

Review

Hypervalent Nonbonded Interactions of a Divalent Sulfur Atom. Implications in Protein Architecture and the Functions

Michio Iwaoka * and Noriyoshi Isozumi

Department of Chemistry, School of Science, Tokai University, Kitakaname, Hiratsuka-shi, Kanagawa 259-1292, Japan

* Author to whom correspondence should be addressed; E-Mail: miwaoka@tokai.ac.jp; Tel.: +81-463-58-1211; Fax: +81-463-50-2094.

Received: 21 May 2012; in revised form: 6 June 2012 / Accepted: 8 June 2012 /

Published: 13 June 2012

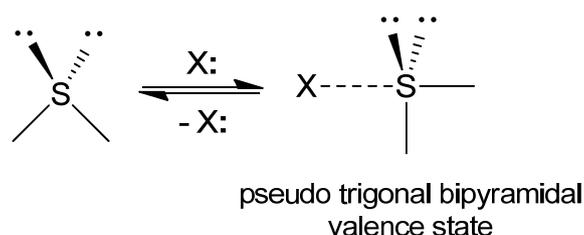
Abstract: In organic molecules a divalent sulfur atom sometimes adopts weak coordination to a proximate heteroatom (X). Such hypervalent nonbonded S···X interactions can control the molecular structure and chemical reactivity of organic molecules, as well as their assembly and packing in the solid state. In the last decade, similar hypervalent interactions have been demonstrated by statistical database analysis to be present in protein structures. In this review, weak interactions between a divalent sulfur atom and an oxygen or nitrogen atom in proteins are highlighted with several examples. S···O interactions in proteins showed obviously different structural features from those in organic molecules (*i.e.*, $\pi_{\text{O}} \rightarrow \sigma_{\text{S}}^*$ versus $n_{\text{O}} \rightarrow \sigma_{\text{S}}^*$ directionality). The difference was ascribed to the HOMO of the amide group, which expands in the vertical direction (π_{O}) rather than in the plane (n_{O}). S···X interactions in four model proteins, phospholipase A₂ (PLA₂), ribonuclease A (RNase A), insulin, and lysozyme, have also been analyzed. The results suggested that S···X interactions would be important factors that control not only the three-dimensional structure of proteins but also their functions to some extent. Thus, S···X interactions will be useful tools for protein engineering and the ligand design.

Keywords: Protein Data Bank; chalcogen bonds; σ -hole bonds; molecular assembly; protein engineering; drug design

1. Introduction

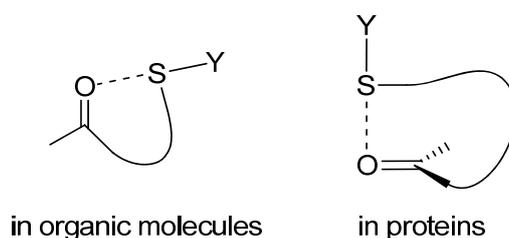
A sulfur atom is usually present in a divalent state in organic molecules. However, it sometimes adopts a weak coordination to a proximate heteroatom (X), such as O, N, and S, in solution as well as in the solid state [1–4]. Such hypervalent weak atomic interactions, so-called nonbonded $S\cdots X$ interactions, are of significant interest because they can control not only the molecular assembly of organic sulfur compounds in advanced materials with unique conductivity or optical properties [5,6] but also the structure and reactivity of isolated molecules [7–9]. The hypervalent (or out-of-octet) nature of a divalent sulfur atom is physic-chemically explained on the basis of the observed apparent directionality by coordination of the lone pair of a ligand (X) to the sulfur center to obtain the pseudo trigonal bipyramidal valence state (Figure 1). Apart from the proposed orbital interaction [10], importance of other factors, such as electrostatic interaction and electron correlation, has also been pointed out for formation of the weak nonbonded interaction [7,11]. Recently, variation of $S\cdots X$ interactions is extending to two-center-three-electron $S\cdots N^+$ [12,13], bifurcated $SS\cdots O$ [14], and $S\cdots CH_2$ (carbene) [15] interactions.

Figure 1. Hypervalent nonbonded $S\cdots X$ interactions of a divalent sulfur atom.



On the other hand, the sulfur atoms of cystine (SSC-type) and methionine (CSC-type) residues in proteins are usually considered as hydrophobic groups, forming no specific nonbonded interaction. However, the recent statistical analyses using the structure database have clearly demonstrated that similar hypervalent $S\cdots X$ interactions to those characterized in organic molecules are widely present in protein structures [16–18]. Interestingly, the directionality of the $S\cdots O$ interactions found in proteins is obviously different from that observed in organic sulfur compounds (*i.e.*, $n_O \rightarrow \sigma_S^*$ versus $\pi_O \rightarrow \sigma_S^*$ as shown in Figure 2) [19], although in both cases the S atom adopts electron coordination in the backside of the S–Y bond (*i.e.*, in the σ hole). Importance of this non-classical atomic interaction of proteins in the stability, function, and evolution has also been pursued by us [18–20] and other groups [16,21–23].

Figure 2. Nonbonded $S\cdots X$ interactions in organic molecules and proteins.



In this review, structural features and functional aspects of weak nonbonded S \cdots X interactions in proteins are highlighted with several examples. In the next two sections, S \cdots X interactions in organic molecules and in proteins are overviewed, respectively. Structural features of the S \cdots X interactions in proteins are discussed in comparison with those observed in organic molecules. In the subsequent sections, the S \cdots X interactions in proteins are implicated in protein architecture and the functions based on the results from structure database analyses using a large set of heterogeneous protein structures and rather small sets of structures of four model proteins, *i.e.*, phospholipase A₂ (PLA₂), ribonuclease A (RNase A), insulin, and lysozyme. It is strongly suggested that the S \cdots X interactions are important factors that control protein structures and functions to some extent. Perspectives of the research on such weak nonbonded interactions are given in the Conclusions. Throughout this review, the term S \cdots X interactions is used for weak hypervalent atomic interactions, although other terminologies, such as σ -hole bonds [24] and chalcogen bonds [25], are frequently used in the recent literature.

2. S \cdots X Interactions in Organic Compounds

2.1. Database Analysis

Attractive interactions between a divalent sulfur atom (S) and nearby heteroatoms (X) have been well recognized in small organic compounds. In 1977, Rosenfield *et al.* [2] surveyed close S \cdots X contacts in organic and inorganic crystals using the Cambridge Crystallographic Database [26] and found an obvious directional preference of X with respect to S, as shown in Figure 1. The directionality was reasonably explained by the presence of specific nonbonded S \cdots X interactions. On the other hand, the directional preference of the S \cdots O interactions with respect to O was studied in detail by Kucsman and Kapovitz [27]. For intramolecular 1,4- and 1,5-type S \cdots O=C interactions, the S atom tended to lie in the direction of the O lone pairs (*i.e.*, the n_O direction) rather than the π electrons (*i.e.*, the π_O direction). These statistical analyses using the structure database demonstrated the importance of the $n_O \rightarrow \sigma_S^*$ orbital interaction for formation of S \cdots O interactions in organic molecules.

