Molecular BioSystems

PAPER

View Article Online
View Journal | View Issue

Cite this: *Mol. BioSyst.*, 2014, **10**, 2597

Features of S-nitrosylation based on statistical analysis and molecular dynamics simulation: cysteine acidity, surrounding basicity, steric hindrance and local flexibility†

Shangli Cheng,[a] Ting Shi,[a] Xiao-Lei Wang,[a] Juan Liang,[a] Hongyi Wu,[a] Lu Xie,[b] Yixue Li[ab] and Yi-Lei Zhao*[ab]

Received 30th May 2014,
Accepted 2nd July 2014

DOI: 10.1039/c4mb00322e

www.rsc.org/molecularbiosystems

S-Nitrosylation is involved in protein functional regulation and cellular signal transduction. Although intensive efforts have been made, the molecular mechanisms of S-nitrosylation have not yet been fully understood. In this work, we carried out a survey on 213 protein structures with S-nitrosylated cysteine sites and molecular dynamic simulations of hemoglobin as a case study. It was observed that the S-nitrosylated cysteines showed a lower pK_a, a higher population of basic residues, a lower population of big-volume residues in the neighborhood, and relatively higher flexibility. The case study of hemoglobin showed that, compared to that in the T-state, Cysβ93 in the R-state hemoglobin possessed the above structural features, in agreement with the previous report that the R-state was more reactive in S-nitrosylation. Moreover, basic residues moved closer to the Cysβ93 in the dep-R-state hemoglobin, while big-volume residues approached the Cysβ93 in the dep-T-state. Using the four characteristics, *i.e.* cysteine acidity, surrounding basicity, steric hindrance, and local flexibility, a 3-dimensional model of S-nitrosylation was constructed to explain 61.9% of the S-nitrosylated and 58.1% of the non-S-nitrosylated cysteines. Our study suggests that cysteine deprotonation is a prerequisite for protein S-nitrosylation, and these characteristics might be useful in identifying specificity of protein S-nitrosylation.

Introduction

Thiol group of cysteine is modified toward S-nitrosothiol in protein S-nitrosylation, and the process is reversible.[1,2] Accumulating evidences suggest that S-nitrosylation plays a key role in regulation of protein functions,[3] human health and diseases[4,5] as well as cellular signaling.[6,7] In particular, protein S-nitrosylation is the molecular basis of NO-related cellular signal transduction.[8,9] Many human diseases, such as Parkinson's disease,[10] neurodegeneration[11] and cancer,[8] and even some physiological processes in plant[12] are related to S-nitrosylation. Biological methods and proteomic experiments are employed to identify the S-nitrosylated cysteine sites in proteins.[13–16]

It is reported that S-nitrosylation is highly specific and selective;[6,17] however, the mechanism of protein S-nitrosylation is still unclear.[18] Many possible pathways have been reported, such as (1) NO-dependent S-nitrosylation,[18,19] (2) *trans*-S-nitrosation,[20] and (3) Cu^{2+}-induced S-nitrosylation in presence of NO.[21] It is likely that deprotonation of thiol group of cysteine is involved in the pathways.[9,18] By using sequence-based bioinformatical method, acid-based motif[22–24] and a revised acid-based motif[25] were proposed. Moreover, structure-based analyses were reported as well.[25,26] For example, based on a few protein structures, S-nitrosylated cysteines were reported to have higher predicted pK_a values; moreover, the modified sites were reported to be located in highly exposed areas of protein.[26] Up to now, few significant characteristics and rare systematic investigations of 3-dimensional structures have been reported to explain the selectivity of S-nitrosylation.

A special case of S-nitrosylation is hemoglobin, in which the S-nitrosylation is preferentially formed on Cysβ93 in R-state hemoglobin rather than T-state.[8] Conformational transition between R- and T-state hemoglobin, which is caused by oxygenation and deoxygenation, led to S-nitrosylation and de-nitrosylation, respectively.[27–29] Cysβ93 in R-state hemoglobin is more reactive for S-nitrosylation compared to the T-state.[30] For the physiological significance, S-nitrosylation of hemoglobin can affect the response of hypoxic vasodilation in human respiratory cycle.[28] Therefore, hemoglobin is selected as a model protein

[a] *State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China. E-mail: yileizhao@sjtu.edu.cn; Fax: +86-21-34207190; Tel: +86-21-34207190*
[b] *Shanghai Center for Bioinformation Technology, Shanghai 201203, China*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c4mb00322e

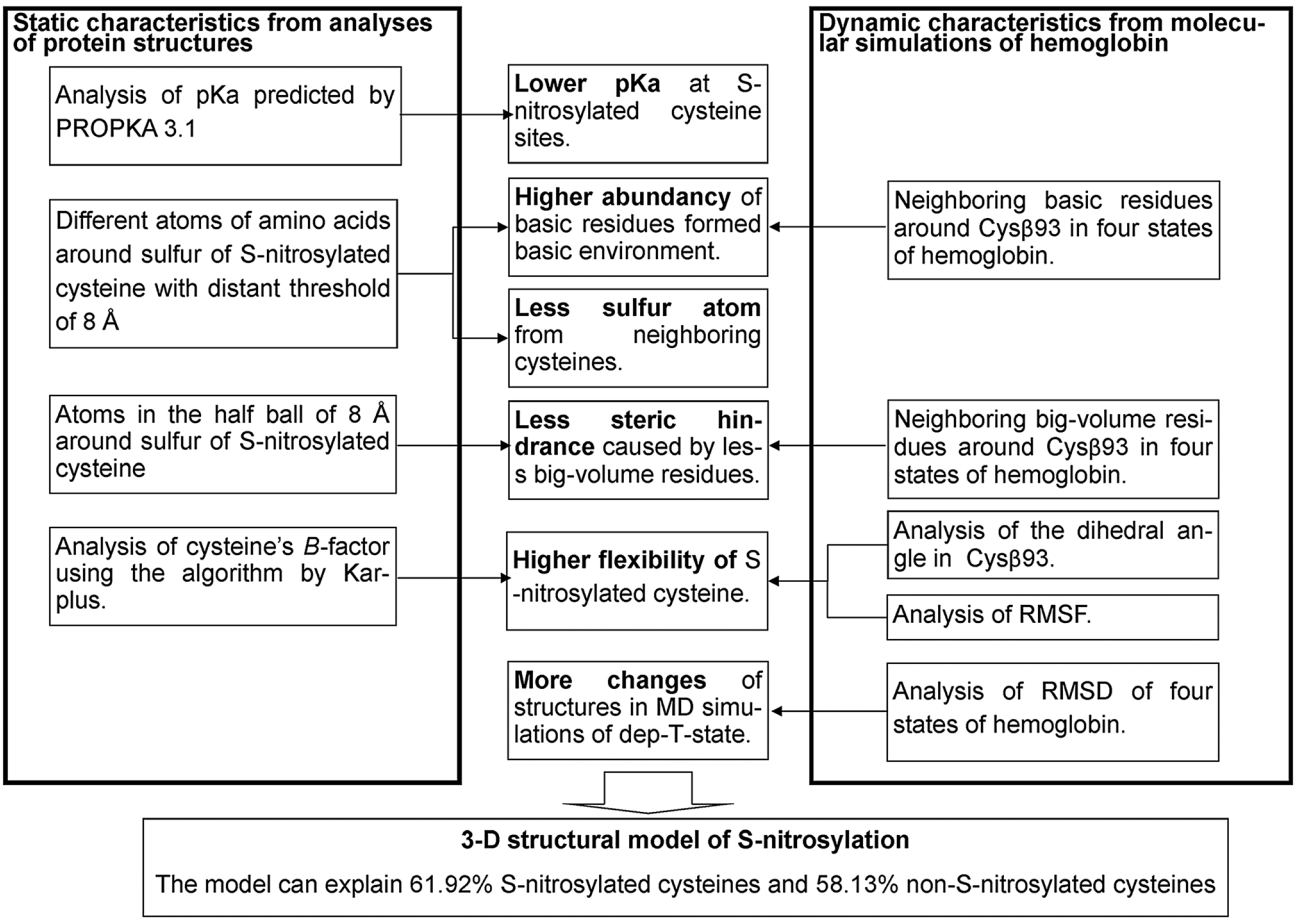

Fig. 1 The flow chart shows correlations between S-nitrosylation and characteristics.

to analyze the process of S-nitrosylation affected by structural changes rather than sequence differences.

