J|A|C|S
ARTICLES
Published on Web 08/13/2002

Statistical and Theoretical Investigations on the Directionality of Nonbonded S···O Interactions. Implications for Molecular Design and Protein Engineering

Michio Iwaoka,* Shinya Takemoto, and Shuji Tomoda*

Contribution from the Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8902, Japan

Received April 9, 2002

Abstract: Weak nonbonded interactions between a divalent sulfur (S) atom and a main-chain carbonyl oxygen (O) atom have recently been characterized in proteins. However, they have shown distinctly different directional propensities around the O atom from the S···O interactions in small organic compounds, although the linearity of the C−S···O or S−S···O atomic alignment was commonly observed. To elucidate the observed discrepancy, a comprehensive search for nonbonded S···O interactions in the Cambridge Structural Database (CSD) and MP2 calculations on the model complexes between dimethyl disulfide (CH_3-$SSCH_3$) and various carbonyl compounds were performed. It was found that the O atom showed a strong intrinsic tendency to approach the S atom from the backside of the S−C or S−S bond (in the σ_S* direction). On the other hand, the S atom had both possibilities of approach to the carbonyl O atom within the same plane (in the n_O direction) and out of the plane (in the π_O direction). In the case of S···O(amide) interactions, the π_O direction was significantly preferred as observed in proteins. Thus, structural features of S···O interactions depend on the type of carbonyl groups involved. The results suggested that S···O interactions may control protein structures to some extent and that the unique directional properties of S···O interactions could be applied to molecular design.

Introduction

Weak nonbonded interaction between a divalent sulfur (S) and an oxygen (O) atom has attracted growing interest in various fields of chemistry, not only because the interaction plays important roles in the structure and the biological activity of some organic sulfur compounds[1] but also because it would possibly regulate enzymatic functions.[2] We have recently suggested that nonbonded S···O interactions may also stabilize folded protein structures.[3] Although the chemical properties of such S···O interactions have been well-characterized,[3,4] their similarity to those in organic sulfur compounds and the discrepancy between them remain to be discussed.

Structural features of intra- and intermolecular S···O interactions were first analyzed by Rosenfield et al.:[5] they investigated the surrounding environments of a divalent S atom (Y−S−Z) in organic and inorganic crystals and showed that a nucleophilic O atom tends to approach the S atom from the backside of S−Y and S−Z bonds (the σ_S* directions). On the other hand, the relative directional preference of the S···O interactions to the O atom was studied in detail by Kucsman and Kapovitz.[6] For intramolecular 1,4- and 1,5-type S···O═C interactions, it was shown that the S atom tends to lie in the direction of the O lone pairs (the n_O direction). According to these statistical analyses, the stabilization mechanism of the typical S···O═C interactions in organic sulfur compounds have been described by the $n_O \rightarrow \sigma_S$* orbital interaction as illustrated in Figure 1a. The electrostatic nature of the 1,4-type S···O interactions between a positively charged S atom and a negatively charged ethereal O atom was also suggested by Burling et al.[1a] based on the substituent effects and the high-level theoretical calculations.

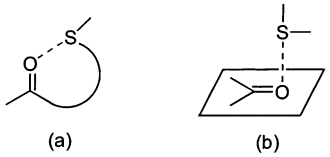

Figure 1. Structural features of nonbonded S···O interactions. (a) Intramolecular S···O interaction in organic sulfur compounds. (b) S···O interaction in proteins.

* To whom correspondence should be addressed: Phone +81 (3) 5454-6575. Fax +81 (3) 5454-6998. E-mail: iwaoka@selen.c.u-tokyo.ac.jp, tomoda@selen.c.u-tokyo.ac.jp.

(1) (a) Burling, F. T.; Goldstein, B. M. *J. Am. Chem. Soc.* **1992**, *114*, 2313−2320. (b) Nagao, Y.; Hirata, T.; Goto, S.; Sano, S.; Kakehi, A.; Iizuka, K.; Shiro, M. *J. Am. Chem. Soc.* **1998**, *120*, 3104−3110. (c) Wu, S.; Greer, A. *J. Org. Chem.* **2000**, *65*, 4883−4887.
(2) (a) Taylor, J. C.; Markham, G. D. *J. Biol. Chem.* **1999**, *274*, 32909−32914. (b) Brandt, W.; Golbraikh, A.; Täger, M.; Lendeckel, U. *Eur. J. Biochem.* **1999**, *261*, 89−97.
(3) (a) Iwaoka, M.; Takemoto, S.; Okada, M.; Tomoda, S. *Chem. Lett.* **2001**, 132−133. (b) Iwaoka, M.; Takemoto, S.; Okada, M.; Tomoda, S. *Bull. Chem. Soc. Jpn.* **2002**, *73*, 1611−1625.
(4) Pal, D.; Chakrabarti, P. *J. Biomol. Struct. Dyn.* **2001**, *19*, 115−128.
(5) Rosenfield, R. E., Jr.; Parthasarathy, R. *J. Am. Chem. Soc.* **1977**, *99*, 4860−4862.
(6) 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, 1985; pp 191−245.

