Characterization and Identification of Protein *S*-Nitrosylation Sites Based on Tertiary Structures

Tzong-Yi Lee[1,2*], Justin Bo-Kai Hsu[3], Tzu-Hao Chang[4]

[1] School of Life and Health Sciences, The Chinese University of Hong Kong, Shenzhen 518172, China.
[2] Warshel Institute for Computational Biology, The Chinese University of Hong Kong, Shenzhen 518172, China.
[3] Department of Medical Research, Taipei Medical University Hospital, Taipei 110, Taiwan.
[4] Graduate Institute of Biomedical Informatics, Taipei Medical University, Taipei 110, Taiwan.

* Corresponding author. Tel.: +86(0) 755-23519551; email: leetzongyi@cuhk.edu.cn

Manuscript submitted August 23, 2019; accepted November 3, 2019.
doi: 10.17706/ijbbb.2020.10.1.15-25

Abstract: *S*-nitrosylation (SNO) is a sulfur atom occurring in cysteine amino acid in the protein connected to nitric oxide (NO), and it is one of the most important and universal Post-translational modifications. Nitroso modification will affect regulating cell function and information transfer. In recent years, there were many studies have developed the method of indentify S-nitrosylation substrate site in silicon. Unfortunately, people did not find significant characteristics for identification in protein sequence. Therefore, this study aims to explore structural characteristics on tertiary structures of S-nitrosylated proteins. As the number of the crystal structure in the PDB increases, also many SNO sites have been experimentally verified, we characterized these substrate sites containing 3D structures by structural analysis methods, such as Spatial amino acid composition, side chain orientation, and DSSP relative solvent accessible area. Besides, the support vector machine (SVM) was employed to generate the predictive model with the consideration of both sequential and spatial features. According to the evaluation of five-fold cross-validation, the spatial model could obtain a higher accuracy than the sequential model. Additionally, this work revealed that the model concerning multiple spatial features could achieve best performance in the prediction of SNO sites.

Key words: *S*-nitrosylation, SNO, nitric oxide, tertiary structure, side chain orientation, support vector machine.

1. Introduction

Protein *S*-nitrosylation (SNO), one of the various posttranslational modifications (PTMs), is a selective and reversible PTM by addition of nitric oxide (NO) moiety to cysteine (Cys) sulfur atom in proteins, critically regulates protein activity, localization and stability. SNO is thought to play a role which similar to phosphorylation, as a pleiotropic regulator that elicits dual effects to regulate diverse pathophysiological processes in plant [1] and human diseases, especially in cardiovascular disease and protection [2], the immune response [3], neurodegenerative disease [4], inflammation [5] and cancers [6]. Differential expression of various SNO targets modulate the localization, stability, and activity of proteins [7]-[9]. The SNO status of proteins may be linked to many cancer therapy outcomes as well as therapeutic-resistance, generating the need to develop SNO-related anti-cancer therapeutics [10]. With the importance of SNO in molecular processes, numerous efforts have been directed toward compute biological studies using protein sequential method to identified S-nitrosylation site in silicon [11]-[15]. Among the numerous studies, there are no significant properties in protein sequence around SNO sites. People try to develop more advanced methods to identify SNO sites.

Due to the progress of the structural genomics projects, there are more than 100,000 structures in Protein Data Bank (PDB) [16]. With the increase in crystal structure, more PTM site can mapped on known structures. And more researches try to dig out useful information from these structures. Durek and co-workers characterized phosphorylation sites by spatial amino acid propensity distributions to generate spatial signature motifs and the subsequent assessment of this information to improve the prediction of phosphorylation sites in proteins [17]. Karabulut *et al*. present the first comprehensive analysis of global and tissue-specific sequence and structure properties of lysine acetylation sites based on recent experimental data [18]. The study of Marino *et al*. have tried to observe characteristics of endogenously SNO modified in WT mouse liver on protein 3D structures, and concluded the endogenous S-nitrosoproteome in the liver has structural features that accommodate multiple mechanisms for selective site-directed S-nitrosylation [19].

As the number of experimental SNO sites and protein crystal structure grows, many SNO sites can be observed in the 3D structure now. The stably growing in vivo or in vitro SNO sites have prompted an increasing interest in the structural characterization of SNO substrate sites. In this study, all experimentally confirmed SNO peptides were used to investigate their spatial context such as spatial amino acid composition, solvent-accessible surface area, secondary structures, side chain orientation and structurally neighboring amino acids of SNO sites. Additionally, we build SVM models by above method of investigation, and we got performance better than SVM models which only use sequence information.