A similar database analysis for analogous nonbonded S \cdots S interactions was carried out by Row and Parthasarathy in 1981 [3]. They reported that S \cdots S interactions in organic crystals would be stabilized by the orbital interaction between the lone pair of one sulfur atom (n_S) and the anti-bonding orbital of the other sulfur atom (σ_S^*) as observed in the S \cdots O interactions. Desiraju and Nalini [4] obtained a similar donor-acceptor interaction mechanism of S \cdots S interactions. The presence of S \cdots C(π) interactions in organic crystals was also suggested by Zauhar *et al.* [28].

2.2. Energetic Elements of S \cdots X Interactions

Database analyses for various types of nonbonded S \cdots X (X = O, S, *etc.*) interactions clearly demonstrated that specific directional preferences are present for the S \cdots X interactions in organic crystals. The directionality strongly indicated the importance of the orbital interaction between the interacting atoms for the stability. For instance, S \cdots O and S \cdots S interactions would be stabilized by the $n_O \rightarrow \sigma_S^*$ and $n_S \rightarrow \sigma_S^*$ orbital interactions, respectively. However, the nonbonded distances between the two interacting atoms are sometimes only marginally shorter than the sum of the van der Waals

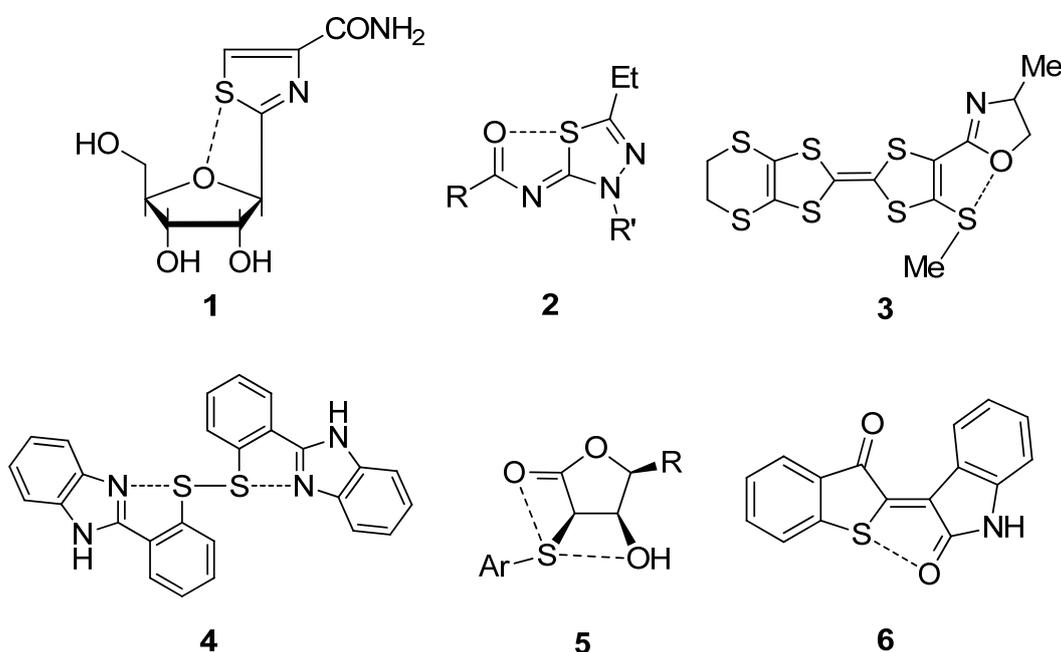
radii [29], and, in such cases, the directionality becomes rather subtle. Therefore, there should be other electronic factors that are contributing to formation of the S \cdots X interactions.

The electrostatic nature of the 1,4-type S \cdots O interactions between a positively charged S atom and a negatively charged ethereal O atom was suggested by Burling and Goldstein [7] based on the substituent effects and the theoretical calculations. Dahaoui *et al.* [11] pointed out importance of van der Waals forces (or electron correlation effects) for the S \cdots S interactions. Thus, several energetic elements, such as orbital interaction, electrostatic interaction, and electron correlation, must be considered for fully understanding S \cdots X interactions. Recent sophisticated theoretical analyses applying the atoms-in-molecules (AIM) method by Nakanishi [30] and the symmetry-adapted perturbation theory (SAPT) by Scheiner [31,32] provided more exact description of S \cdots X interactions in terms of the total electron energy density and Laplacian of the electron density at the bond critical points of AIM as well as the electrostatic, induction, and dispersion components of SAPT.

2.3. Examples

Intramolecular nonbonded S \cdots X interactions have been extensively studied for some organic sulfur compounds in relation to the biological activities as well as the physical properties as advanced materials. Examples are shown in Figure 3.

Figure 3. Examples of S \cdots O interactions in organic molecules.



Burling and Goldstein [7] demonstrated the importance of an intramolecular 1,4-type S \cdots O interaction of thiazole nucleoside analogues (**1**) for their antitumor activity. Nagao [8,33] reported that 1,5-type S \cdots O interaction plays important roles in the antagonism of (acylimino)thiadiazoline derivatives (**2**) towards an angiotensin II receptor. Similar S \cdots O and S \cdots N interactions are responsible for the molecular structures and functions of TTF-oxazoline derivatives (**3**) [6], bis[2-(1*H*-benzimidazol-2-yl)phenyl]disulfide (**4**) [34], β -hydroxy- α -sulfenyl- γ -butyrolactones (**5**) [35], and thioindirubin (**6**) [36].

3. S··X Interactions in Proteins

Weak nonbonded interactions are important physicochemical forces that control the structure of proteins [37]. Ionic interaction, hydrogen bond, van der Waals forces, and hydrophobic interaction are mainly considered as this class of interactions, but some novel interaction patterns, such as C–H···O hydrogen bond [38–40], cation- π interaction [41–43], and CH/ π hydrogen bond [44–46], have recently been characterized in folded protein structures. Importance of these new interactions for the stability and functions of proteins has also been pointed out. The S··X interactions can be another member of such non-classical interactions.

The SSC and CSC groups involved in cystine and methionine residues, however, were usually considered just as hydrophobic moieties in folded protein structures until recently, except for S··C(π) interactions [47] and weak NH··S and OH··S hydrogen bonds [48]. The S··C(π) interactions in proteins were first pointed out by Morgan and co-workers [49], who analyzed the close atomic contact between S and a π -plane in eight protein structures and found that the S atoms have a propensity to come over the π -plane. On the other hand, Reid *et al.* [50] suggested by using a larger set of protein structures that the close S··C(π) contact in proteins can also be explained by CH··S interactions because the S atoms access to the π -plane from the side rather than the top. According to several experimental and theoretical studies having been reported to date [22,28,51,52], however, the nature of S··C(π) interactions in proteins would be well rationalized by the interaction between the aromatic π electrons and the S atom [22,47]. Meanwhile, NH··S and OH··S hydrogen bonds were suggested to play some roles in particular proteins [48], but the interactions were rarely found in protein structures. The S atoms of cystine and methionine would have only a weak character of a hydrogen-bond acceptor.