Nonbonded S···O interactions in proteins have recently been pursued by us[3] and other research groups.[4,7] The stereochemistry of the nonbonded S···O interactions involving a methionine S atom was statistically analyzed by Carugo,[7] but no strong directional preference was observed. It was therefore suggested that the S···O interactions in proteins are either very weak or physicochemically different from those in small molecules. On the other hand, using a larger set of heterogeneous protein structures, we have found distinct directional preferences of the S···O interactions, as shown in Figure 1b: both a methionine S atom (a CSC group) and a cystine S atom (an SSC group) tend to approach a main-chain O atom perpendicularly to the amide plane (the π_O directions), and the O atom tends to approach the S atom from the backside of the S−C and S−S covalent bonds (the σ_S* directions).[3] Hence, the importance of the $\pi_O \rightarrow \sigma_S$* orbital interaction was proposed. Similar directionality of the S···O interactions involving a methionine S atom was also reported by Pal and Chakrabarti.[4] Moreover, our theoretical calculations using the Møller−Plesset method (MP2)[8] suggested that dispersion and/or long-range electrostatic forces are of primary importance for the stability of the S···O interactions and that coexisting C−H···O and N−H···O hydrogen bonds would stabilize them cooperatively (up to about 3.2 kcal/mol for the case involving an SSC group).[3]

The structural features of S···O interactions in proteins are thus obviously different from those characterized in organic molecules in that the S atom approaches the O atom out of the amide plane in proteins but within the π-plane in organic compounds (Figure 1), while the directional properties of the interactions relative to the S atom are commonly found. It should also be noted that the former S···O interactions are normally formed in the nonbonded region (r_S···O ≥ 3.25 Å) according to Kucsman and Kapovitz's definition,[6] while the latter are formed in a rather short distance range (r_S···O ≤ 3.25 Å). The discrepancy may possibly arise from some structural reasons: (1) The S···O interactions in organic compounds are usually formed between the nonbonded atoms separated by three or four covalent bonds (corresponding to 1,4- or 1,5-type S···O interaction, respectively), while those in proteins are much more distant throughout the polypeptide backbone. This structural restriction would not only prohibit the perpendicular approach of the S atom to the O atom in organic compounds but also reduce steric congestion around the O atom upon in-plane approach of the S atom. (2) The S···O interactions in proteins involve an amide O atom in most cases, while those in organic compounds involve various types of carbonyl groups, such as ketones and esters. The diversity would weaken the directional properties of S···O interactions in organic compounds. (3) In organic crystals, the packing force may be so strong that they affect the directional preferences of S···O interactions, while protein structures would be generally more flexible, even in the solid state, because they are governed only by weak noncovalent interactions.

Our main interest in the present work is to elucidate intrinsic directional properties of S···O interactions, which would provide valuable information in the fields of molecular design[9] and protein engineering.[10] To approach the goal, we have carried out extensive database analysis of close *intermolecular* S···O interactions in organic crystals and have compared the structural features with those of the S···O interactions in proteins as well as those of the *intramolecular* S···O interactions in organic crystals. We have also performed MP2 calculations on the model complexes between dimethyl disulfide (CH_3SSCH_3, **1**) and formaldehyde (HCHO, **2**), acetaldehyde (CH_3CHO, **3**), acetone (CH_3COCH_3, **4**), 2-butanone ($CH_3COCH_2CH_3$, **5**), methyl acetate (CH_3COOCH_3, **6**), and *N*-methylacetamide ($CH_3CONHCH_3$, **7**) to analyze the effects of the type of carbonyl O atoms on the strength and directionality of the S···O interactions. On the basis of these statistical and theoretical investigations, the possibility of S···O interactions as useful chemical tools for molecular design and protein engineering is discussed.