In our work, we collected 213 structures of S-nitrosylated proteins from PDB database by BLAST tool (Standard Protein BLAST in webserver of NCBI). Structure-based investigations on the S-nitrosylated proteins were carried out, including pK_a, atomic distribution, steric hindrance and local flexibility. Since the process of S-nitrosylation was related to deprotonation of cysteine, the hemoglobin with deprotonated Cysβ93 was studied in this work. Overall, four states of hemoglobin, including R-state, T-state, dep-R-state (R-state with deprotonated Cysβ93) and dep-T-state (T-state with deprotonated Cysβ93) were selected to detect dynamic characteristics. According to the discovered characteristics (Fig. 1), a three-dimensional structure-based S-nitrosylation model was constructed, explaining 61.92% of the S-nitrosylated and 58.13% of the non-S-nitrosylated cysteine sites in the collected proteins.

Materials and methods

Collection of S-nitrosylated proteins

The S-nitrosylated proteins studied in this work were collected in sequence from the previously reported GPS-SNO paper[31] and references therein (Table S8, ESI†). The structural information of the S-nitrosylated proteins was then obtained by the BLAST tool and the PDB database. BLAST thresholds of identity and positivity were set to be greater than 0.95. Overall, 213 proteins containing 323 S-nitrosylated cysteine sites and 965 non-S-nitrosylated cysteine sites were obtained. In the redundancy analysis using CD-HIT,[32] sequence similarity of 179 S-nitrosylated proteins was less than 0.7. CD-HIT was used to cluster and compare protein sequences by similarity tolerance.

Analysis of pK_a

Acidity constant pK_a can be the quantitative measurement of dissociation of thiol group (–SH). Because some PDB files contained non-standard amino acids, they could not be correctly recognized by the program of PROPKA 3.1,[33] in which pK_a value was estimated according to coulomb interactions and the description of internal and surface residues. In our study, pK_a values of 276 out of S-nitrosylated cysteines and 685 out of non-S-nitrosylated cysteines were calculated based on protein structures by the program.

Analysis of neighboring atoms

Atoms of neighboring residues were analyzed within a series of distance thresholds, including 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5 and 8.0 Å, in which distance was from the sulfur of all cysteine sites to the neighboring atoms. The 20 types of amino

acids were grouped into five categories, *i.e.*, polar amino acid (Ser, Thr, Cys, Pro, Asn, and Gln), acidic amino acid (Asp and Glu), basic amino acid (Lys, Arg, and His), aromatic amino acid (Phe, Tyr and Trp) and aliphatic amino acid (Gly, Ala, Val, Leu, Ile and Met).[34] Mann–Whitney U test was used for the distributional analysis of the different types of atoms. The percentage difference value was defined in formula 1.

$$\text{Percentage difference value} = \frac{PN(\text{i}) - NN(\text{i}) \times \frac{P}{N}}{P} \times 100\% \quad (1)$$

where *PN*(i), *NN*(i), *P* and *N* represent the number of atoms(i) in the S-nitrosylated set, the number of atoms(i) in the non-S-nitrosylated set, the number of S-nitrosylated cysteines and the number of non-S-nitrosylated cysteines, respectively. If the percentage difference value was greater than 0, the type of atoms was highly abundant in the S-nitrosylated cysteine set and *vice versa*.

Analysis of steric hindrance

To analyze steric hindrance, atoms (X) located in front of the cysteine residues were selected with the distance to the sulfur atom of cysteine less than 8 Å, and the angle of C–S···X larger than 90°, where the C atom was the side-chain carbon, and the X atom was located in the half ball of 8 Å (Fig. 4). The X atom could cause steric hindrance in the process of S-nitrosylation, which might prevent oxidant agents from attacking the thiol group.

Analysis of flexibility

The *B*-factor reflects local structural fluctuations. In order to evaluate local flexibility of S-nitrosylated or non-S-nitrosylated cysteine, the *B*-factor (*B*-value) of cysteine was calculated using Karplus algorithm based on protein sequences.[35]

Hemoglobin and molecular dynamics simulations

Since the R- and T-state of hemoglobin shared the identical amino acid sequence but with different capacity for S-nitrosylation,[36] interactions between neighboring amino acid residues and Cysβ93 site were investigated by molecular dynamics (MD) simulations on four types of hemoglobins: the R-state, T-state, dep-R-state and dep-T-state hemoglobins. The initial structures for R- and T-states of hemoglobin were chosen from 1HHO[37] and 2HHB[38] in the PDB database. The full α2β2 structure of R-states was constructed with α and β subunits (1HHO) based on symmetry.[39]

In the preparation of four hemoglobin structures, histidine residues connecting to the heme were protonated at the δ-position, while the other histidine residues were protonated at the ε-position. Hemoglobin molecules were immersed in octahedral boxes of TIP3P water with 10 Å to the edge. Therefore, the systems of R-state, T-state, dep-R-state and dep-T-state contained 8870, 8956, 8864 and 8955 water molecules, respectively. Six sodium ions in the cases of R-state and T-state[40] and eight sodium ions in the cases of dep-R-state and dep-T-state were added into the water boxes for charge neutralization. In the MD simulations, the AMBER force field 99SB was used for all amino acids and heme motif.[41] The cut-off of 10.0 Å[40] and SHAKE algorithm[42,43] were used for short-range non-bonded and hydrogen bonds under periodic boundary condition. After the conjugate gradient method was performed under the minimization step, the system was gradually heated from 0 to 300 K. Finally, the 20 ns simulations