3.1. Database Analysis

Nonbonded S··X interactions in proteins have recently been pursued by several research groups [16–18,21–23]. The stereochemistry of the nonbonded S(CSC)··O interactions for methionine residues was first analyzed by Carugo [21] using a small set of protein structures. Although no strong directional preference was observed, the result suggested that the S··O interactions in proteins would have either a very weak or physicochemically different character from those observed in small molecules. On the other hand, Iwaoka *et al.* [17] thoroughly surveyed close S··X (X = O, N, S, C, *etc.*) atomic contacts involved in 604 high-resolution (≤ 2.0 Å) heterogeneous X-ray structures selected from the Protein Data Bank [53]. Statistical analyses of the relative nonbonded S··X distance ($d = r_{S\cdots X} - vdw_S - vdw_X$), the directionality around the S and X atoms, and the location along the amino acid sequence revealed distinct structural features of the S··X interactions. In case of the most frequent S··O interactions, both SSC and CSC S atoms tend to approach a main-chain O atom perpendicularly to the amide plane (the π_O direction), and the O atom tends to approach the S atom from the backside of the S–S or S–C covalent bonds (the σ_S^* direction). Similar directionalities of the S(CSC)··O interactions were also reported by Pal and Chakrabarti [16]. The structural propensity observed is in striking contrast to the S··O interactions in organic crystals, in which the n_O orbital, not π_O , is usually used to form the interaction.

Table 1 shows numbers of S(SSC)⋯X contacts in selected protein structures [17]. For short contacts ($d \leq 0.0 \text{ \AA}$), a probability of S⋯O contacts increases significantly, suggesting the presence of specific S⋯O interactions in proteins. According to the statistical analysis for the obtained data, four types of nonbonded S⋯X interactions have been clearly characterized, *i.e.*, S–S⋯O=C, C–S⋯O=C, C–S⋯N, and S–S⋯S–S interactions [18].

Table 1. Numbers of S(SSC)⋯X contacts in selected proteins [17]^a.

X =	$d^b \leq 0.0 \text{ \AA}$	$0.0 < d \leq 0.5 \text{ \AA}$
O	100	664
N	33	359
S	15	68
C	134	1478
Others	2	0
Total	284	2569

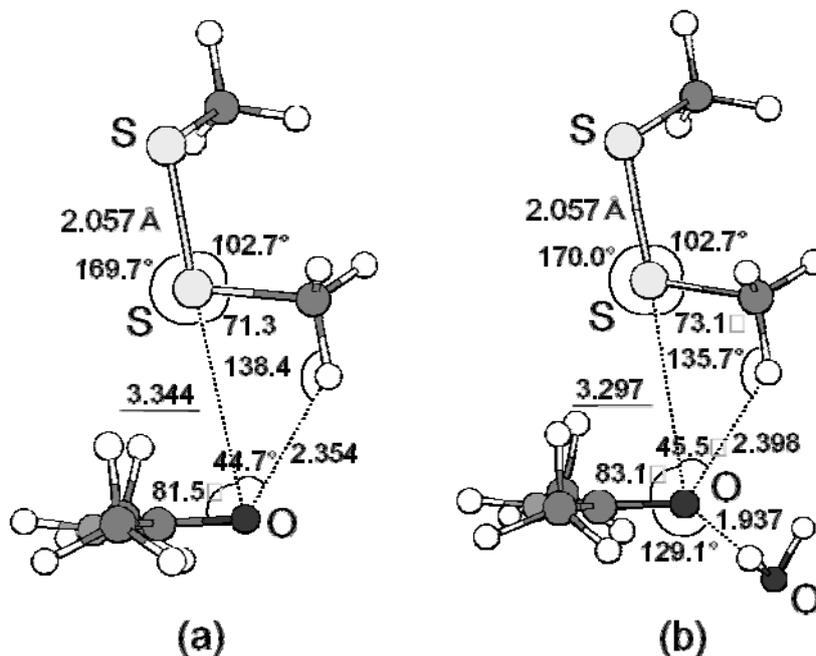
^a 604 heterogeneous proteins with high resolution ($\leq 2 \text{ \AA}$) were selected. The total number of Cys residues was 790; ^b A relative nonbonded S⋯X distance defined as $d = r_{S\cdots X} - vdw_S - vdw_X$.

A majority of close S⋯O contacts for an SSC group in proteins is assigned to S–S⋯O=C interactions, which can be characterized by the linear S–S⋯O atomic alignment and the vertical access of the S atom to the carbonyl plane as shown in Figure 2. According to the directionality, a significant contribution from the $\pi_O \rightarrow \sigma_S^*$ orbital interaction is obvious. The S–S⋯O=C interactions are most frequently observed in helices, suggesting that the S⋯O interactions would support the stability. The C–S⋯O=C interaction formed between a methionine side-chain and a main-chain peptide group has a similar character to the S–S⋯O=C interaction, but the strength of the interaction is weaker with attenuated directionality. Similarly, the C–S⋯N interaction between a methionine side-chain and a main-chain peptide group can also be characterized by the linear C–S⋯N atomic alignment and the vertical access of the S atom to the amide plane, suggesting a contribution from the $\pi_N \rightarrow \sigma_S^*$ orbital interaction. On the other hand, most of close S⋯S contacts in proteins can be assigned to S–S⋯S–S interaction, which would be stabilized by the $n_S \rightarrow \sigma_S^*$ orbital interaction in a similar manner to that observed in organic crystals [3].

3.2. Energetic Elements of S⋯X Interactions

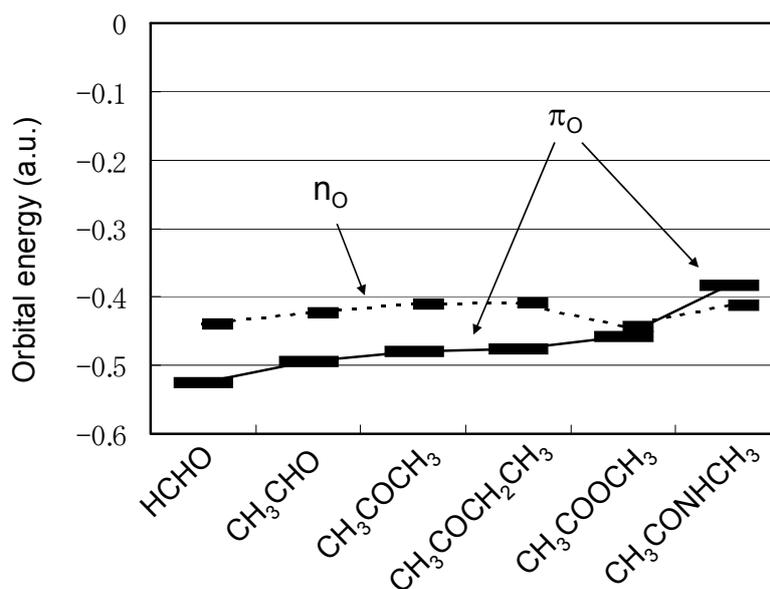
Ab initio calculation was carried out for the model complexes ($\text{CH}_3\text{SSCH}_3 + \text{CH}_3\text{CONHCH}_3$ and $\text{CH}_3\text{SCH}_3 + \text{CH}_3\text{CONHCH}_3$) to investigate the nature of the S⋯X interactions observed in proteins [17,18]. The calculation using the Møller-Plesset method (MP2) [54] suggested the importance of the dispersion force or the electron correlation for the stability of the S⋯X interactions. The S–S⋯O=C interaction was estimated to be as strong as 3.2 kcal/mol at MP2/6-31G(d) and would be predominantly stabilized by electron correlation with a significant contribution from the $\pi_O \rightarrow \sigma_S^*$ orbital interaction (Figure 4). The presence of a CH⋯O hydrogen bond and an additional hydrogen bond (such as NH⋯O or OH⋯O) at the main-chain carbonyl O atom would stabilize the complex cooperatively with the S⋯O interaction. On the other hand, the C–S⋯O=C and C–S⋯N interactions should be weaker than the S–S⋯O=C interaction with strengths up to 2.5 and 2.9 kcal/mol, respectively.

