.; Kidera, A. Protein structural change upon ligand binding: Linear response theory. *Phys. Rev. Lett.* **2005**, *94*, 078102.
- (34) Dutta, A.; Krieger, J.; Lee, J. Y.; Garcia-Nafria, J.; Greger, I. H.; Bahar, I. Cooperative Dynamics of Intact AMPA and NMDA Glutamate Receptors: Similarities and Subfamily-Specific Differences. *Structure* **2015**, *23*, 1692–1704.
- (35) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410.
- (36) Oliveira, A.; Bleicher, L.; Schrago, C. G.; Silva, F. P., Jr. Conservation analysis and decomposition of residue correlation networks in the phospholipase A2 superfamily (PLA2s): Insights into the structure-function relationships of snake venom toxins. *Toxicon* **2018**, *146*, 50–60.
- (37) Thompson, J. D.; Gibson, T. J.; Higgins, D. G. Multiple Sequence Alignment Using ClustalW and ClustalX. *Curr. Protoc. Bioinf.* **2003**, *00*, 2.3.1–2.3.22.
- (38) Simonetti, F. L.; Elin, T.; Ariel, C.; Morten, N.; Cristina, M. B. MISTIC: mutual information server to infer coevolution. *Nucleic Acids Res.* **2013**, *41*, W8.
- (39) Yang, L.; Song, G.; Carriquiry, A.; Jernigan, R. L. Close correspondence between the motions from principal component analysis of multiple HIV-1 protease structures and elastic network modes. *Structure* **2008**, *16*, 321–330.
- (40) Cao, J.; Burke, J. E.; Dennis, E. A.; Burke, E. A.; Chemistry, D. J. J. o. B. Using Hydrogen/Deuterium Exchange Mass Spectrometry to Define the Specific Interactions of the Phospholipase A2 Superfamily with Lipid Substrates, Inhibitors, and Membranes. *J. Biol. Chem.* **2013**, *288*, 1806–13.
- (41) Singh, N.; Somvanshi, R.; Sharma, S.; Dey, S.; Kaur, P.; Singh, T. Structural Elements of Ligand Recognition Site in Secretory Phospholipase A2 and Structure-Based Design of Specific Inhibitors. *Curr. Top. Med. Chem.* **2007**, *7*, 757–764.
- (42) Gur, M.; Zomot, E.; Bahar, I. Global motions exhibited by proteins in micro- to milliseconds simulations concur with anisotropic network model predictions. *J. Chem. Phys.* **2013**, *139*, 121912.
- (43) Yang, L.-W.; Bahar, I. Coupling between catalytic site and collective dynamics: A requirement for mechanochemical activity of enzymes. *Structure* **2005**, *13*, 893–904.
- (44) Dupreux, C. M.; Yu, B. Z.; Jain, M. K.; Noel, J. P.; Deng, T.; Li, Y.; Byeon, I. J. L.; Tsai, M. D. Phospholipase A2 engineering: Structural and functional roles of highly conserved active site residues tyrosine-52 and tyrosine-73. *Biochemistry* **1992**, *31*, 6402–6413.
- (45) Huang, T.-T.; Marcos, M. d. V.; Hwang, J.-K.; Echave, J. A mechanistic stress model of protein evolution accounts for site-specific evolutionary rates and their relationship with packing density and flexibility. *BMC Evol. Biol.* **2014**, *14*, 78.

- (46) Tiwari, S. P.; Reuter, N. Conservation of intrinsic dynamics in proteins—what have computational models taught us? *Curr. Opin. Struct. Biol.* **2018**, *50*, 75–81.
- (47) Davidson, F. F.; Dennis, E. A. Evolutionary relationships and implications for the regulation of phospholipase A2 from snake venom to human secreted forms. *J. Mol. Evol.* **1990**, *31*, 228–238.
- (48) Six, D. A.; Dennis, E. A. The expanding superfamily of phospholipase A2 enzymes: classification and characterization. *Biochim. Biophys. Acta* **2000**, *1488*, 1–19.
- (49) Bezzine, S.; Bollinger, J. G.; Singer, A. G.; Veatch, S. L.; Keller, S. L.; Gelb, M. H. On the binding preference of human groups IIA and X phospholipases A2 for membranes with anionic phospholipids. *J. Biol. Chem.* **2002**, *277*, 48523–48534.
- (50) Hou, S.; Xu, T.; Xu, J.; Qu, L.; Xu, Y.; Chen, L.; Liu, J. Structural basis for functional selectivity and ligand recognition revealed by crystal structures of human secreted phospholipase A2 group IIE. *Sci. Rep.* **2017**, *7*, 10815.
- (51) Liu, Y.; Bahar, I. Sequence evolution correlates with structural dynamics. *Mol. Biol. Evol.* **2012**, *29*, 2253–2263.
- (52) Wako, H.; Endo, S. Dynamic properties of oligomers that characterize low-frequency normal modes. *Biophys. Physicobiol.* **2019**, *16*, 220–231.
- (53) Arni, R. K.; Ward, R. J. Phospholipase A2—a structural review. *Toxicon* **1996**, *34*, 827–841.
- (54) Ward, R. J.; Alves, A. R.; Ruggiero Neto, J.; Arni, R. K.; Casari, G. A SequenceSpace analysis of Lys49 phospholipases A2: clues towards identification of residues involved in a novel mechanism of membrane damage and in myotoxicity. *Protein Eng.* **1998**, *11*, 285–294.
- (55) Haliloglu, T.; Keskin, O.; Ma, B.; Nussinov, R. How Similar Are Protein Folding and Protein Binding Nuclei? Examination of Vibrational Motions of Energy Hot Spots and Conserved Residues. *Biophys. J.* **2005**, *88*, 1552–1559.
- (56) Su, J. G.; Jiao, X.; Sun, T. G.; Li, C. H.; Chen, W. Z.; Wang, C. X. Analysis of domain movements in glutamine-binding protein with simple models. *Biophys. J.* **2007**, *92*, 1326–1335.
- (57) Bahar, I.; Attilgan, A. R.; Demirel, M. C.; Erman, B. Vibrational dynamics of folded proteins: Significance of slow and fast motions in relation to function and stability. *Phys. Rev. Lett.* **1998**, *80*, 2733–2736.
- (58) Vishveshwara, S.; Ghosh, A.; Hansia, P. Intra and Inter-Molecular Communications Through Protein Structure Network. *Curr. Protein Pept. Sci.* **2009**, *10*, 146–160.
- (59) Tekpinar, M.; Yildirim, A. Only a Subset of Normal Modes is Sufficient to Identify Linear Correlations in Proteins. *J. Chem. Inf. Model.* **2018**, *58*, 1947–1961.
- (60) Bouvignies, G.; Bernadó, P.; Meier, S.; Cho, K.; Grzesiek, S.; Bruschweiler, R.; Blackledge, M. Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13885–13890.