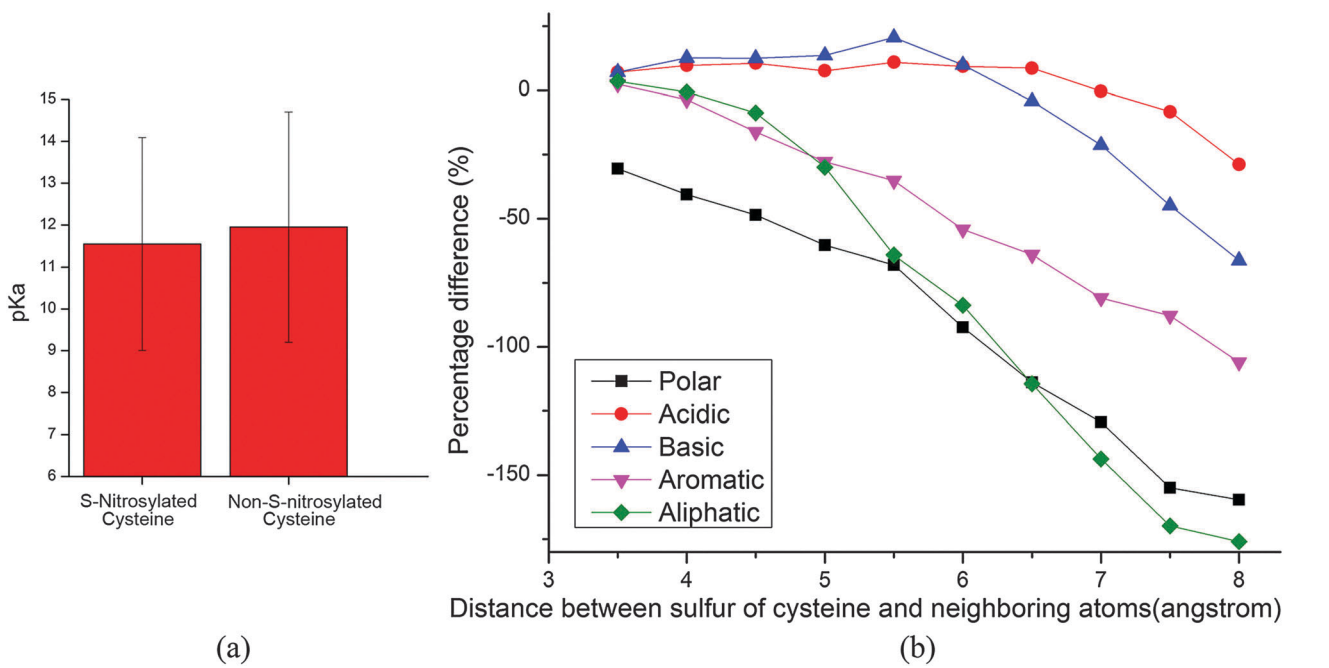

Fig. 2 pK_a of S-nitrosylated and non-S-nitrosylated cysteines and atomic distribution around sulfur of S-nitrosylated cysteines. (a) pK_a of S-nitrosylated and non-S-nitrosylated cysteines was calculated using PROPKA 3.1. *y*-axis represents the value of pK_a. (b) *x*-axis represents distance (3.5 to 8.0 Å) from sulfur of cysteine to the neighboring atoms. *y*-axis represents percentage difference values between the S-nitrosylated and non-S-nitrosylated cysteine sets. The Mann–Whitney U test was used (Tables S1–S5, ESI†).

of the R-, T-, dep-R- and dep-T-state hemoglobins were carried out in the absence of any restraint under conditions of 300 K, NTP and a time-step of 2 fs. Each simulation was repeated 14 times with rearranged random number. For each state of hemoglobin, 280 ns MD simulations were carried out. In total, 1120 ns MD trajectory were obtained. Typically, frames in a range of 5 to 20 ns in each trajectory were used for analysis.

Results and discussions

S-nitrosylated cysteine has lower pK_a

In our work, the pK_a values of thiol group in cysteine residue were calculated using PROPKA 3.1. The results showed that the pK_a value of the S-nitrosylated cysteine (11.54 ± 2.54, using 276 S-nitrosylated cysteine sites) was lower than that of the

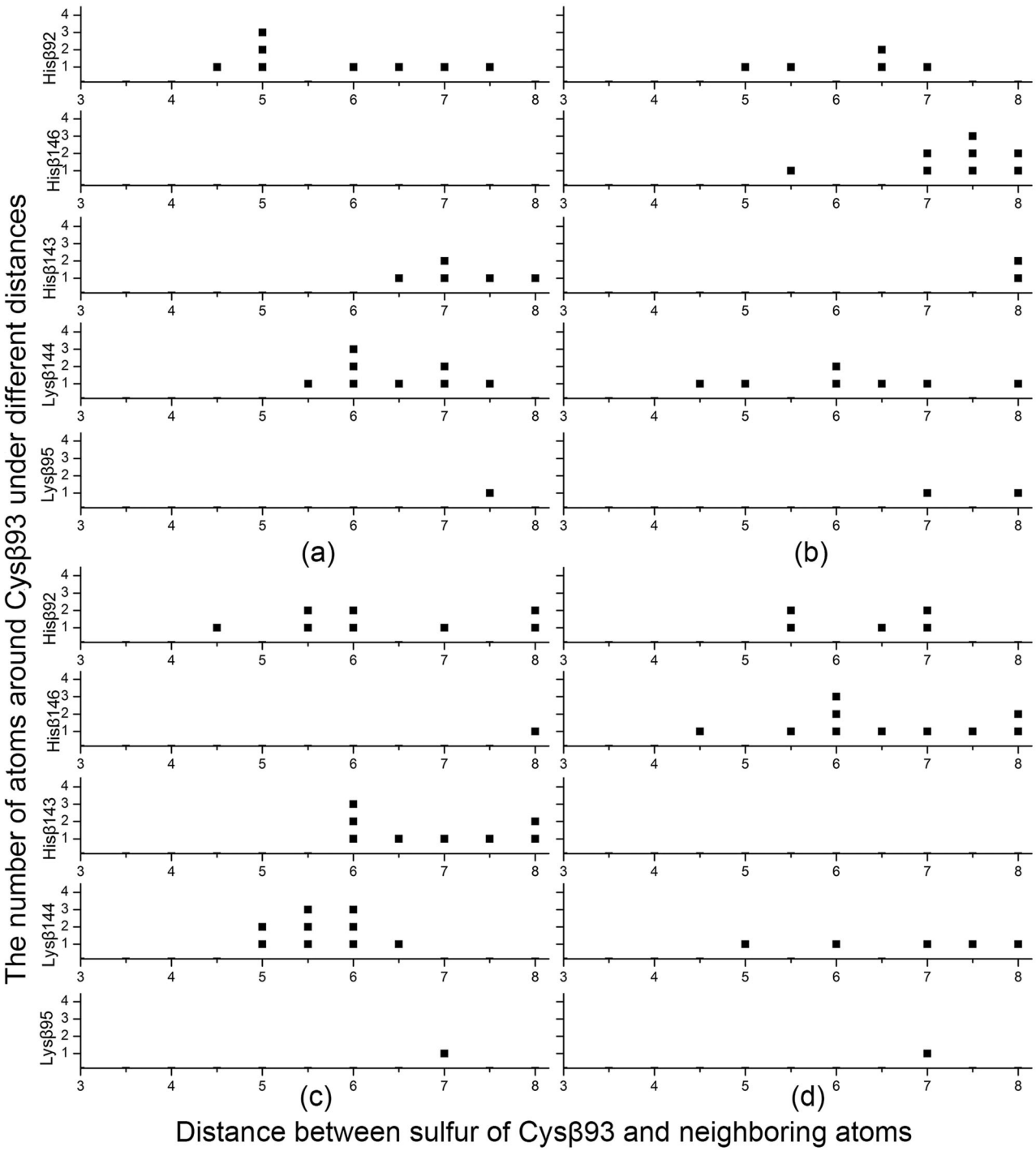

Fig. 3 The neighboring basic residues around Cysβ93 in R-state (a), T-state (b), dep-R-state (c) and dep-T-state (d) hemoglobin. *x*-axis represents the distance between the sulfur of Cysβ93 and neighboring atoms from different residues. *y*-axis (■) represents the number of atoms.

non-S-nitrosylated cysteine (11.95 ± 2.75, using 685 non-S-nitrosylated cysteine sites) (Fig. 2a). Compared to P.-T. Doulias' study,[26] our data set covered and expanded the S-nitrosylated cysteines in his work, where the pK_a of S-nitrosylated cysteines (10.0 ± 2.10, using 142 S-nitrosylated cysteine sites) was higher than non-S-nitrosylated cysteines' pK_a (9.88 ± 2.20, using 559 non-S-nitrosylated cysteine sites) using PROPKA 2.0. When the complete dataset was used with the version 2.0 (PROPKA 2.0), the pK_a values of S-nitrosylated (227 sites) and non-S-nitrosylated (638 sites) were calculated as 9.36 ± 3.24 and 9.89 ± 2.68, respectively. We speculated that a larger dataset and updated software might be responsible for the differences.