Figure 4. Optimized structures of the complexes between CH_3SSCH_3 and $\text{CH}_3\text{CONHCH}_3$ without H_2O (a) or with H_2O (b) at MP2/6-31G(d) [17].



The reason for the observed discrepancy in the directionality between the $\text{S}\cdots\text{O}$ interactions in proteins and organic molecules can be explained on the basis of the HOMO levels of various carbonyl compounds [19]. As graphically shown in Figure 5, for most of the carbonyl compounds the HOMO is assigned to the oxygen lone pair (n_{O}) lying in the carbonyl plane rather than the π orbital expanding perpendicular to the plane.

Figure 5. Energy levels of n_{O} and π_{O} orbitals for various carbonyl compounds calculated at MP2/6-31G(d) [19].



However, the π_0 level of amide $\text{CH}_3\text{CONHCH}_3$ is remarkably raised compared with other carbonyl compounds, while the n_0 level remains almost unchanged. The elevation of the π_0 orbital would be due to the conjugation between the N lone pair and the carbonyl group. Thus, inversion of the energy levels of n_0 and π_0 would be responsible for the observed directional preferences of the $\text{S}\cdots\text{O}$ interactions in proteins.

3.3. Examples

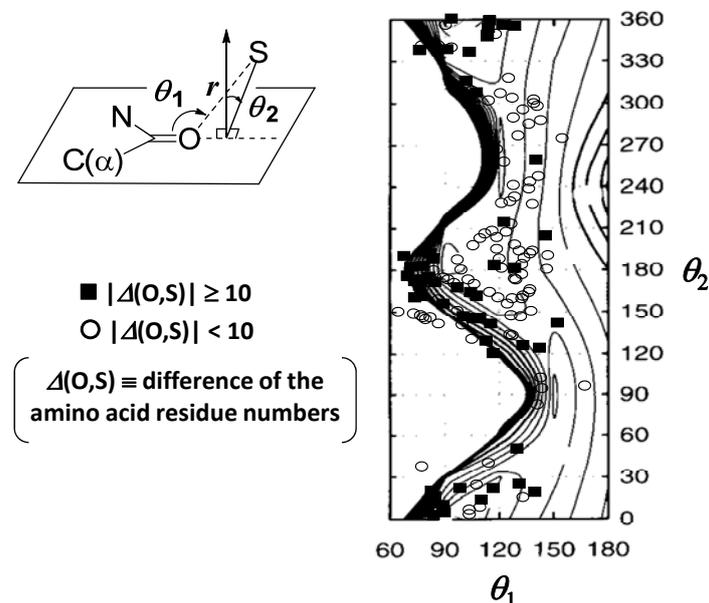
The importance of $\text{S}\cdots\text{O}$ interactions in the enzymatic function of proteins has been pointed out for some particular cases. Taylor and Markham [55] suggested that the electrostatic $\text{S}\cdots\text{O}$ interaction between the S atom of a methionine substrate and the carboxylate O atom of Asp118 plays a major role in the enzymatic activity of S-adenosylmethionine synthetase. Brandt *et al.* [56] reported that the cleavage of a disulfide bond in the extracellular region of the G-protein receptors, which is an important process for the receptor activation, is catalyzed by the $\text{S}\cdots\text{O}$ interaction between Cys121 and the carboxylic group of Asp288. The existence of $\text{S}\cdots\text{O}$ interactions was suggested for the complexes between *N*-acetylglucosamine-thiazoline and β -hexosaminidase and between benzophenone and porcine odorant-binding protein [33]. Importance of the possible $\text{S}\cdots\text{O}$ interaction between the Met10 and Thr317 residues of adenylosuccinate lyase for the stability of the protein structure was also suggested using the M10L mutant [57]. More recently, the $\text{S}\cdots\text{N}$ interactions between an OSCN⁻ ligand and the imidazole ring of His109 in lactoperoxidase [58] and between the sulfenic acid form of Cys50 and the imidazole ring of His42 in peroxiredoxin [59] were suggested to play roles in the functions.

4. Implications of $\text{S}\cdots\text{X}$ Interactions in Protein Architecture

According to the comparison between the results from database analyses and *ab initio* calculation, it was clear that the directional preferences of the $\text{S}\cdots\text{O}$ interactions observed in protein structures are in accord with the profiles of the potential surfaces calculated for the isolated model complexes [18]. An example is shown in Figure 6 for the case of the directionality around the O atom in the $\text{S}-\text{S}\cdots\text{O}=\text{C}$ interactions in proteins. The remarkable agreement strongly suggested that these nonbonded interactions are important determinants for protein architecture. Similar phenomena have been found between the Ramachandran plots of the amino acid residues and the single amino acid potentials (SAAP) in water, which has already been applied to developing a new all-atom force field for molecular simulation of peptide molecules [60,61].

Protein structures are generally considered to be flexible because they are governed only by weak nonbonded interactions. Therefore, coincidence of the statistical conformational preference for the interaction of proteins with the potential surface calculated for the isolated model is to be noticed and would have important implications in protein architecture. The detailed analysis revealed the following features [19]. The linearity of $\text{Y}-\text{S}\cdots\text{O}$ ($\text{Y} = \text{S}$ or C) alignment in both organic molecules and proteins is not affected by the crystal packing force but can be disturbed by the structural constraint between the interacting fragments. On the other hand, the vertical nature of $\text{Y}-\text{S}\cdots\text{O}=\text{C}$ interactions is not affected by the presence of other weak nonbonded interactions in protein structures

Figure 6. The directionality around the O atom in the S–S⋯O=C interactions in proteins superimposed on the potential surface calculated for the complex of between CH₃SSCH₃ and CH₃CONHCH₃ at MP2/6-31G(d). The contour lines are drawn with an interval of 0.25 kcal/mol. This figure was modified from reference [18].



The directionality, however, is easily affected by crystal packing force for organic molecules. Thus, the order of the factors that control molecular structure of organic molecules and proteins in the solid state can be summarized as shown in Figure 7 [19].

Figure 7. The factors that control molecular structure of organic molecules and proteins in the solid state.



The O atom has strong tendency to approach the S atom from the backside of the S–C or S–S bond (in the σ_{S^*} direction), irrespective of the types of carbonyl groups. On the other hand, the S atom tends to approach the O atom either within the carbonyl plane (in the n_{O} direction) or from the vertical direction (in the π_{O} direction). In the case of S⋯O(amide) interactions, the vertical direction is significantly preferred, due probably to elevation of the π_{O} orbital. The linearity of the S⋯O interactions in organic molecules would overcome the crystal packing force, whereas the vertical nature of the S⋯O(amide) interactions may be affected by the packing force. The verticality, however, would survive in protein structures. These structural features will be informative for protein engineering and molecular design of functional organic sulfur compounds.