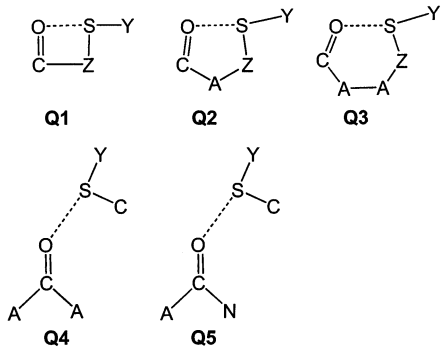

Figure 2. Structures of the fragments used for the CSD analysis. An upper-case A is indicative of any atom including hydrogen. Y and Z are indicative of C or S (excluding the case that both Y and Z are S at the same time).

Experimental Section

Intra- and intermolecular S···O contacts with the nonbonded S···O atomic distance between 2.5 and 5.0 Å were comprehensively searched in the structural data retrieved from the Cambridge Structural Database (CSD version 5.21, April 2001) using the Quest3D program[11] installed on the SPP computer system of Data Processing Center of Kyoto University. The quested interaction patterns (fragments **Q1**−**Q5**) are shown in Figure 2, where an upper-case A is indicative of any atom. Y and Z are indicative of C or S. To compare the structural features of the S···O interactions in organic crystals with those in proteins, we have investigated only the interactions formed between a divalent organic S atom (i.e., only CSC and SSC groups) and a carbonyl O atom. For fragments **Q1**−**Q3**, the minimum bond path between the S and O atoms was only considered. The queries were made for the crystal structures selected according to the following criteria: (a) the crystallographic *R* factor ≤ 10%, (b) with error-free coordinates according to the criteria used in the CSD system, (c) with no crystallographic disorder, and (d) no polymeric structures. The total numbers of the fragments found in the selected crystal structures were also counted.

The data of intra- and intermolecular S···O interactions (fragments **Q1**−**Q3** and **Q4**−**Q5**, respectively) were subsequently analyzed by the use of structural parameters defined in Figure 3. As for distance parameters, a nonbonded S···O atomic distance *r* and the relative distance *d* [= *r* − vdw(S) − vdw(O), where vdw(X) means the van der Waals radius of atom X; 1.80 Å for S and 1.52 Å for O[12]] were utilized. As for angular parameters, θ_1−θ_4 were defined. Angles θ_1 and θ_2 were utilized to show the spatial locations of O relative to S of

(7) Carugo, O. *Biol. Chem.* **1999**, *380*, 495−498.
(8) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618−622.
(9) (a) Glusker, J. P. *Curr. Chem.* **1998**, *198*, 1−56. (b) Inoue, Y.; Wada, T. *Adv. Supramol. Chem.* **1997**, *4*, 55−96.
(10) (a) Betz, S. F.; Bryson, J. W.; Passador, M. C.; Brown, R. J. *Acta Chem. Scand.* **1996**, *50*, 688−96. (b) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. *Science* **1995**, *270*, 935−941.
(11) Allen, F. H.; Davies, J. E.; Galloy, J. J.; Johnson, O.; Kennard, O.; Macrae, C. F.; Mitchell, E. M.; Mitchell, G. F.; Smith, J. M.; Watson, D. G. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 187−204.
(12) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441−451.

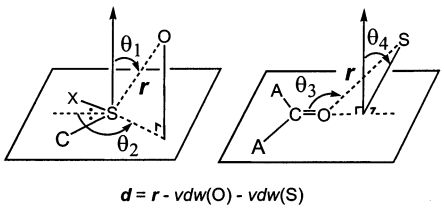

Figure 3. Structural parameters for analyzing the directionality of S···O interactions. An upper-case A is indicative of any atom including hydrogen. X is indicative of C or S. r is an S···O atomic distance. d is a relative S···O distance to the sum of the van der Waals radii.

CSC and SSC groups. Angles θ_3 and θ_4 were used to indicate the spatial locations of S relative to the C═O plane.

For comparison, the S···O interactions in proteins were also surveyed according to the method described previously.[3] Nonredundant protein structures with the structural homology less than 25% were extracted from Protein Data Bank (PDB) by the use of the PDB_SELECT program.[13] The statistical data obtained for the S···O interactions involving CSC and SSC groups (fragments **Q5**, where atom Y is C and S, respectively) were analyzed separately.