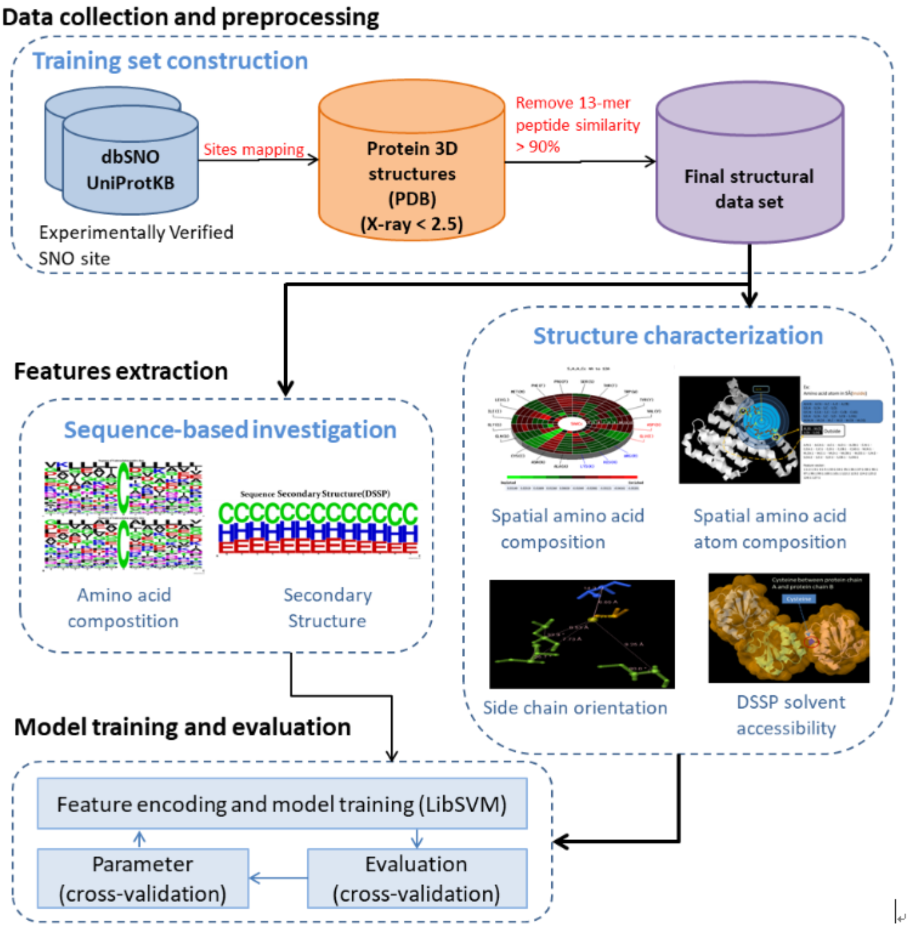
Data collection and preprocessing
Training set construction
dbSNO
UniProtKB
Experimentally Verified SNO site
Sites mapping
Protein 3D structures (PDB) (X-ray < 2.5)
Remove 13-mer peptide similarity > 90%
Final structural data set
Features extraction
Sequence-based investigation
Sequence Secondary Structure(DSSP)
Amino acid compostition
Secondary Structure
Structure characterization
Spatial amino acid composition
Spatial amino acid atom composition
Side chain orientation
DSSP solvent accessibility
Model training and evaluation
Feature encoding and model training (LibSVM)
Parameter (cross-validation)
Evaluation (cross-validation)

Fig. 1. Flowchart of this work.

2. Material and Methods

2.1. Data Collection and Preprocessing

Fig. 1 depicts the system flow of the proposed method, including data collection and preprocessing, sequence-based investigation, structural characterization, model training and evaluation. The experimentally verified *S*-nitrosylation sites are mainly extracted from dbSNO [20] and release 20140711 of UniProtKB [21]. In this study, the data set extracted from dbSNO and UniProtKB is regarded as the training set for sequential and structural investigation of *S*-nitrosylation sites. After removing the redundant sites between dbSNO and UniProtKB, the number of cystein (C) substrate sites is 4165. Based on structure region annotation from PDB and UniProtKB, we removed sites which aren't located on known 3D structure regions. Finally we got 298 cysteine sites which can map on 3D structure as positive set. On the other hand, we view non- s-nitrosylation cysctein sites which aren't around 12Å from positive sites in spatial as negative data set. Final training data set as shown in **Table 1** and each similarity of 13-mer flanking peptide of cysteine central wasn't over 90%.

Table 1. Statistics of Data Set in This Investigation

	Number of positive site	Number of negative site	Number of SNO proteins
Data set	298	763	191

2.2. Sequence-Based Investigation

In order to compare sequence-based investigation with structure characterization, some basic sequential peptide character also need to be investigated. The 13-mer flanking sequences (from -6 to +6) from the substrate sites (position 0) are extracted for sequence-based investigation. Since the flanking sequences of the substrate sites are graphically visualized as the entropy plots of sequence logo, the conservation of amino acids surrounding the s-nitrosylation sites can be easily observed. We used amino acid composition method to investigate sequential peptide of SNO sites because it is a popular way for SNO site in past research. Each SNO peptide was encoded to 20 dimensions for SVM model training. In addition to the composition of amino acids of 13-mer peptide, the secondary structure (SS) around the s-nitrosylation sites was also investigated. Since these subtract sites have mapped on PDB 3D structure, the flanking sequence must have experimentally secondary structure. We extracted the secondary structure from DSSP files. The output of DSSP explained extensively under 'explanation'. The very short summary of the output is: H = α-helix, B = residue in isolated β-bridge, E = extended strand, participates in β ladder, G = 3-helix (310 helix), I = 5 helix (π-helix), T = hydrogen bonded turn, S = bend. We consolidated the output into three terms, "H," "E" and "C" which stand for helix, sheet and coil, respectively, and it was encoded to "0," "1" and "2" for SVM model training.