5. Implications of S⋯X Interactions in Protein Functions

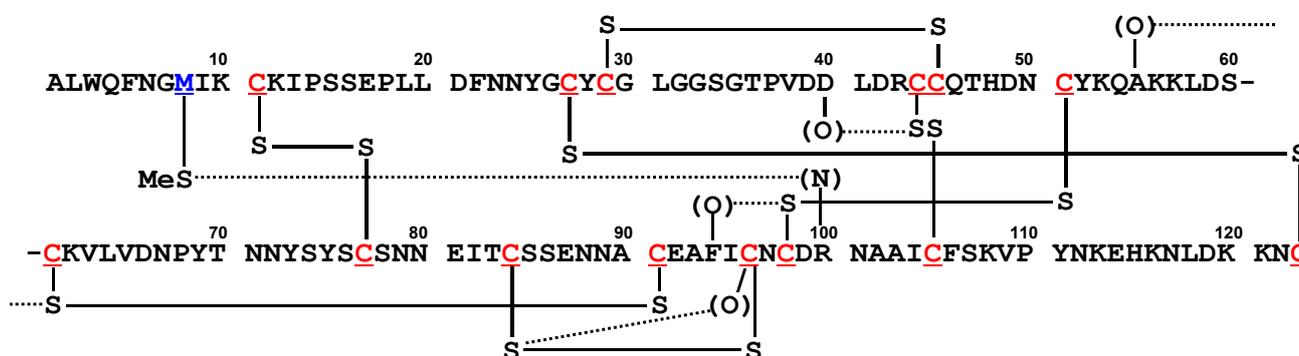
With a success in characterization of S⋯X interactions in proteins, we subsequently sought out particular protein families or domains, for which specific S⋯X interactions are commonly present in a

wide range of the structures registered in the Protein Data Bank [53]. Phospholipase A₂ (PLA₂) was the first example of such proteins [20]. We have thus far found three more examples by applying the same method of the database analysis that was applied for the PLA₂ family. The results are summarized in this section.

5.1. Phospholipase A₂ (PLA₂) [20]

PLA₂ [62,63], a small globular protein consisting of about 130 amino acid residues, is a SS-rich enzyme that catalyzes hydrolysis of the 2-acyl ester bond of phosphoglycerides in the presence of a calcium ion. There are two domain groups in the vertebrate PLA₂ family; PLA₂ and snake PLA₂ (sPLA₂). Comprehensive search for close S··X (X=O, N, and S) contacts in the structures of the PLA₂ domain, which were retrieved from protein data bank, revealed the presence of four common S··O interactions, *i.e.*, S(C44)··O(D40), S(C61)··O(A55), S(C84)··O(C96), and S(C98)··O(F94), and one common S··N interaction, *i.e.*, S(M8)··N(R100), as shown in Figure 8. Most of the S··X interactions were found in the vicinity of the active site and to tolerate the conformational changes caused by binding of the substrate. It was reported in the literature that the enzymatic activity of porcine PLA₂ is decreased in the M8,20L mutant [64]. This would be explained by elimination of the S(M8)··N(R100) interaction.

Figure 8. The amino acid sequence of bovine PLA₂ with notations of common S··X interactions.



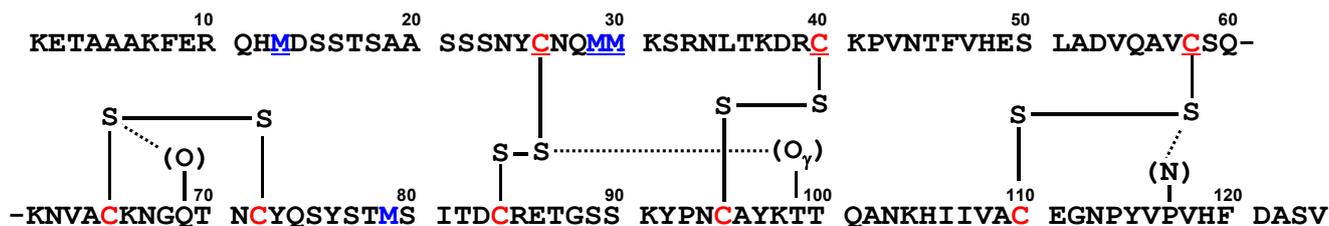
On the other hand, an evolutionary aspect of the S··X interactions was analyzed for the sPLA₂ domain. For this domain group, the phylogenetic dendrogram was already analyzed by Ohno *et al.* [65] using the amino acid sequences of various PLA₂ involved in venom of snakes inhabiting the southern islands of Japan. Mapping the common S··O and S··N interactions observed by the database analysis, we found that most of the S··X interactions make clusters on the dendrogram. The results suggested a possible role of S··X interactions in molecular evolution of proteins.

5.2. Ribonuclease A (RNase A)

RNase A [66] is a typical globular protein of 124 amino acid residues having four SS bonds (C40–C95, C65–C72, C26–C84, and C58–C110) in the native state. Its biological functions as well as the three-dimensional structure and the folding pathways have been extensively studied [67,68]. According to the standardized statistical analysis method [20], close S··X contacts in the structures belonging to the RNase A domain were thoroughly sought. The 43 high-resolution (≤ 2.0 Å) structures

were extracted from the protein data bank, and two common S...O interactions, *i.e.*, S(C26)...O_γ(T99) and S(C65)...O(Q69), and one common S...N interaction, *i.e.*, S(C58)...N(P117), were characterized. The locations along the amino acid sequence are shown in Figure 9. The SS loop of C65–C72 is one of the important sites that fold in the beginning of the oxidative folding [69] and gives significant thermodynamic stability to the native structure [70]. The presence of the S(C65)...O(Q69) interaction in this loop may be responsible in part for the stability.

Figure 9. The amino acid sequence of bovine RNase A with notations of common S...X interactions.

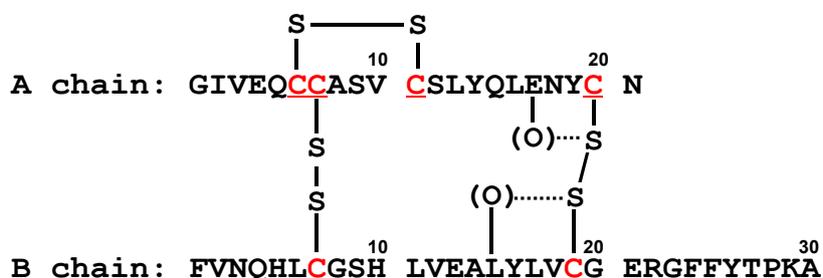


In some RNase A structures complexed with a substrate, close S...X contacts between the S atom of the C65 residue and the substrate were found. Examples are S(C65)...O(ADT) in 8RSA [71], S(C65)...N(PAP) in 1AFK [72], S(C65)...N(ATR) in 1AFL [72], and S(C65)...N(PUA) in 1QHC [73]. These interactions would be cooperatively stabilized by the S(C65)...O(Q69) interaction.