Molecular orbital calculations were carried out by using the Gaussian 98 program[14] installed on DEC Alpha21164A Unix workstations. All complex structures shown in this paper were fully optimized at the MP2/6-31G(d) level. The second-order Møller−Plesset (MP2) perturbation theory[8] was applied, because the reasonable complex structures could not be located with the Hartree−Fock method. The 6-31G(d) basis sets were chosen to obtain reliable results within an acceptable computation time. We have indeed applied larger basis sets, such as 6-31+G(d,p), for some complexes, but only marginal changes were observed in the structures and the complexation energies. All complexation energies were corrected for the basis set superposition error (BSSE) by applying the counterpoise method.[15]

Results and Discussion

Database Analysis. The statistical data obtained for intra- and intermolecular S···O═C interactions (fragments **Q1**−**Q5**) in CSD are summarized in Table 1, along with the corresponding data obtained for the S···O interactions in proteins. N_Q represents the number of fragment **Q** found in the database. $N_{0.0}$ and $N_{0.2}$ represent the numbers of close S···O contacts matching the criteria of $d \leq 0.0$ and 0.2 Å, respectively, in the total fragments. For instance, a total of 8347 independent fragments were found for fragment **Q4** in CSD, among which 292 fragments (3.5%) had the intermolecular nonbonded S···O distance r less than 3.32 Å ($d \leq 0.0$ Å) and 626 fragments (7.5%) had r less than 3.52 Å ($d \leq 0.2$ Å).

For fragments **Q4** and **Q5**, the observed S···O contacts were further classified to S(CSC)···O and S(SSC)···O interactions, depending on the type of S, in order to compare them with those in proteins. However, the dihedral angle through a disulfide (S−S) bond in some organic sulfur compounds was found to be structurally constrained to ∼0°, whereas proteins usually possessed unconstrained S−S bonds with a normal dihedral angle of ∼90°.[16] Therefore, the S···O contacts involving an S−S bond, whose dihedral angle is larger than 70°, were selected from the observed S(SSC)···O interactions. Such S···O interactions are denoted by symbol S* in Table 1.

Table 1. Summary of the Database Analysis of S···O═C Interactions[a]

fragment	Y	Z	N_Q[b]	$N_{0.0}$ (%)[c]	$N_{0.2}$ (%)[c]
Q1	C/C/S	C/S/C	1215	551 (45.3)	715 (58.8)
Q2	C/C/S	C/S/C	2320	316 (13.6)	399 (17.2)
Q3	C/C/S	C/S/C	1417	46 (3.2)	86 (5.9)
Q4	C/S		8347	292 (3.5)	626 (7.5)
	C		7786	208 (2.7)	497 (6.4)
	S		561	84 (15.0)	129 (23.0)
	S*[d]		362	3 (0.8)	25 (6.9)
Q5	C		2686	75 (2.8)	158 (5.9)
	S		264	29 (11.0)	51 (19.3)
	S*[d]		150	2 (1.3)	9 (6.0)
Q5 in proteins	C		2124	28 (1.3)	94 (4.4)
Q5 in proteins	S		790	70 (8.9)	154 (19.5)

[a] The data of **Q1**−**Q5** and **Q5** in proteins were obtained from CSD and PDB, respectively. See the text for details. [b] The total numbers of the corresponding independent fragments found in the database. [c] $N_{0.0}$ and $N_{0.2}$ represent the numbers of close S···O contacts with $d \leq 0.0$ and 0.2 Å, respectively. The values in parentheses represent percentages of the close contacts to the total number of the fragment ($N/N_Q \times 100$). [d] S atoms involved in the SSC group whose S−S bond dihedral angle is larger than 70°.

Intramolecular S···O Interactions (Q1−Q3). The ratios of close S···O contacts to the total number of the corresponding fragment ($N_{0.0}/N_Q$ and $N_{0.2}/N_Q$) significantly decreased upon increasing the number of covalent bonds that intervene between the S and O atoms. The ratios for 1,6-type S···O interactions (**Q3**) were almost the same as those for intermolecular S···O interactions (**Q4**). The results suggested that the nature of S···O interactions is significantly affected by the length of the covalent linkage between the interacting S and O atoms.