2.3. Spatial Amino Acid Composition

This characterized method was first used in identified kinase specific phosphorylation site [17], [22]. After that, many people used similar method to identify other PTM site which doesn't have signal sequence motif and they also got some significant characteristics from 3D structure. In this study, a spatial amino acid composition (Spatial AAC) is determined by calculating the relative frequencies of 20 amino acid types within radial distances, ranging from 4 to 12 Å and centralized by s-nitrosylated amino acid residue. In addition, the radial distances is determined by calculating any two amino acid C-alpha atom. A radial cumulative propensity plot was applied to display the spatial AAC. In order to identify the significant difference of spatial AAC between *S*-nitrosylation sites (positive data) and non-s-nitrosylation sites (negative data), a measurement of F-score has been applied to calculate a statistical value for each radial distance.

2.4. Spatial Amino Acid-Atom Composition

Since NO is a very small molecule, we think even the difference of atom level may affect the modification of cysteine done by NO. A study of Lin *et al*. in 2013, mentioned that using amino acid atom distribution was to help predict DNA binding site on protein [23]. Based on above method, we further observed distribution of the atoms of each amino acid. We divided 20 amino acids to 167 amino acid & atom combinations (AA-atom). For example, Glycine contains 4 atoms, "N", "CA", "C" and "O", and Glycine became 4 combinations: G-N, G-CA, G-C, G-O. Same as spatial amino acid composition, it is determined by calculating the relative frequencies of 167 AA-atom types within radial distances ranging from 4 to 12 Å and centralized by s-nitrosylated cystein Sulfur-Gama(SG) atom.

2.5. Side Chain Orientation

The side chain orientation is an important indicator to determine a catalytic site on protein structure. Chien *et al*. have used this method to predict protein catalytic residues [24]. According to their study, catalytic residue is a residue whose distance to the catalytic central is less than 10Å and side chain angle to catalytic central is less than 80 degree. In this study, the side chain orientations of the amino acids which spatially surround the SNO substrate sites are determined to investigate the functional roles and to bind effects of the spatially neighboring amino acids to the substrate sites of NO attachment.

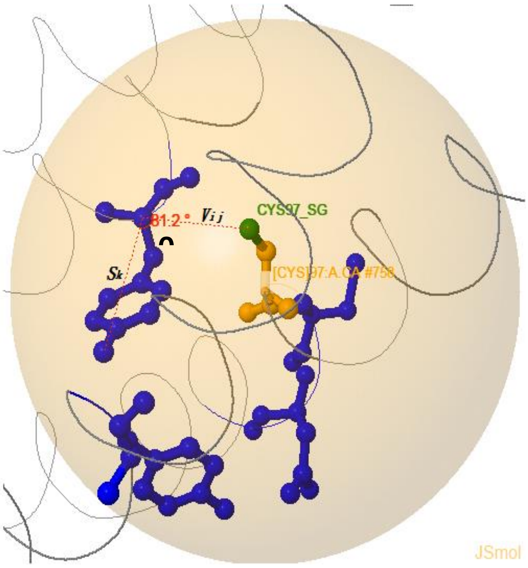

Fig. 2. Schematic representation of side chain orientation.

As illustrated in **Fig. 2**, given a spatially neighboring amino acid k which distance is less than 10Å from cysteine SG atom, the side chain direction of residue k is the vector s_k from its Cα atom to its functional atom:

$$s_k = X_k^F - X_k^{CA} \tag{1}$$

where $\boldsymbol{X}_k^F$ and $\boldsymbol{X}_k^{CA}$ are the functional atom and Cα atom of residue *k*. Table 2 lists the amino acid types whose functional atom is on the side chain (side chain functional amino acids). And we decided to use the functional atom for calculating side chain vector of the amino acids. For each cysteine SG atom i and any one of its surrounding residue j, the vector between point i and Cα atom of residue j is defined as:

$$v_{ij} = X_i - X_j \tag{2}$$

where X_i and X_j are the SG atom i and Cα atom of residue j. We computed the angle θ_{ij} between v_{ij} and s_j , which is the side chain vector of residue j,

$$\theta_{ij} = \mathbf{acos}\frac{v_{ij}\cdot s_j}{\|v_{ij}\|\,\|s_j\|} \tag{3}$$

the angle θ k less than 80∘ is defined as a functional residue to the cysteine residue of the SNO substrate site.