5.3. Insulin

Insulin [74] is a peptide hormone composed of two peptides (A and B chains), which are connected together by two SS bridges. Another SS bond is present in the A chain. For this model, two common S...O interactions, *i.e.*, S(C20)...O(E17) and S(C19)...O(L15), were found within A and B chains, respectively, among the 23 high-resolution (≤ 2.0 Å) structures. Figure 10 shows the locations of the interactions. By the complex formation with a ligand, the common S...O interactions were not disrupted, while new S...O interactions were emerged in the A chain; S(C6)...O(I2), S(C7)...O(V3), and S(C11)...O(Q5). These allosteric effects may possibly contribute to the function of insulin.

Figure 10. The amino acid sequence of bovine insulin with notations of common S...X interactions.



5.4. Lysozyme

Lysozyme [75] consists of 129 or 130 amino acid residues with four SS bonds in the native state. A large number of the structures with various ligand have been registered in the Protein Data Bank. We selected 28 high-resolution (≤ 1.5 Å) structures from the lysozyme domain and analyzed close

Sulfur-containing functional groups of cystine (an SSC group) and methionine (a CSC group) were previously considered to be just hydrophobic moieties in protein structures, but they are indeed able to form specific nonbonded interactions with nearby polar non-hydrogen atoms ($X = O, N, S$) in folded proteins. A unique directionality of the $S \cdots X$ interactions (see Figure 2) would be largely controlled by the orbital interaction between the interacting atoms, while a significant contribution from the electron correlation seems to be important for the stability according to the *ab initio* calculation. For four particular proteins, *i.e.*, PLA₂, RNase A, insulin, and lysozyme, unique $S \cdots X$ interactions have been characterized, and some were suggested to play roles in the stability of the native structures and the functions to some extent.

Finally, the statistical analyses using the both protein and organic molecule structure databases demonstrated that the order of the strength of the factors that control molecular structures in the solid state can be expressed as shown in Figure 7. The order will be useful for versatile fields of chemistry, such as development of advanced materials built by molecular assembly, molecular design of sulfur-containing ligands or medicines, and protein engineering. The statistical method employed for the $S \cdots X$ interactions, *i.e.*, integration of the database analyses and theoretical calculation, will be useful for characterization of other weak nonbonded interactions hidden in molecules as well as protein structures.

Supplementary Materials

Supplementary materials can be accessed at: <http://www.mdpi.com/1420-3049/17/6/7266/s1>.

Acknowledgments

This work was supported by Grants-in-aid for Scientific Research on Priority Areas (C) “Genomic Information Science” (Nos. 15014211 and 16014223) and Scientific Research (B) (No. 16350092) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Conflict of Interest

The authors declare that they have no personal or financial conflict of interest.

References and Notes

1. Ciuffarin, E.; Guaraldi, G. Nucleophilic substitution at dicoordinated sulfur. Effect of the leaving group on the reaction between triphenylmethyl sulfenyl derivatives and *n*-butylamine. *J. Org. Chem.* **1970**, *35*, 2006–2010.
2. Rosenfield, R.E., Jr.; Parthasarathy, R.; Dunitz, J.D. Directional preferences of nonbonded atomic contacts with divalent sulfur. 1. Electrophiles and nucleophiles. *J. Am. Chem. Soc.* **1977**, *99*, 4860–4862.
3. Row, T.N.G.; Parthasarathy, R. Directional preferences of nonbonded atomic contacts with divalent sulfur in terms of its orbital orientations. 2. $S \cdots S$ interactions and nonspherical shape of sulfur in crystals. *J. Am. Chem. Soc.* **1981**, *103*, 477–479.

- Desiraju, G.R.; Nalini, V. Database analysis of crystal-structure-determining interactions involving sulphur: Implications for the design of organic metals. *J. Mater. Chem.* **1991**, *1*, 201–203.
- Turbiez, M.; Frère, P.; Allain, M.; Videlot, C.; Ackermann, J.; Roncali, J. Design of organic semiconductors: Tuning the electronic properties of π -conjugated oligothiophenes with the 3,4-ethylenedioxythiophene (EDOT) building block. *Chem. Eur. J.* **2005**, *11*, 3742–3752.
- Réthoré, C.; Madalan, A.; Fourmigué, M.; Canadell, E.; Lopes, E.B.; Almeida, M.; Clérac, R.; Avarvari, N. O \cdots S vs. N \cdots S intramolecular nonbonded interactions in neutral and radical cation salts of TTF-oxazoline derivatives: synthesis, theoretical investigations, crystalline structures, and physical properties. *New J. Chem.* **2007**, *31*, 1468–1483.
- Burling, F.T.; Goldstein, B.M. Computational studies of nonbonded sulfur-oxygen and selenium-oxygen interactions in the thiazole and selenazole nucleosides. *J. Am. Chem. Soc.* **1992**, *114*, 2313–2320.
- Nagao, Y.; Hirata, T.; Goto, S.; Sano, S.; Kakehi, A.; Iizuka, K.; Shiro, M. Intramolecular nonbonded S \cdots O interaction recognized in (acylimino)thiadiazoline derivatives as angiotensin II receptor antagonists and related compounds. *J. Am. Chem. Soc.* **1998**, *120*, 3104–3110.
- Nakanishi, W.; Nakamoto, T.; Hayashi, S.; Sasamori, T.; Tokitoh, N. Atoms-in-molecules analysis of extended hypervalent five-center, six-electron (5c-6e) C₂Z₂O interactions at the 1,8,9-positions of anthraquinone and 9-methoxyanthracene systems. *Chem. Eur. J.* **2007**, *13*, 255–268.
- Bleiholder, C.; Werz, D.B.; Köppel, H.; Gleiter, R. Theoretical investigations on chalcogen-chalcogen interactions: What makes these nonbonded interactions bonding? *J. Am. Chem. Soc.* **2006**, *128*, 2666–2674.
- Dahaoui, S.; Pichon-Pesme, V.; Howard, J.A.K.; Lecomte, C. CCD charge density study on crystals with large unit cell parameters: The case of hexagonal L-cystine. *J. Phys. Chem. A* **1999**, *103*, 6240–6250.
- Nauser, T.; Jacoby, M.; Koppenol, W.H.; Squier, T.C.; Schöneich, C. Calmodulin methionine residues are targets for one-electron oxidation by hydroxyl radicals: Formation of S \cdot -N three-electron bonded radical complexes. *Chem. Commun.* **2005**, 587–589.
- Mishra, B.; Sharma, A.; Naumov, S.; Priyadarsini, K.I. Novel reactions of one-electron oxidized radicals of selenomethionine in comparison with methionine. *J. Phys. Chem. B* **2009**, *113*, 7709–7715.
- Zhang, Y.; Wang, W.Z. The bifurcate chalcogen bond: Some theoretical observations. *J. Mol. Struct. (Theochem)* **2009**, *916*, 135–138.
- Zhao, Q.; Feng, D.C.; Sun, Y.M.; Hao, J.C.; Cai, Z.T. Theoretical investigations on the weak nonbonded C=S \cdots CH₂ interactions: Chalcogen-bonded complexes with singlet carbene as an electron donor. *Int. J. Quantum Chem.* **2011**, *111*, 3881–3887.
- Pal, D.; Chakrabarti, P. Non-hydrogen bond interactions involving the methionine sulfur atom. *J. Biomol. Struct. Dynam.* **2001**, *19*, 115–128.
- Iwaoka, M.; Takemoto, S.; Okada, M.; Tomoda, S. Statistical characterization of nonbonded S \cdots O interactions in proteins. *Chem. Lett.* **2001**, 132–133.
- Iwaoka, M.; Takemoto, S.; Okada, M.; Tomoda, S. Weak nonbonded S \cdots X (X=O, N, and S) interactions in proteins. Statistical and theoretical studies. *Bull. Chem. Soc. Jpn.* **2002**, *75*, 1611–1625.