Since the average pK_a values of S-nitrosylated and non-S-nitrosylated cysteine were all larger than the physiological pH, the cysteines existed mostly in the protonated state. However, the lower pK_a of S-nitrosylated cysteines suggested that it would be relatively more feasible to deprotonate the thiol group (–SH) of S-nitrosylated cysteine than the non-S-nitrosylated cysteines. Moreover, in the case of hemoglobin, S-nitrosylation of both R- and T-state Cysβ93 was accelerated under a more basic environment in previous study.[30] Such a pK_a (>7.0) does not grant the formation of thiol anion (RS^-) yet; thus, a basic environment was necessary to enhance the process.

Basic residues around S-nitrosylated cysteines improve and stabilize deprotonated cysteines

To analyze the physicochemical environments around the sulfur of cysteine, we extracted the characteristics of atomic distribution of the 323 S-nitrosylated and 965 non-S-nitrosylated sites. According to five groups of amino acids, namely, polar (S, T, C, P, N and Q), acidic (D and E), basic (K, R and H), aromatic (F, Y and W) and aliphatic (G, A, V, L, I and M) groups, the percentage difference values (PDVs)[44] between S-nitrosylated and non-S-nitrosylated cysteines were calculated (Fig. 2b, Tables S6.1–S6.10, ESI†). PDVs of different types of atoms were tested using Mann–Whitney U test (Tables S1–S5, ESI†). Within a range of less than 6 Å around the sulfur of cysteines, basic and acidic residues were of a high abundance in S-nitrosylated cysteine sites. It was suggested that S-nitrosylated cysteines distributed in highly charged environments. This result was consistent with the previous study of –SNO group surrounded by charged residues;[45–47] *i.e.*, the basic residues were present more frequently around the S-nitrosylated cysteine sites. Among the basic residues, histidine was abundant in S-nitrosylated cysteines within 5.0 Å ($P < 0.05$, Tables S6.1–S6.4, ESI†). These results suggested that a basic environment could facilitate the process of cysteine deprotonation under physiological pH. Moreover, for aromatic and aliphatic groups, Phe ($P < 0.05$, Tables S6.3–S6.10, ESI†), Leu ($P < 0.05$, Tables S6.3–S6.10, ESI†) and Tyr (Tables S6.3–S6.10, ESI†) were of high abundance in non-S-nitrosylated cysteines in a distance range between 4.5 and 8 Å. These results demonstrated that a nonpolar environment would be disfavored for S-nitrosylation.

Correspondingly, hemoglobin MD simulations showed that more basic residues surrounded Cysβ93 in R- and dep-R-state hemoglobin. The average distance from sulfur of Cysβ93 to neighboring atoms was calculated by repeated trajectories of MD simulations. Neighboring basic residues, including Hisβ92, Hisβ143, Hisβ146, Lysβ95 and Lysβ144, were analyzed based on different radii around the sulfur atom of Cysβ93. When the distance was set at less than 6 Å, it was observed that basic residues, especially Hisβ92 and Lysβ144, presented more frequently around Cysβ93 in R-state than in T-state hemoglobin (Fig. 3a and b). When the Cysβ93 was deprotonated in dep-R-state, basic residues of Hisβ143 and Lysβ144 moved closer to the Cysβ93 (Fig. 3c), and additional hydrogen bonds were formed between N–H and thiol anion in certain frames (Fig. S1, ESI†), while only Hisβ146 approached the Cysβ93 in dep-T-state hemoglobin (Fig. 3d). Thus, the deprotonated Cysβ93 immersed in a more basic environment in the dep-R-state hemoglobin. This indicated that basic amino acid residues likely stabilized the deprotonated Cysβ93, which might be important in the process of S-nitrosylation.

Integrating the analysis of 323 S-nitrosylated cysteines (Fig. 2b) and MD simulations of hemoglobin (Fig. 3), we proposed that basic amino acids, especially histidine, were critical for S-nitrosylated cysteines. It was suggested that basic amino acids could form a basic environment that could contribute for the deprotonation of S-nitrosylated cysteines and stabilization of deprotonated cysteines. In the collected S-nitrosylated proteins, 49 S-nitrosylated proteins with 53 S-nitrosylated cysteines were in line with this characteristic (Table S7, ESI†).

Steric hindrance inhibits S-nitrosylation

In order to evaluate the steric hindrance caused by the neighboring atoms around the sulfur of S-nitrosylated and non-S-nitrosylated cysteines, we analyzed distributional characteristics of the neighboring atoms in the half ball of 8 Å (Fig. 4). The S-nitrosylated cysteines had fewer neighboring atoms than the non-S-nitrosylated cysteines, for which medians and interquartile ranges were estimated at 48 and 19 for the former and at 50 and 12 for the latter,

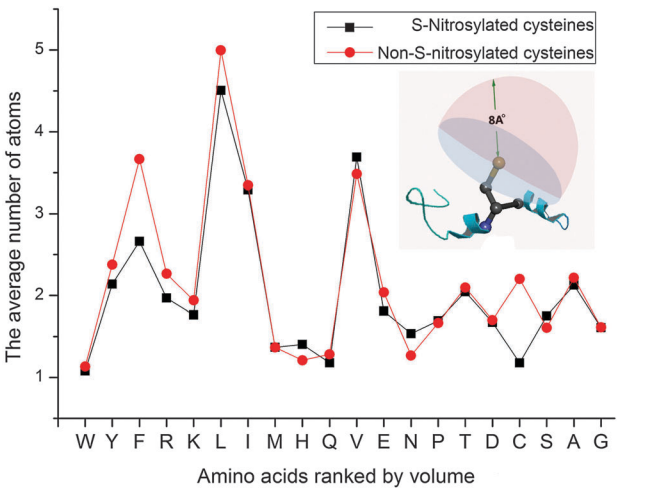

Fig. 4 Steric hindrance caused by atoms in half ball of 8 Å. *x*-axis represents the different types of amino acids ranked by volume (BIGC670101 in AAindex) from large to small. *y*-axis represents the average number of atoms in S-nitrosylated cysteine sites and non-S-nitrosylated cysteine sites.

respectively ($P < 0.01$, with Mann–Whitney U test. See Fig. S2, ESI†). When the atoms were attributed to amino acid residues and ranked by volume of amino acids, it was found that atoms of amino acids with big volume, such as Tyr, Phe, Arg, and Leu, were present more in non-S-nitrosylated cysteines than in S-nitrosylated cysteines (Fig. 4). The steric hindrance caused by these big-volume residues would affect the attacking of oxidant agents and the binding of NO group. Moreover, the previous study reported that there were more exposing S-nitrosylated cysteines than non-S-nitrosylated cysteines in protein structure.[25] Our results suggested that steric hindrance could be a disadvantage for the process of S-nitrosylation.