2.6. Model Construction and Evaluation

This study incorporates support vector machines (SVMs) with the sequential and structural features to generate the predictive models for the identification of S-nitrosylation sites. A public SVM library, namely LIBSVM, is applied for training the predictive models. The radial basis function (RBF): K(Si,Sj)=exp(−γ‖Si−Sj‖2) is selected as the kernel function of SVM. Five-fold cross-validation is used to evaluate the predictive performance of the models. The following measures of predictive performance of the trained models are defined as: Precision (Pre) = TP/(TP+FP), Sensitivity (Sn) = TP/(TP+FN), Specificity (Sp) = TN/(TN+FP) and Accuracy (Acc) = (TP + TN)/(TP+FP+TN+FN), where TP is true positive prediction; TN is true negative prediction; FP is false positive prediction and FN is false negative prediction respectively.

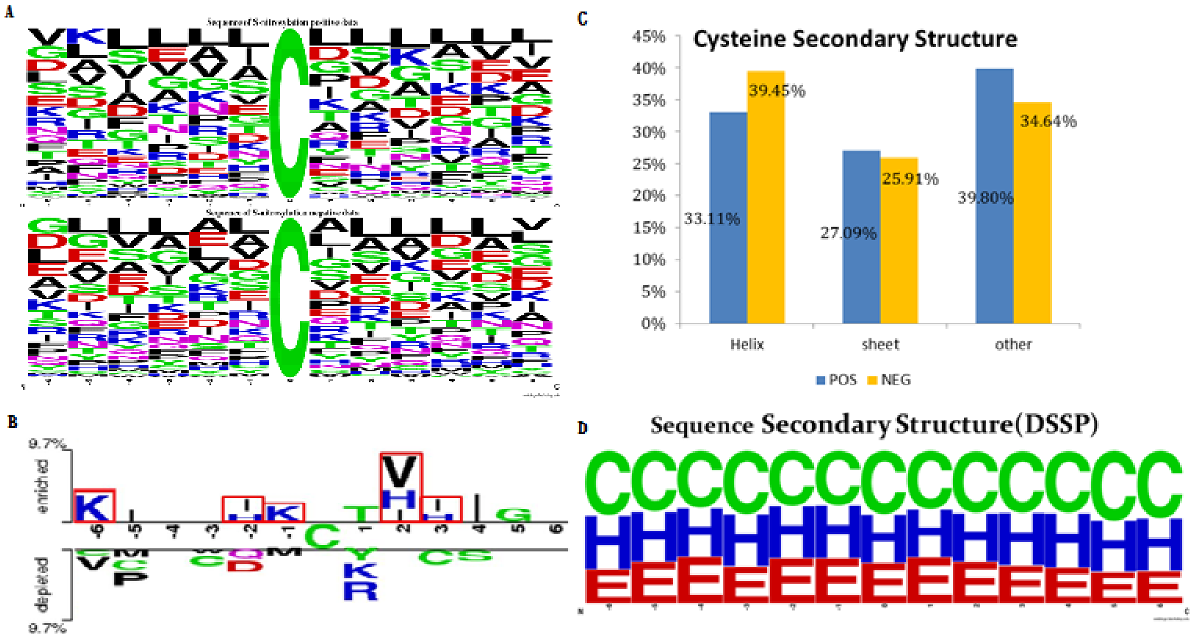

Fig. 3. Amino acid composition and secondary structures of SNO sites.

3. Results and Discussion

3.1. Sequential and Structural Amino Acid Composition of SNO Sites

In sequence-based investigation of substrate site, **Fig. 3A** shows that there is no significant motif in flank of SNO sites. However, the Two Sample Logo as presented in **Fig. 3B**, is a comparison of position-specific amino acid composition, between SNO sites (top) and non-SNO sites (bottom), indicates that the positively charged Lysine (K) and Histidine (H) residues are highly abundant around SNO sites. Consistent with previous studies, the SNO cysteine has the preference to locate in the regions flanking with acidic and basic amino acids. **Fig. 3C** and **3D** show the frequency of secondary structure of center substrate site and flanking sequence. Sequence logo shows that there is no significant secondary structure motif surrounding the SNO sites. There are still differences found on center cysteine that more positive site located on coil region and more negative site located on helix region, as presented in **Fig. 3C**.

When it comes to Structure characterization, **Fig. 4A** shows the radial cumulative propensity plot (spatial neighborhood) of amino acid composition, surrounding 298 SNO substrate sites based on tertiary structures in the range from 4 to 12 Å respectively. Found in the observation from Two Sample Logo, SNO has the significant enrichments of K and H residues in the sequential neighborhood of substrate sites. However, the radial cumulative propensity plots present that there is an additional enrichment of amino acids in the spatial neighborhood. In addition to the K and D residues, an enrichment of hydrophilic residues exists and includes Asparagine (N) and Proline (P), accompanied by a remarkable depletion of hydrophobic Cysteine (C) residue in the spatial neighborhood.