Directional preferences of 1,4-, 1,5-, and 1,6-type S···O interactions were analyzed by the use of angles θ_1−θ_4. Figure 4a displays the spatial distribution of the O atoms ($d \leq 0.0$ Å) relative to the CSC and SSC S atoms. The plots for 1,4-S···O interactions showed up like a circle, due to internal rotation of fragment **Q1**. When the interactions were formed in a short distance range ($d \leq -0.4$ Å), most plots concentrated in the area of $\theta_1 = 70°-110°$ and $\theta_2 = 90°-110°$, which corresponded to an orientation slightly hindered from the backside of the S−C or S−S bond (a bent direction of the antibonding σ_S* orbital) due to the structural constraint of fragment **Q1**. On the other hand, the plots for 1,5-S···O interactions (**Q2**) made a strong cluster in the area of $\theta_1 = 70°-110°$ and $\theta_2 = 110°-130°$, the center of which was closer to a linear extension of the S−C or S−S bond ($\theta_1 \sim 90°$ and $\theta_2 \sim 130°$). The most ideal directional property for the effective orbital interaction with σ_S* was obtained for 1,6-S···O interactions (**Q3**), although the number of the plots was small ($N_{0.0} = 46$). Thus, it appeared that the carbonyl O atom tends to approach the S atom in the direction of the σ_S* orbital with a decrease of structural restriction.

Figure 4b shows directional preferences of intramolecular S···O interactions relative to the carbonyl O atom by using

(13) (a) Hobohm, U.; Scharf, M.; Schneider, R.; Sander, C. *Protein Sci.* **1992**, *1*, 409−417. (b) Hobohm, U.; Sander, C. *Protein Sci.* **1994**, *3*, 522−524.
(14) Fisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
(15) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503−506.
(16) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167−339.

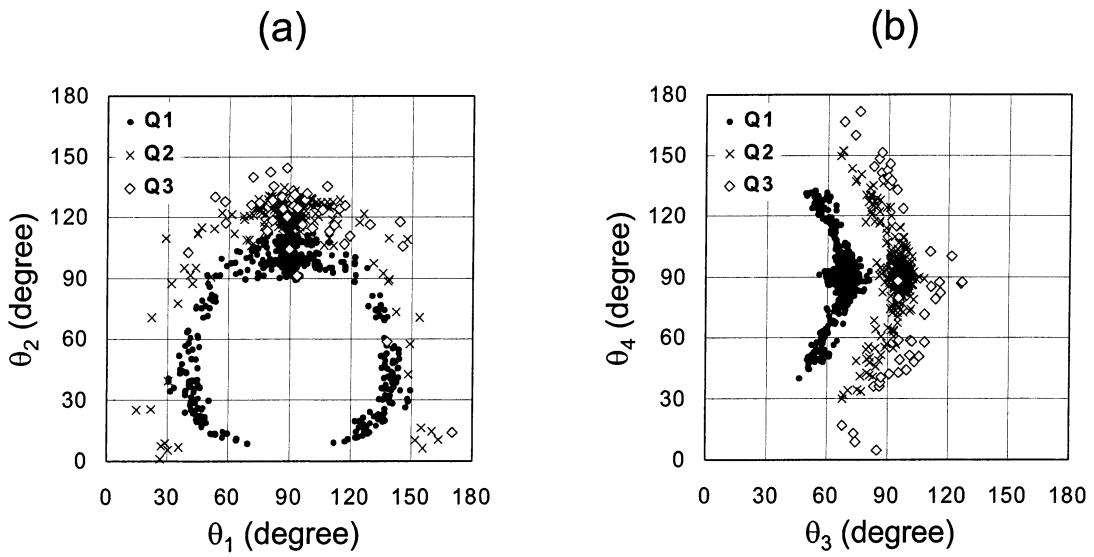

Figure 4. Directionality of intramolecular 1,4- (**Q1**), 1,5- (**Q2**), and 1,6-type (**Q3**) S···O═C interactions with $d \leq 0.0$ Å. (a) Spatial distribution of O relative to S determined by using angles θ_1 and θ_2. (b) Spatial distribution of S relative to O determined by using angles θ_3 and θ_4.