For comparison, similar statistical analyses were carried out on the hemoglobin trajectories of MD simulations. Neighboring aromatic and aliphatic residues with big volume, including Pheβ103, Tyrβ145, Leuβ91 and Leuβ141, were analyzed based on different cut-off radius around the sulfur of Cysβ93. For R-, T-, dep-R- and dep-T-state hemoglobin, MD simulations were repeated 14 times. Average distances were calculated from the sulfur of Cysβ93 to the neighboring atoms. The results showed that these big-volume residues were closer to the Cysβ93 in T-state hemoglobin than that in R-state hemoglobin, especially Tyrβ145 (Fig. 5a and b). When the Cysβ93 was deprotonated, the big-volume residues, including Pheβ103, Tyrβ145 and Leuβ141,

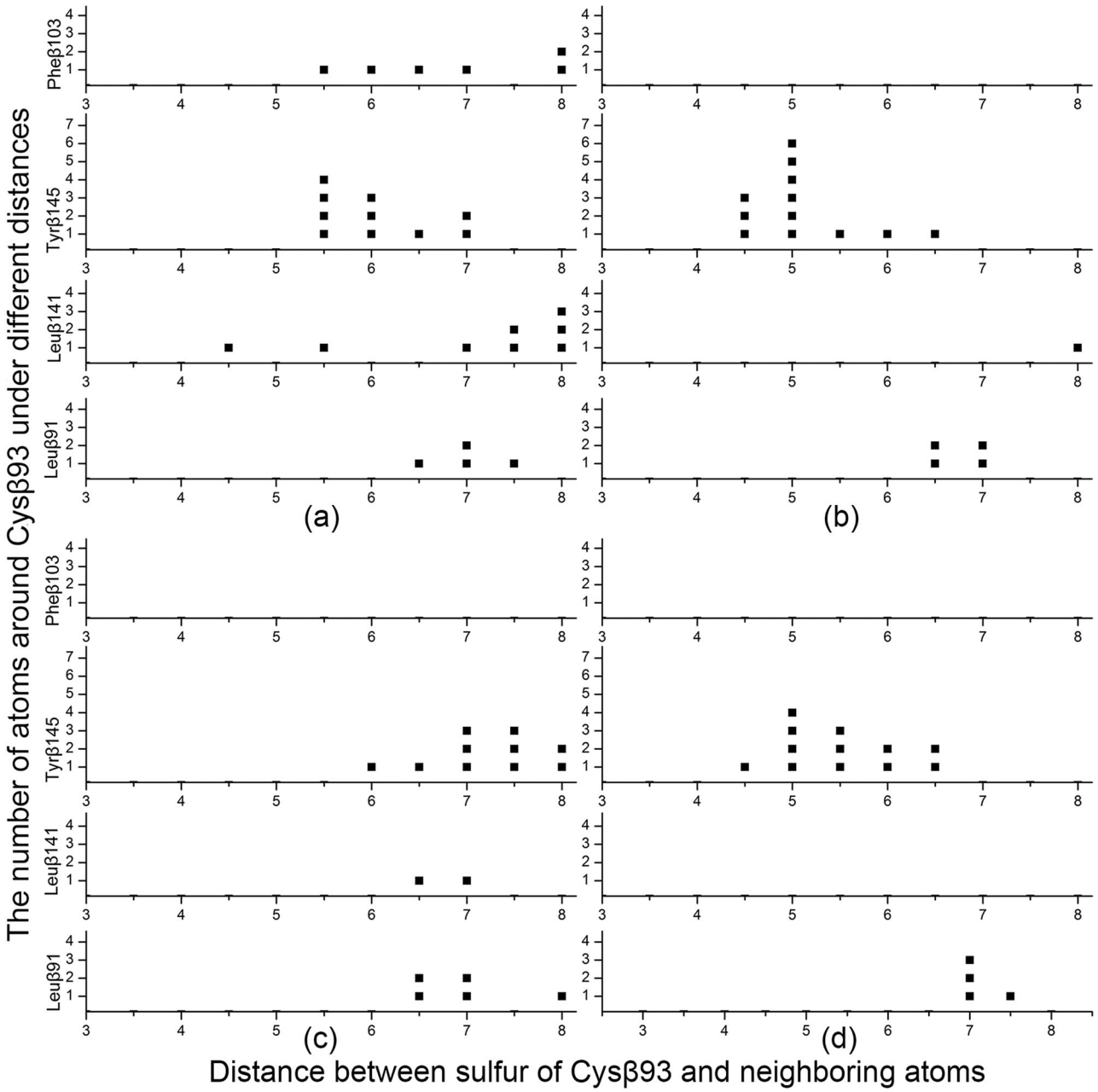

Fig. 5 The neighboring big-volume residues around Cysβ93 in R-state (a), T-state (b), dep-R-state (c) and dep-T-state (d) hemoglobin. *x*-axis represents the distance between sulfur of Cysβ93 and the neighboring atoms from different residues. *y*-axis (■) represents the number of atoms.

moved away in dep-R-state hemoglobin. Moreover, these big-volume amino acids were not observed around the Cysβ93 within the radius of 6 Å in dep-R-state hemoglobin (Fig. 5c). Compared to dep-R-state hemoglobin, Tyrβ145 was closer to Cysβ93 in dep-T-state hemoglobin (Fig. 5c and d).

Based on these analyses, we extracted the characteristic that few residues with big volume were located around S-nitrosylated cysteine sites in protein structures. This represented less steric hindrance occurring around S-nitrosylated cysteines. Combined with the mechanism of oxidant-mediated S-nitrosylation, the big-volume residues would prevent the oxidant from attacking cysteine. In summary, it was indicated that steric hindrance would inhibit protein S-nitrosylation. This result was in good agreement with another oxidant-mediated post-translational modification of tyrosine nitration.[44]

S-nitrosylated cysteines are more flexible

The *B*-factor can reflect cysteine's fluctuation, and it can be used to evaluate the flexibility of cysteine. The *B*-factor values of the S-nitrosylated and non-S-nitrosylated cysteines were calculated using Karplus algorithm. Our calculations indicated that the S-nitrosylated cysteines (score was 0.965 ± 0.036) were more flexible than the non-S-nitrosylated cysteines (score was 0.958 ± 0.035).

For hemoglobin, there were three cysteines, including Cysα104, Cysβ93 and Cysβ112. The Cysβ93 in R-state can be S-nitrosylated. Cysteines of Cysα104, Cysβ112 in R-state, as well as cysteines of Cysα104, Cysβ112 and Cysβ93, in T-state cannot be S-nitrosylated; thus, these five cysteines were used as non-S-nitrosylated cysteines. In the MD simulations of hemoglobin, the results showed that the Cysβ93 in the R-state had a larger RMSF than the other five cysteines in both R- and T-state hemoglobin (Table 1), suggesting that S-nitrosylated Cysβ93 was more flexible. It was also found that there was a free space (26.25 $Å^3$) near the Cysβ93 in the R-state hemoglobin (Fig. S3, ESI†), which would allow the Cysβ93 to move freely. This indicated that the Cysβ93 in the R-state hemoglobin would have more opportunity to make contact with oxidant agents.