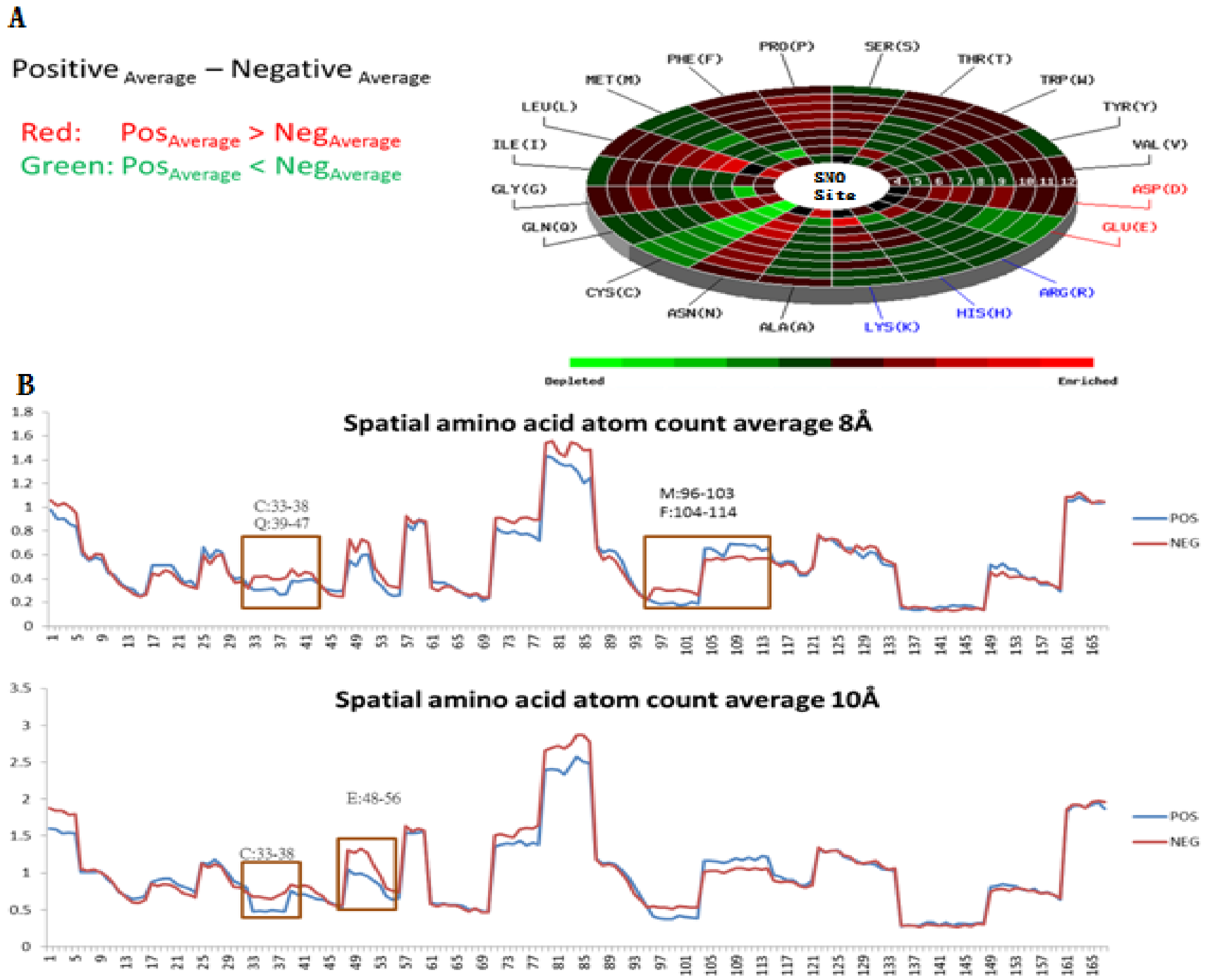

Fig. 4. Spatial amino acid and detailed atom composition.

Fig. 4B shows 167 AA-atom (X-axis) and the average of composition (y-axis). We found average composition for spatial range 8Å and 10Å has more difference between positive and negative. In 8Å range, significant difference between atom of number 33 to number 47(atom of Cysteine and Glutamine), and 96 to 114(atom of Methionine and Phenylalanine). When range zoomed to 10Å, trend of number 48 to number 56(atom of Glutamic acid) became significant. However, we found that not all atoms belong to same kind of residue can have same difference value between positive and negative data. Differences indeed exist at atomic level.