19. Iwaoka, M.; Takemoto, S.; Tomoda, S. Statistical and theoretical investigations on the directionality of nonbonded S \cdots O interactions. Implications for molecular design and protein engineering. *J. Am. Chem. Soc.* **2002**, *124*, 10613–10620.
20. Iwaoka, M.; Isozumi, N. Possible roles of S \cdots O and S \cdots N interactions in the functions and evolution of phospholipase A₂. *Biophysics* **2006**, *2*, 23–34.
21. Carugo, O. Stereochemistry of the interaction between methionine sulfur and the protein core. *Biol. Chem.* **1999**, *380*, 495–498.
22. Bhattacharyya, R.; Pal, D.; Chakrabarti, P. Disulfide bonds, their stereospecific environment and conservation in protein structures. *Protein Eng. Des. Sel.* **2004**, *17*, 795–808.
23. Lu, J.M.; Lu, Y.X.; Yang, S.B.; Zhu, W.L. Theoretical and crystallographic data investigations of noncovalent S \cdots O interactions. *Struct. Chem.* **2011**, *22*, 757–763.
24. Murray, J.S.; Lane, P.; Politzer, P. Simultaneous σ -hole and hydrogen bonding by sulfur- and selenium-containing heterocycles. *Int. J. Quantum Chem.* **2008**, *108*, 2770–2781.
25. Wang, W.Z.; Ji, B.M.; Zhang, Y. Chalcogen bond: A sister noncovalent bond to halogen bond. *J. Phys. Chem. A* **2008**, *113*, 8132–8135.
26. Allen, F.H. The Cambridge Structural Database: A quarter of a million crystal structures and rising. *Acta Crystallogr.* **2003**, *B58*, 380–388.
27. Kucsman, A.; Kapovits, I. Nonbonded Sulfur-Oxygen Interaction in Organic Sulfur Compounds. In *Organic Sulfur Chemistry*; Bernardi, F., Csizmadia, I.G., Mangini, A., Eds.; Elsevier: Amsterdam, The Netherlands, 1985; pp. 191–245.
28. Zauhar, R.J.; Colbert, C.L.; Morgan, R.S.; Welsh, W.J. Evidence for a strong sulfur-aromatic interaction derived from crystallographic data. *Biopolymers* **2000**, *53*, 233–248.
29. Cozzolino, A.F.; Vargas-Baca, I.; Mansour, S.; Mahmoudkhani, A.H. The nature of the supramolecular association of 1,2,5-chalcogenadiazoles. *J. Am. Chem. Soc.* **2005**, *127*, 3184–3190.
30. Nakanishi, W.; Hayashi, S.; Narahara, K. Atoms-in-molecules dual parameter analysis of weak to strong interactions: Behaviors of electronic energy densities versus Laplacian of electron densities at bond critical points. *J. Phys. Chem. A* **2008**, *112*, 13593–13599.
31. Scheiner, S. On the properties of X \cdots N noncovalent interactions for first-, second-, and third-row X atoms. *J. Chem. Phys.* **2011**, *134*, 164313.
32. Adhikari, U.; Scheiner, S. The S \cdots N noncovalent interaction: Comparison with hydrogen and halogen bonds. *Chem. Phys. Lett.* **2011**, *514*, 36–39.
33. Nagao, Y.; Honjo, T.; Iimori, H.; Goto, S.; Sano, S.; Shiro, M.; Yamaguchi, K.; Sei, Y. Intramolecular nonbonded S \cdots O interaction in acetazolamide and thiadiazolinethione molecules in their dimeric crystalline structures and complex crystalline structures with enzymes. *Tetrahedron Lett.* **2004**, *45*, 8757–8761.
34. Esparza-Ruiz, A.; Peña-Hueso, A.; Hernández-Díaz, J.; Flores-Parra, A.; Contreras, R. Effect of weak sulfur interactions and hydrogen bonds in the folded or unfolded conformation of bis 2-(1H-benzimidazol-2-yl)phenyl disulfide derivatives. *Cryst. Growth Des.* **2007**, *7*, 2031–2040.
35. González, F.V.; Jain, A.; Rodríguez, S.; Sáez, J.A.; Vicent, C.; Peris, G. Stereoisomerization of β -hydroxy- α -sulfenyl- γ -butyrolactones controlled by two concomitant 1,4-Type nonbonded sulfur–oxygen interactions as analyzed by X-ray crystallography. *J. Org. Chem.* **2010**, *75*, 5888–5894.

36. Shishkin, O.V.; Omelchenko, I.V.; Kalyuzhny, A.L.; Paponov, B.V. Intramolecular S \cdots O chalcogen bond in thioindirubin. *Struct. Chem.* **2010**, *21*, 1005–1011.
37. Dobson, C.M.; Sali, A.; Karplus, M. Protein folding: A perspective from theory and experiment. *Angew. Chem. Int. Ed. Engl.* **1998**, *37*, 869–893.
38. Derewenda, Z.S.; Lee, L.; Derewenda, U. The occurrence of C–H \cdots O hydrogen bonds in proteins. *J. Mol. Biol.* **1995**, *252*, 248–262.
39. Bella, J.; Berman, H.M. Crystallographic evidence for C $^{\alpha}$ –H \cdots O=C hydrogen bonds in a collagen triple helix. *J. Mol. Biol.* **1996**, *264*, 734–742.
40. Vargas, R.; Garza, J.; Dixon, D.A.; Hay, B.P. How strong is the C $^{\alpha}$ –H \cdots O=C hydrogen bond? *J. Am. Chem. Soc.* **2000**, *122*, 4750–4755.
41. Flocco, M.M.; Mowbray, S.L. Planar stacking interactions of arginine and aromatic side-chains in proteins. *J. Mol. Biol.* **1994**, *235*, 709–717.
42. Gallivan, J.P.; Dougherty, D.A. Cation- π interactions in structural biology. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9459–9464.
43. Gallivan, J.P.; Dougherty, D.A. A computational study of cation- π interactions vs salt bridges in aqueous media: Implications for protein engineering. *J. Am. Chem. Soc.* **2000**, *122*, 870–874.
44. Umezawa, Y.; Nishio, M. CH/ π interactions as demonstrated in the crystal structure of guanine-nucleotide binding proteins, Src homology-2 domains and human growth hormone in complex with their specific ligands. *Bioorg. Med. Chem.* **1998**, *6*, 493–504.
45. Takahashi, O.; Kohno, Y.; Nishio, M. Relevance of weak hydrogen bonds in the conformation of organic compounds and bioconjugates: Evidence from recent experimental data and high-level ab initio MO calculations. *Chem. Rev.* **2010**, *110*, 6049–6076.
46. Brandl, M.; Weiss, M.S.; Jabs, A.; Sühnel, J.; Hilgenfeld, R. C–H \cdots π -interactions in proteins. *J. Mol. Biol.* **2001**, *307*, 357–377.
47. Chakrabarti, P.; Bhattacharyya, R. Geometry of nonbonded interactions involving planar groups in proteins. *Prog. Biophys. Mol. Biol.* **2007**, *95*, 83–137.
48. Ippolito, J.A.; Alexander, R.S.; Christianson, D.W. Hydrogen bond stereochemistry in protein structure and function. *J. Mol. Biol.* **1990**, *215*, 457–471.
49. Morgan, R.S.; Tatsch, C.E.; Gushard, R.H.; McAdon, J.M.; Warne, P.K. Chains of alternating sulfur and π -bonded atoms in eight small proteins. *Int. J. Peptide Protein Res.* **1978**, *11*, 209–217.
50. Reid, K.S.C.; Lindley, P.F.; Thornton, J.M. Sulfur-aromatic interactions in proteins. *FEBS Lett.* **1985**, *190*, 209–213.
51. Viguera, A.R.; Serrano, L. Side-chain interactions between sulfur-containing amino acids and phenylalanine in α -helices. *Biochemistry* **1995**, *34*, 8771–8779.
52. Cheney, B.V.; Schulz, M.W.; Cheney, J. Complexes of benzene with formamide and methanethiol as models for interactions of protein substructures. *Biochim. Biophys. Acta* **1989**, *996*, 116–124.
53. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The protein data bank. *Nucl. Acids Res.* **2000**, *28*, 235–242.
54. Møller, C.; Plesset, M.S. Note on the approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618–622.