Furthermore, the distribution of dihedral angles (SG-CB-CA-C) in the Cysβ93 of hemoglobin was obtained, where SG, CB, CA and C denoted the sulfur atom, the carbon of side chain, the alpha carbon and the carbon of the main chain, respectively. The dihedral angles in the R-state hemoglobin were stabilized at $\sim 0^\circ$ (Fig. 6a), while the dihedral angles in the T-state hemoglobin moved back and forth at $\sim 120^\circ$ and $\sim 240^\circ$ (Fig. 6b). Moreover, the dihedral angles of the Cysβ93 in the dep-R-state hemoglobin fluctuated much more, and they could adopt values of $\sim 0^\circ$, $\sim 120^\circ$ and $\sim 240^\circ$ compared to that in the dep-T-state (Fig. 6c and d).

Table 1 RMSF of cysteines in R-state and T-state hemoglobin. Cysβ93 in the R-state can be S-nitrosylated. Cysteines in the T-state, as well as Cysα104 and Cysβ112 in the R-state, cannot be S-nitrosylated

	R-state (Å)	T-state (Å)
Cysα104	0.445	0.455
Cysβ93	**0.995**	0.755
Cysβ112	0.464	0.448

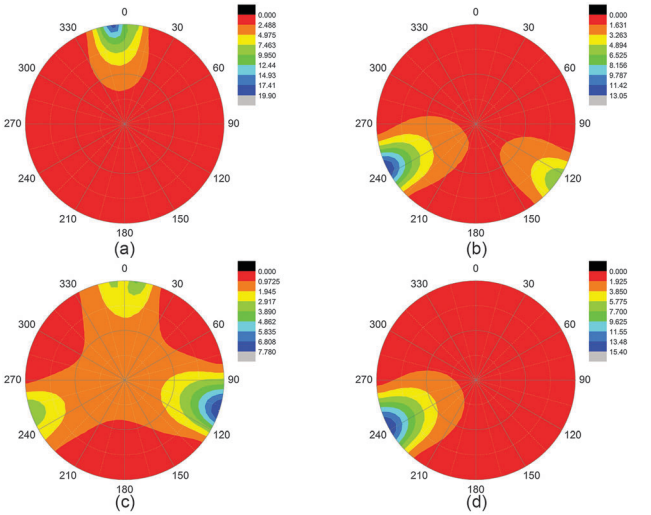

Fig. 6 The dihedral angle of Cysβ93 in R-state (a), T-state (b), dep-R-state (c) and dep-T-state (d) hemoglobin. Dihedral angle was analysed from 0° to 360°. The color bar represents frequency (%) of dihedral angle.

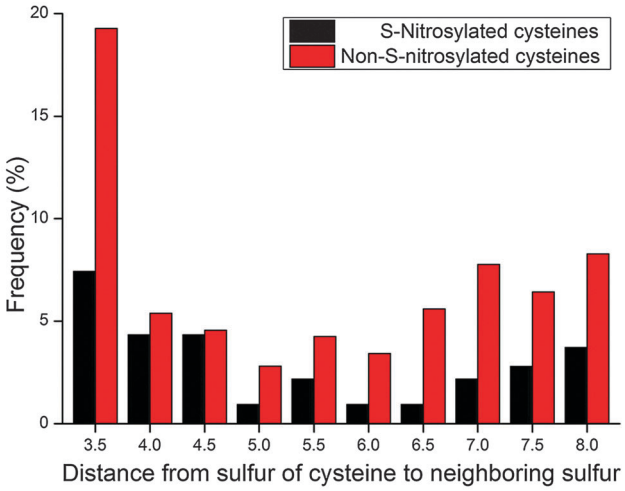

Fig. 7 The atomic distribution of sulfur atoms of neighboring cysteines in S-nitrosylated and non-S-nitrosylated cysteine sets. *x*-axis represents the distance from the sulfur of S-nitrosylated or non-S-nitrosylated cysteines to the sulfur of neighboring cysteines. *y*-axis represents the frequency of neighboring sulfur atoms, which was normalized by the number of S-nitrosylated cysteines and the number of non-S-nitrosylated cysteines, respectively.

In brief, S-nitrosylated cysteines were more flexible in protein structures. When cysteines were deprotonated, the dihedral angles of the S-nitrosylated cysteines were more active and conducive to the sulfur of the S-nitrosylated cysteines making contact with oxidants.

Less Cys residues locate around S-nitrosylated cysteines

By calculating the atomic distribution of sulfur atoms from neighboring cysteines, we found that the sulfur atoms of these residues were less available around the S-nitrosylated cysteines (Fig. 7). The previous study also showed a similar phenomenon,

in which there were fewer other cysteines in the neighboring region.[24] Especially, when the distance was less than 3.5 Å (disulfide bond length is about 2.05 Å), the frequency of sulfur atoms of Cys around S-nitrosylated cysteine sites was much lower than that around the non-S-nitrosylated (Table S6, ESI†). It was considered that these reductive cysteines would compete with the cysteines for oxidant agents in the process of S-nitrosylation.

Deprotonation of non-S-nitrosylated cysteines causes instability of protein structures

In addition to the consistent characteristics found in both statistical and molecular simulations, the instability of hemoglobin was noted when the non-S-nitrosylated cysteine of Cysβ93 was deprotonated in the dep-T-state hemoglobin. Based on the results of MD simulations, which were repeated 14 times, the root-mean-square deviation (RMSD) of each trajectory was analyzed to evaluate structural changes in R-, T-, dep-R- and dep-T-state (Fig. S4, ESI†). To assess the rate of change in hemoglobin structure during MD simulations, the average differential coefficient of the RMSD of each trajectory was calculated, where a larger average differential coefficient meant a larger change in protein structure. Thus, for each state of hemoglobin (R-, T-, dep-R-, and dep-T-state), there were 14 values of average differential coefficient. According to the distribution of the average differential coefficient, the structure of dep-T-state hemoglobin changed greatly compared to the dep-R-, T- and R-state (Fig. 8), suggesting a protein structure would become unstable when a "non-S-nitrosylated" cysteine was deprotonated. Furthermore, RMSD of each amino acid residue was calculated, and amino acids with larger RMSF gathered at the loop between Aspβ79 and Glyβ83 (Fig. S5, ESI†). It was proposed that the deprotonation of Cysβ93 might enhance the movement of the loop through allosteric regulation.