3.2. Side Chain Orientation of SNO Sites

For characterization of side chain orientation, we calculated the average of side chain angle of top 3 residues, which are nearest from central cysteine by above method. As shown in **Table 2**, the average is closely between positive and negative and both they are over than 80^0. Then we calculated average of side chain angle of each residue, which locates in range 10Å from central cysteine SG atom. The average of side chain angle of positive sites has been smaller than negative, and average of negative is more than 80^0. This means side chain angle of spatial neighbor residue have tendency to point to cystein SG atom (NO binding atom). In before researches, it was said that side chain of residue can occur attract for small molecular. If the side chains of surrounding amino acids tend to be in the same direction, this position will be easy to form a catalytic location. Concerning PTM issue, we can also use side chain orientation as the combination of molecules even this kind of molecules are much smaller than some ligand. There is no difference found in small range for residue side chain angle, only the range zoomed to 10Å, the difference can be tell. Overall, the combination of NO should be affected by the side chain angle of surrounding residues.

Table 2. Average of Side Chain Angle of Cysteine Surround Residues

	Nearest three surround residues in positive data set	Nearest three surround residues in negative data set	Distance < 10 Å surround residues in positive data set	Distance < 10 Å surround residues in negative data set
Average of side chain angle	94.97^0	93.28^0	53.17^0	81.36^0

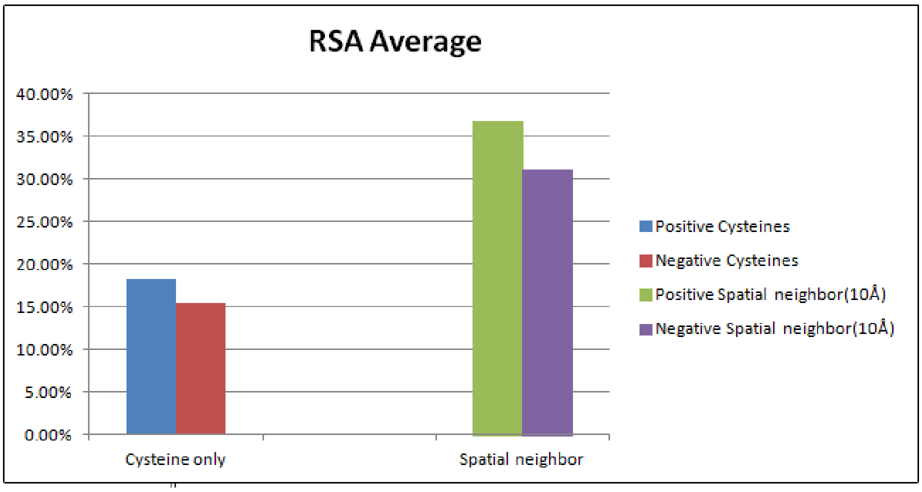

Fig. 5. Characteristics of relative solvent accessible area around SNO sites.

3.3. Relative Solvent Accessible Area of SNO Sites

Fig. 5 shows average of relative solvent accessible area of cysteine and spatial neighbor residues in 10Å, positive and negative respectively. Average value presented that positive RSA is bigger than both negative cysteine and surrounding residues. Even so, we observed the average of RSA in positive cysteine is still low, so we tried to observe the appearance of protein 3D structures. As shown in **Fig. 6**, many cysteine residues are coated with surrounded residues even distributed on the surface of proteins and this makes its RSA value to be low. We also found a pair of same protein structure which one of them has modified cysteine (SNC). Even cysteine has been modified, the residue still be buried by surrounded residues (RSA: 0.03). The

RMSD of two structures alignment is 0.241; this means these two structures have highly structure similarity. Thus we found even cysteine be buried in protein NO molecular, still it can bind on cysteine.