55. Taylor, J.C.; Markham, G.D. The bifunctional active site of *S*-adenosylmethionine synthetase—Roles of the active site aspartates. *J. Biol. Chem.* **1999**, *274*, 32909–32914.
56. Brandt, W.; Golbraikh, A.; Täger, M.; Lendeckel, U. A molecular mechanism for the cleavage of a disulfide bond as the primary function of agonist binding to G-protein-coupled receptors based on theoretical calculations supported by experiments. *Eur. J. Biochem.* **1999**, *261*, 89–97.
57. Palenchar, J.B.; Crocco, J.M.; Colman, R.F. The characterization of mutant *Bacillus subtilis* adenylosuccinate lyases corresponding to severe human adenylosuccinate lyase deficiencies. *Protein Sci.* **2003**, *12*, 1694–1705.
58. Singh, A.K.; Singh, N.; Sharma, S.; Shin, K.; Takase, M.; Kaur, P.; Srinivasan, A.; Singh, T.P. Inhibition of lactoperoxidase by its own catalytic product: Crystal structure of the hypothyocyanate-inhibited bovine lactoperoxidase at 2.3-Å resolution. *Biophys. J.* **2009**, *96*, 646–654.
59. Nakamura, T.; Yamamoto, T.; Abe, M.; Matsumura, H.; Hagihara, Y.; Goto, T.; Yamaguchi, T.; Inoue, T. Oxidation of archaeal peroxiredoxin involves a hypervalent sulfur intermediate. *Proc. Nat. Acad. Sci. USA* **2008**, *105*, 6238–6242.
60. Iwaoka, M.; Tomoda, S. The SAAP force field. A simple approach to a new all-atom protein force field by using single amino acid potential (SAAP) functions in various solvents. *J. Comput. Chem.* **2003**, *24*, 1192–1200.
61. Iwaoka, M.; Kimura, N.; Yosida, D.; Minezaki, T. The SAAP force field: Development of the single amino acid potentials for 20 proteinogenic amino acids and Monte Carlo molecular simulation for short peptides. *J. Comput. Chem.* **2009**, *30*, 2039–2055.
62. Dennis, E.A. Diversity of group types, regulation, and function of phospholipase A₂. *J. Biol. Chem.* **1994**, *269*, 13057–13060.
63. Jain, M.K.; Gelb, M.H.; Rogers, J.; Berg, O.G. Kinetic basis for interfacial catalysis by phospholipase A₂. *Methods Enzymol.* **1995**, *249*, 567–614.
64. Yu, B.Z.; Pan, Y.H.; Janssen, M.J.W.; Bahnson, B.J.; Jain, M.K. Kinetic and structural properties of disulfide engineered phospholipase A₂: Insight into the role of disulfide bonding patterns. *Biochemistry* **2005**, *44*, 3369–3379.
65. Ohno, M.; Chijiwa, T.; Oda-Ueda, N.; Ogawa, T.; Hattori, S. Molecular evolution of myotoxic phospholipase A₂ from snake venom. *Toxicon* **2003**, *42*, 841–854.
66. Raines, R.T. Ribonuclease A. *Chem. Rev.* **1998**, *98*, 1045–1065.
67. Rothwarf, D.M.; Li, Y.J.; Scheraga, H.A. Regeneration of bovine pancreatic ribonuclease A: Identification of two nativelylike three-disulfide intermediates involved in separate pathways. *Biochemistry* **1998**, *37*, 3760–3766.
68. Welker, E.; Narayan, M.; Wedemeyer, W.J.; Scheraga, H.A. Structural determinants of oxidative folding in proteins. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2312–2316.
69. Xu, X.; Rothwarf, D.M.; Scheraga, H.A. Nonrandom distribution of the one-disulfide intermediates in the regeneration of ribonuclease A. *Biochemistry* **1996**, *35*, 6406–6417.
70. Li, Y.J.; Rothwarf, D.M.; Scheraga, H.A. Mechanism of reductive protein unfolding. *Nat. Struct. Mol. Biol.* **1995**, *2*, 489–494.
71. Nachman, J.; Miller, M.; Gilliland, G.L.; Carty, R.; Pincus, M.; Wlodawer, A. Crystal structure of two covalent nucleoside derivatives of ribonuclease A. *Biochemistry* **1990**, *29*, 928–937.

72. Leonidas, D.D.; Shapiro, R.; Irons, L.I.; Russo, N.; Acharya, K.R. Crystal structures of ribonuclease A complexes with 5'-diphosphoadenosine 3'-phosphate and 5'-diphosphoadenosine 2'-phosphate at 1.7 Å resolution. *Biochemistry* **1997**, *36*, 5578–5588.
73. Leonidas, D.D.; Shapiro, R.; Irons, L.I.; Russo, N.; Acharya, K.R. Toward rational design of ribonuclease inhibitors: High-resolution crystal structure of a ribonuclease A complex with a potent 3',5'-pyrophosphate-linked dinucleotide inhibitor. *Biochemistry* **1999**, *38*, 10287–10297.
74. de Meyts, P. Insulin and its receptor: Structure, function and evolution. *BioEssays* **2004**, *26*, 1351–1362.
75. Merlini, G.; Bellotti, V. Lysozyme: A paradigmatic molecule for the investigation of protein structure, function and misfolding. *Clin. Chim. Acta* **2005**, *357*, 168–172.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).