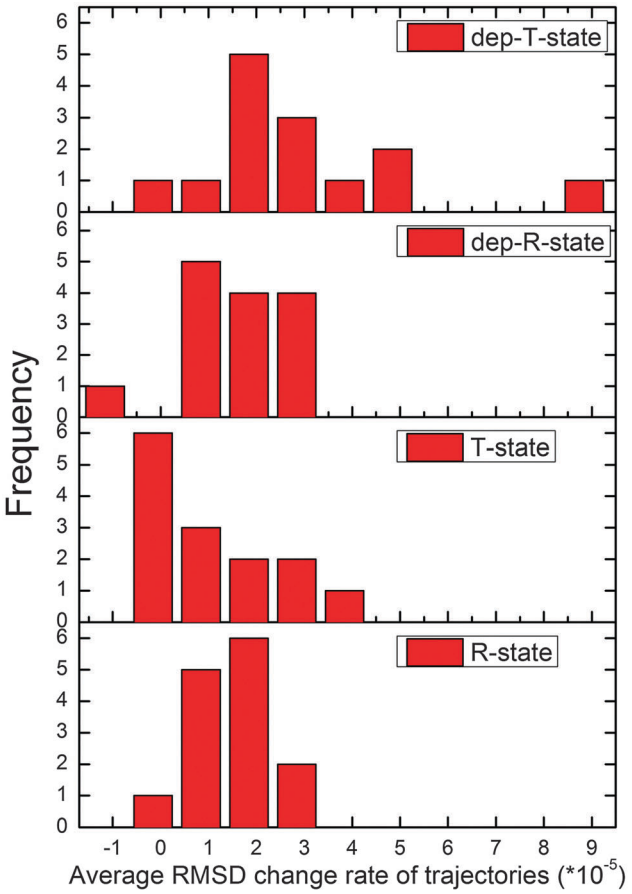

Fig. 8 The rate of change for RMSDs of R-state, T-state, dep-R-state and dep-T-state hemoglobin in repeated MD simulations. *x*-axis represents average rate of change calculated by RMSD curve in MD simulations. *y*-axis represents the frequency of different rate of change.

Discussion

In this work, we collected 213 S-nitrosylated proteins with structural information, containing 323 S-nitrosylated cysteine sites and 965 non-S-nitrosylated cysteine sites. First, compared to the non-S-nitrosylated cysteines, the S-nitrosylated cysteines had a lower pK_a, a higher abundance of basic residues, such as His and Lys, and a lower abundance of aromatic and aliphatic residues such as Phe, Tyr and Leu. The basic residues formed a basic environment around the S-nitrosylated cysteines in S-nitrosylated proteins, while the aromatic and aliphatic residues formed a non-polar environment around the non-S-nitrosylated. It was suggested that a basic environment could enhance cysteines' deprotonation, which might be an important step in S-nitrosylation. Furthermore, in MD simulations of hemoglobin, basic residues moved closer to the deprotonated Cysβ93 in the dep-R-state hemoglobin, suggesting that basic environments formed by basic residues can further stabilize deprotonated cysteines. Second, the S-nitrosylated cysteines had a lower population of big-volume residues, such as Phe, Tyr and Leu, in the half ball of 8 Å. In MD simulations of hemoglobin, Pheβ103 and Tyrβ145 moved away when Cysβ93 was deprotonated in the dep-R-state hemoglobin, while Tyrβ145 was closer to Cysβ93 in dep-T state hemoglobin. Based on the mechanism of oxidant-mediated S-nitrosylation, the big-volume residues would prevent oxidant agents from approaching to make contact with the cysteine sites. This indicated that steric hindrance might make significant contributions to the process of S-nitrosylation. Third, by analysis of the *B*-factor, S-nitrosylated cysteines were more flexible in structure. In MD simulations of hemoglobin, the Cysβ93 in the R-state hemoglobin had a higher RMSF than that in the T-state. It was speculated that the flexibility of cysteines could affect S-nitrosylation as well. In particular, the proper fluctuation of the SG-CB-CA-C conformation in dep-R-state allowed the Cysβ93 sulfur atom to contact NO reagent by chance. In addition, a low abundance of the sulfur atoms of neighboring Cys was observed in S-nitrosylated cysteines.

Hemoglobin adopts the R-state under oxygenation in lung, and the T-state under deoxygenation in vein.[27] A probable process of S-nitrosylation was proposed in hemoglobin according to our analyses. First, under physiological pH, basic residues would enhance the deprotonation of the Cysβ93 in R-state hemoglobin. Second, the basic residues moved closer to the deprotonated Cysβ93 and stabilized it. Third, the big-volume residues moved

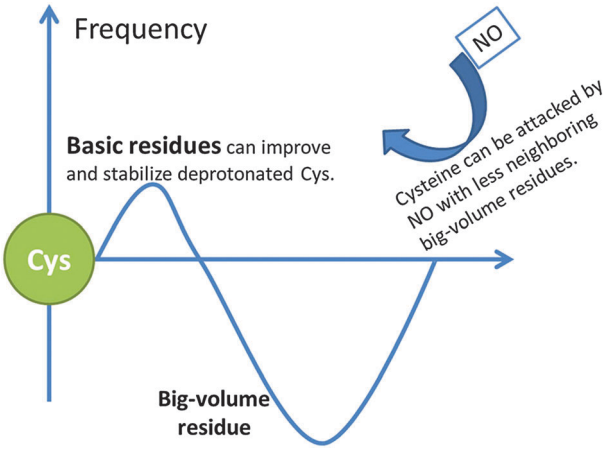

Fig. 9 The 3-D structural model of S-nitrosylation. *x*-axis represents the distance to the sulfur of cysteine (green ball). *y*-axis represents the frequency of the neighboring residues.

away from the deprotonated Cysβ93. Then, in coordination with high flexibility, the Cysβ93 was S-nitrosylated by adding an NO group in lung. However, these S-nitrosylated characteristics were not present in T-state hemoglobin. Since the S-nitrosylation was reversible, NO was released from the T-state hemoglobin in vein. This showed that the blood flow was regulated by S-nitrosylation.

In a previous study, S. M. Marino proposed a revised acid-based motif and a structural model (covered 15 S-nitrosylated proteins).[24] In our work, we proposed a 3-dimensional model of S-nitrosylation associated with the distribution of amino acid residues in protein structures (Fig. 9). If there were more basic residues (including His and Lys) and less big-volume residues (including Phe, Tyr and Leu) around a cysteine, it would be easier for the cysteine to be S-nitrosylated. The reason was that a cysteine could be deprotonated and stabilized by neighboring basic residues. The charged environment formed by basic residues was necessary for S-nitrosylation.[44] Furthermore, less big-volume residues, which were located around cysteine sites in protein structures, would allow NO agent to access the target cysteines. This model might shed light on further understanding of the process of S-nitrosylation.

In the model, there were two major conditions: (I) atom number of basic amino acids (His and Lys) were non-zero in a distance of less than 5 Å (the distance was from the sulfur of cysteine to the neighboring atoms); (II) atom number of big-volume residues (Phe, Tyr and Leu) were less than 22 in a distance of less than 8 Å (22 was the average number of big-volume atoms in the data set). In addition to the two main conditions, a minor condition from statistical analysis (Fig. 7, Table S1, ESI†) was extracted: *i.e.* no other cysteines existed in a distance less than 5 Å. Under these three conditions, the 3-dimensional S-nitrosylation model explained 200 out of 323 S-nitrosylated cysteine sites (61.92%) and 561 out of 965 non-S-nitrosylated cysteine sites (58.13%).

Conclusion

In this work, 213 protein structures with 323 S-nitrosylated cysteine sites and 965 non-S-nitrosylated cysteine sites were collected to study characteristics of S-nitrosylation using statistical analyses and MD simulations, where hemoglobin was employed as a case. Four major characteristics were observed, including (1) lower pK_a, (2) higher population of basic residues, (3) lower population of big-volume residues in the neighborhood, and (4) higher flexibility. MD simulations of hemoglobin showed that basic residues could enhance deprotonation of S-nitrosylated cysteines, and basic environments could further stabilize deprotonated cysteines. Moreover, steric hindrance caused by big-volume residues would play an important role in the process of S-nitrosylation, and the flexibility of cysteines could also affect the S-nitrosylation. In conclusion, a 3-dimensional model of S-nitrosylation was proposed, which could explain 61.9% of the S-nitrosylated and 58.1% of the non-S-nitrosylated cysteines. Our study suggested that deprotonation was a prerequisite for protein S-nitrosylation, and the model would improve the understanding of S-nitrosylation in structural characteristics. It needs more structural characteristics based on a larger dataset to make a more powerful model.