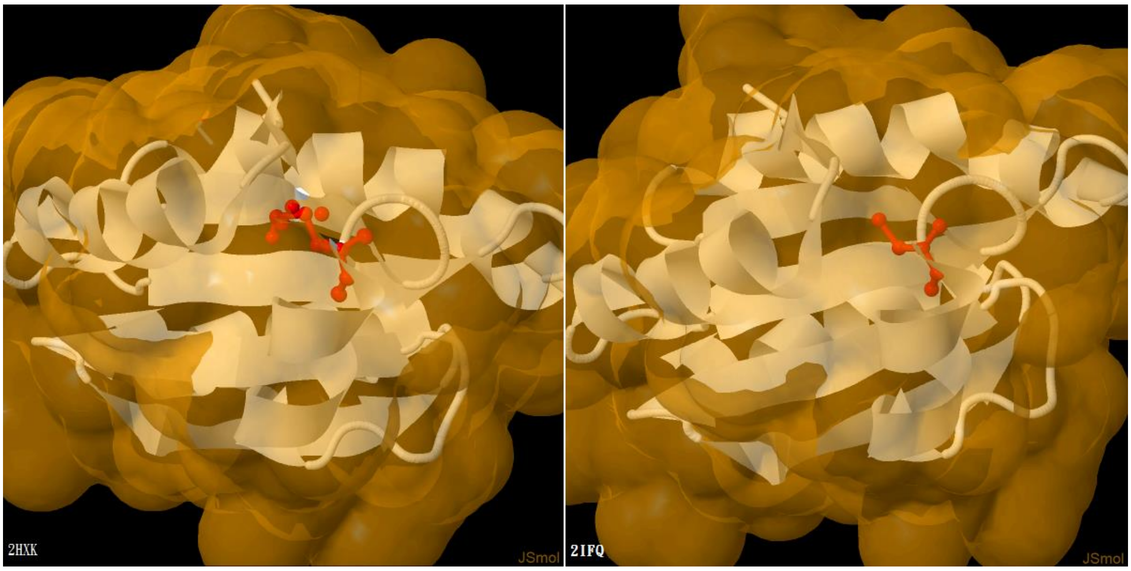

Fig. 6. Structure and solvent area comparison between same proteins. Left-hand side is the cysteine modified structure and its RSA value is 0.37. Right-hand side is the non-modified cysteine with RSA value as 0. RMSD of two structure alignment is 0.241.

Table 3. SVM Performance

Training features	Spatial Range (Å)	TP	FP	TN	FN	Sensitivity	Specificity	Precision	Accuracy
AAC	-	174	323	440	124	58.39%	57.67%	35.01%	57.87%
Secondary Structure	-	162	333	430	136	54.36%	56.37%	32.73%	55.79%
SPAAC	12	177	315	448	121	59.40%	58.72%	35.98%	58.91%
SPAAAC	10	180	296	467	118	60.40%	61.21%	37.82%	60.98%
Side chain orientation	10	165	316	447	133	55.37%	58.58%	34.30%	57.68%
SPAAAC + Side chain orientation	10	190	280	483	108	63.76%	63.30%	40.43%	63.43%
AAC + SPAAAC+ Side chain orientation	10	183	292	471	115	61.41%	61.73%	38.53%	61.64%

3.4. Cross-validation Performance

Table 3 shows all cross-validation performance of each sequential and spatial features and hybrid combination. For sequence feature, the predictive of amino acid composition of 13-mer peptide (AAC) sensitivity, specificity, precision, and accuracy are 58.39%, 57.67%, 35.01%, 57.87% respectively. The performance of secondary structure is less than AAC. This performance shows there is no significant characteristic in frank sequence for identified *S*-nitrosylation. For spatial feature, each method was adopted to do the test from range 4Å to 12 Å. Each measures of predictive performance increased when the spatial AAC at 12Å, as shown in **Table 3**. And then, spatial AA-atom compositions (SPAAAC) at 10 Å proposed better performance than above methods. This indicates that the difference in atomic level is preferably be used to identify cysteine *S*-nitrosylation. Performance of side chain orientation is 55.37%, 58.58%, 34.30% and 57.68% respectively. But the model trained with a combination of SPAAAC and side chain orientation performed best. Finally, we combined sequential and spatial best feature, the predictive performance was slightly declined but still better than sequential performance only. Overall, the features of spatial characteristic were more powerful than sequential characteristic as for identified *S*-nitrosylation site.

4. Conclusion

As many studies have described, it is hard to find significant features of s-nitrosylation site in sequence, we used protein structural attributes to dig out useful characterizations such as spatial amino acid composition, side chain orientation. Also, we used this method to identify s-nitrosylation site in silicon. Apart from s-nitrosylation site problem in sequence structure, there are many other issues share the same problem. As the number of crystalline structures grows rapidly, more and more researchers turn to study 3D structure. For now, many PTM sites can mapped on known PDB structures, such as glycosylaiton, acetylation, methylation etc. Karabulut *et al*. have used Spatial amino acid composition to identify tissue specific acetylation site and they also discovered that characteristic of 3D structure are different from the one of amino acid sequence [18]. To sum up, we proposed that spatial amino acid composition and side chain orientation methods can be used to enable the identification of PTM sites on protein 3D structure.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

TYL conducted the research; JBKH and THC analyzed the data; TYL and JBKH wrote the paper; all authors had approved the final version.

Acknowledgment

The authors sincerely appreciate the Warshel Institute for Computational Biology, The Chinese University of Hong Kong, Shenzhen, for financially supporting this research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] Lindermayr, C., Saalbach, G., & Durner, J. (2005). Proteomic identification of S-nitrosylated proteins in Arabidopsis. *Plant Physiol*, *137(3)*, 921-30.

[2] Lima, B., *et al*. (2010). S-nitrosylation in cardiovascular signaling. *Circ Res*, *106(4)*, 633-46.

[3] Tripathi, P. (2007). Nitric oxide and immune response. *Indian J Biochem Biophys*, *44(5)*, 310-9.

[4] Uehara, T., *et al*. (2006). S-nitrosylated protein-disulphide isomerase links protein misfolding to neurodegeneration. *Nature*, *441(7092)*, 513-7.

[5] Clancy, R. M., Amin, A. R., & Abramson, S. B. (1998). The role of nitric oxide in inflammation and immunity. *Arthritis Rheum*, *41(7)*, 1141-51.

[6] Xu, W., *et al*. (2002). The role of nitric oxide in cancer. *Cell Res*, *12(5-6)*, 311-20.

[7] Cardinale, A., Chiesa, R., & Sierks, M. (2014). Protein misfolding and neurodegenerative diseases. *Int J Cell Biol*, 217371.

[8] Hess, D. T., *et al*. (2005). Protein S-nitrosylation: Purview and parameters. *Nat Rev Mol Cell Biol*, *6(2)*, 150-66.

[9] Lam, Y. W., *et al*. (2010). Comprehensive identification and modified-site mapping of S-nitrosylated targets in prostate epithelial cells. *PLoS One*, *5(2)*, e9075.

[10] Wang, Z. (2012). Protein S-nitrosylation and cancer. *Cancer Lett*, *320(2)*, 123-9.

[11] Xue, Y., *et al*. (2010). GPS-SNO: Computational prediction of protein S-nitrosylation sites with a modified GPS algorithm. *PLoS One*, *5(6)*, e11290.

[12] Li, B.Q., *et al*. (2012). Predict and analyze S-nitrosylation modification sites with the mRMR and IFS approaches. *J Proteomics*, *75(5)*, 1654-65.

[13] Xu, Y., *et al*. (2013). iSNO-PseAAC: Predict cysteine S-nitrosylation sites in proteins by incorporating position specific amino acid propensity into pseudo amino acid composition. *PLoS One*, *8(2)*, e55844.

[14] Lee, T. Y., *et al*. (2011). SNOSite: Exploiting maximal dependence decomposition to identify cysteine S-nitrosylation with substrate site specificity. *PLoS One*, *6(7)*, e21849.

[15] Xu, Y., *et al*. (2013). iSNO-AAPair: Incorporating amino acid pairwise coupling into PseAAC for predicting cysteine S-nitrosylation sites in proteins. *PeerJ*, *1*, e171.

[16] Berman, H. M., *et al*. (2002). The protein data bank. *Acta Crystallogr D Biol Crystallogr*, *58(Pt 6 No 1)*, 899-907.

[17] Durek, P., *et al*. Detection and characterization of 3D-signature phosphorylation site motifs and their contribution towards improved phosphorylation site prediction in proteins. (2009). *BMC Bioinformatics*, 10.

[18] Karabulut, N. P., & Frishman, D. (2015). Tissue-specific sequence and structural environments of lysine acetylation sites. *J Struct Biol*, *191(1)*, 39-48.

[19] Marino, S. M., & Gladyshev, V. N. (2010). Structural analysis of cysteine S-nitrosylation: A modified acid-based motif and the emerging role of trans-nitrosylation. *J Mol Biol*, *395(4)*, 844-59.

[20] Lee, T. Y., *et al*. (2012). dbSNO: A database of cysteine S-nitrosylation. *Bioinformatics*, *28(17)*, 2293-5.

[21] (2014). Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res*, *42(Database issue)*, D191-8.

[22] Su, M. G., & Lee, T. Y. (2013). Incorporating substrate sequence motifs and spatial amino acid composition to identify kinase-specific phosphorylation sites on protein three-dimensional structures. *BMC Bioinformatics*, *14 Suppl 16*, S2.

[23] Lin, C. K., & Chen, C. Y. (2013). PiDNA: Predicting protein-DNA interactions with structural models. *Nucleic Acids Res*, *41(Web Server issue)*, W523-30.

[24] Chien, Y. T., & Huang, S. W. (2012). Accurate prediction of protein catalytic residues by side chain orientation and residue contact density. *PLoS One*, *7(10)*, e47951.

Tzong-Yi Lee received the B.S. degree in computer science and information engineering from National Central University (NCU), Taiwan and the Ph.D. degree in bioinformatics from National Chiao Tung University (NCTU), Taiwan, in 2002 and 2008, respectively. From 2009 to 2017, Dr. Lee was an assistant professor, associate professor (2012), and professor (2015) with the Department of Computer Science and Engineering, Yuan Ze University, Taiwan. Currently he is an associate professor in the School of Life and Health Sciences, the Chinese University of Hong Kong, Shenzhen, China. Dr. Lee's research interests include bioinformatics, genomics and proteomics, network biology, data mining in Omics science, shallow and deep learning, database design and software development.

Justin Bo-Kai Hsu obtained his Ph.D. degree at Institute of Bioinformatics and Systems Biology of the National Chiao Tung University. From 2013 to 2017, he worked for the biotechnology company YourGene Bioscience as Taiwan BioBank Project Manager. Because of the project, he is familiar in analyzing and integrating NGS data from various sequencing platforms (Illumina and Ion proton). In 2017, he also worked as an adjunct assistant professor in Yuan Ze University and gave a lecture of “genomics and proteomics” for graduate students. Now he is a researcher in the Department of Medical Research of Taipei Medical University Hospital, a major focus recently is combining genomic data with medical imaging features to improve the prediction of glioma patient survival and to identify prognostic genes which might be possible to refine the treatment.