Acknowledgements

YLZ, LX, and YXL conceived and designed the investigation. SLC, JL, and HYW performed the structure analyses and MD simulations. SLC, TS, XLW, and YLZ wrote the paper. This work is supported in part by the National High-Tech R&D Program of China "863" (No. 2012AA020403), the National Basic Research Program of China "973" (No. 2012CB721005, 2013CB966802), the National Science Foundation of China (No. 21377085, 21303101, 31121064, J1210047, and 91230105), the MOE New Century Excellent Talents in University (No. NCET-12-0354), and the SJTU-HPC computing facility award.

Notes and references

1 D. T. Hess, A. Matsumoto, S. O. Kim, H. E. Marshall and J. S. Stamler, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 150–166.
2 M. Benhar, M. T. Forrester and J. S. Stamler, *Nat. Rev. Mol. Cell Biol.*, 2009, **10**, 721–732.
3 Y. Iwakiri, *Nitric Oxide*, 2011, **25**, 95–101.
4 M. W. Foster, T. J. McMahon and J. S. Stamler, *Trends Mol. Med.*, 2003, **9**, 160–168.
5 A. M. Evangelista, M. J. Kohr and E. Murphy, *Antioxid. Redox Signaling*, 2013, **19**, 1209–1219.
6 S. R. Tannenbaum and J. E. Kim, *Nat. Chem. Biol.*, 2005, **1**, 126–127.
7 S. R. Jaffrey, H. Erdjument-Bromage, C. D. Ferris, P. Tempst and S. H. Snyder, *Nat. Cell Biol.*, 2001, **3**, 193–197.
8 Z. Q. Wang, *Cancer Lett.*, 2012, **320**, 123–129.
9 D. D. Thomas and D. Jourd'heuil, *Antioxid. Redox Signaling*, 2012, **17**, 934–936.
10 J. G. Fang, T. Nakamura, D. H. Cho, Z. Z. Gu and S. A. Lipton, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 18742–18747.
11 D. H. Cho, T. Nakamura, J. G. Fang, P. Cieplak, A. Godzik, Z. Gu and S. A. Lipton, *Science*, 2009, **324**, 102–105.

12 C. Lindermayr and J. Durner, *J. Proteomics*, 2009, **73**, 1–9.
13 H. Li, X. Xing, G. Ding, Q. Li, C. Wang, L. Xie, R. Zeng and Y. Li, *Mol. Cell. Proteomics*, 2009, **8**, 1839–1849.
14 G. Hao, B. Derakhshan, L. Shi, F. Campagne and S. S. Gross, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 1012–1017.
15 F. Torta, V. Usuelli, A. Malgaroli and A. Bachi, *Proteomics*, 2008, **8**, 4484–4494.
16 B. Derakhshan, P. C. Wille and S. S. Gross, *Nat. Protoc.*, 2007, **2**, 1685–1691.
17 D. A. Mitchell and M. A. Marletta, *Nat. Chem. Biol.*, 2005, **1**, 154–158.
18 K. A. Broniowska, *Antioxid. Redox Signaling*, 2012, **17**, 969.
19 G. Czapski and S. Goldstein, *Free Radical Biol. Med.*, 1995, **19**, 785–794.
20 Y. Yang and J. Loscalzo, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 117–122.
21 G. Stubauer, A. Giuffre and P. Sarti, *J. Biol. Chem.*, 1999, **274**, 28128–28133.
22 Y. J. Chen, W. C. Ku, P. Y. Lin, H. C. Chou, K. H. Khoo and Y. J. Chen, *J. Proteome Res.*, 2010, **9**, 6417–6439.
23 A. Martinez-Ruiz, L. Villanueva, C. Gonzalez de Orduna, D. Lopez-Ferrer, M. A. Higueras, C. Tarin, I. Rodriguez-Crespo, J. Vazquez and S. Lamas, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 8525–8530.
24 Y. Xu, J. Ding, L. Y. Wu and K. C. Chou, *PLoS One*, 2013, **8**, e55844.
25 S. M. Marino and V. N. Gladyshev, *J. Mol. Biol.*, 2010, **395**, 844–859.
26 P. T. Doulias, J. L. Greene, T. M. Greco, M. Tenopoulou, S. H. Seeholzer, R. L. Dunbrack and H. Ischiropoulos, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 16958–16963.
27 T. J. McMahon, *Nat. Med.*, 2002, **8**, 711–717.
28 J. S. Stamler, L. Jia, J. P. Eu, T. J. McMahon, I. T. Demchenko, J. Bonaventura, K. Gernert and C. A. Piantadosi, *Science*, 1997, **276**, 2034–2037.
29 J. P. Pezacki, N. J. Ship and R. Kluger, *J. Am. Chem. Soc.*, 2001, **123**, 4615–4616.
30 L. Jia, C. Bonaventura, J. Bonaventura and J. S. Stamler, *Nature*, 1996, **380**, 221–226.
31 Y. Xue, Z. X. Liu, X. J. Gao, C. J. Jin, L. P. Wen, X. B. Yao and J. A. Ren, *PLoS One*, 2010, **5**, e11290.
32 W. Li and A. Godzik, *Bioinformatics*, 2006, **22**, 1658–1659.
33 M. H. M. Olsson, C. R. Sondergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525–537.
34 T.-Y. Lee, Y.-J. Chen, T.-C. Lu, H.-D. Huang and Y.-J. Chen, *PLoS One*, 2011, **6**, e21849.
35 P. Karplus and G. Schulz, *Naturwissenschaften*, 1985, **72**, 212–213.
36 D. J. Singel and J. S. Stamler, *Annu. Rev. Physiol.*, 2005, **67**, 99–145.
37 B. Shaanan, *J. Mol. Biol.*, 1983, **171**, 31.
38 G. Fermi, *J. Mol. Biol.*, 1984, **175**, 159.
39 L. Mouawad, D. Perahia, C. H. Robert and C. Guilbert, *Biophys. J.*, 2002, **82**, 3224–3245.
40 O. K. Yusuff, J. O. Babalola, G. Bussi and S. Raugei, *J. Phys. Chem. B*, 2012, **116**, 11004–11009.
41 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 712–725.
42 J. P. Ryckaert, *J. Comput. Phys.*, 1977, **23**, 327.
43 R. Gnanasekaran, Y. Xu and D. M. Leitner, *J. Phys. Chem. B*, 2010, **114**, 16989–16996.
44 S. Cheng, B. Lian, J. Liang, T. Shi, L. Xie and Y. L. Zhao, *Mol. Biosyst.*, 2013, **9**, 2860–2868.
45 M. R. Talipov and Q. K. Timerghazin, *J. Phys. Chem. B*, 2013, **117**, 1827–1837.
46 N. Gould, P. T. Doulias, M. Tenopoulou, K. Raju and H. Ischiropoulos, *J. Biol. Chem.*, 2013, **288**, 26473–26479.
47 J. Liang, S. Cheng, J. Hou, Z. Xu and Y. L. Zhao, *Sci. China Chem.*, 2012, **55**, 2081–2088.