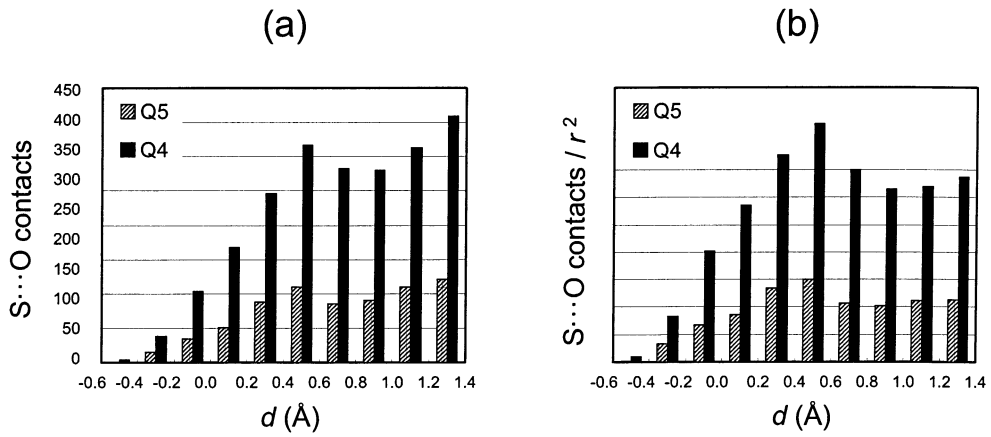

Figure 5. The numbers of intermolecular S···O contacts [fragments **Q4** (filled bars) and **Q5** (striped bars)] as a function of the relative nonbonded distance *d*. (a) Observed. (b) Normalized with the possible spherical surface area by a factor of $1/r^2$.

angles θ_3 and θ_4. For 1,4- and 1,5-type S···O interactions, the crescent clusters appeared. Centers of the clusters were found at around $\theta_3 \sim 70°$ and $\theta_4 \sim 90°$ and around $\theta_3 \sim 100°$ and $\theta_4 \sim 90°$, respectively, when the nonbonded S···O distance was short. The orientations might be a little distorted from the direction of the O lone pair on the carbonyl plane (the n_O direction): the n_O direction would be at $\theta_3 \sim 120°$ and $\theta_4 \sim 90°$ for the sp^2-hybridized O atom. However, the situation was distinctly different for 1,6-S···O interaction: the plots spread like a vertical sine curve, and some plots were found even in the direction perpendicular to the carbonyl plane (the π_O direction) with $\theta_4 \sim 0°$ or 180°. The results suggested that the preference for an n_O direction may be specific only to 1,4- and 1,5-S···O interactions.

A similar database analysis of intramolecular 1,4-, 1,5-, and 1,6-S···O interactions was previously carried out by Kucsman and Kapovitz[6] in 1985. However, fewer close S···O contacts were observed at that time (150, 130, and 0 contacts for 1,4-, 1,5-, and 1,6-S···O interactions, respectively, with $r \leq 3.25$ Å), compared with the corresponding $N_{0.0}$ values in Table 1 (551, 316, and 46 contacts, respectively, with $r \leq 3.32$ Å). Structural preferences previously demonstrated for 1,4- and 1,5-S···O interactions (Figure 1a) have been reasonably reproduced in the present work by using larger sets of S···O contact data as described above. They were also in good agreement with previous quantum chemical calculations.[17] Furthermore, the data in Table 1 have statistically revealed the presence of 1,6-S···O interactions, which were not detected in the previous work[6] but were reported in some of the individual compounds.[18] Since the formation of 1,6-S···O interactions is less advantageous from an entropy standpoint than that of 1,4- and 1,5-S···O interactions, 1,6-S···O interactions may have rarely been observed in CSD.

Intermolecular S···O Interactions (Q4). The number of intermolecular S···O contacts (fragment **Q4**) is graphically shown in Figure 5 with filled bars as a function of the relative nonbonded distance (*d*). When the number was normalized with a possible spherical surface area by a factor of $1/r^2$ (Figure 5b), a broad peak appeared at $d \sim 0.5$ Å. The presence of long-range intermolecular S···O interactions was thus suggested in

(17) (a) Ángáyan, J. G.; Poirier, R. A.; Kucsman, Á.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1987**, *109*, 2237−2245. (b) Gough, K. M.; Millington, J. *Can. J. Chem.* **1995**, *73*, 1287−1293. (c) Minyaev, R. M.; Minkin, V. I. *Can. J. Chem.* **1998**, *76*, 776−788.

(18) (a) Raymond, J. F.; Kemmitt, D. W.; Russell, D. R.; Serindag, O. *Acta Crystallogr. Sect. C* **1993**, *49*, 1434−1436. (b) Neidlein, R.; Hartz, G.; Gieren, A.; Betz, H.; Hübner, T. *Chem. Ber.* **1985**, *118*, 1455−1462.

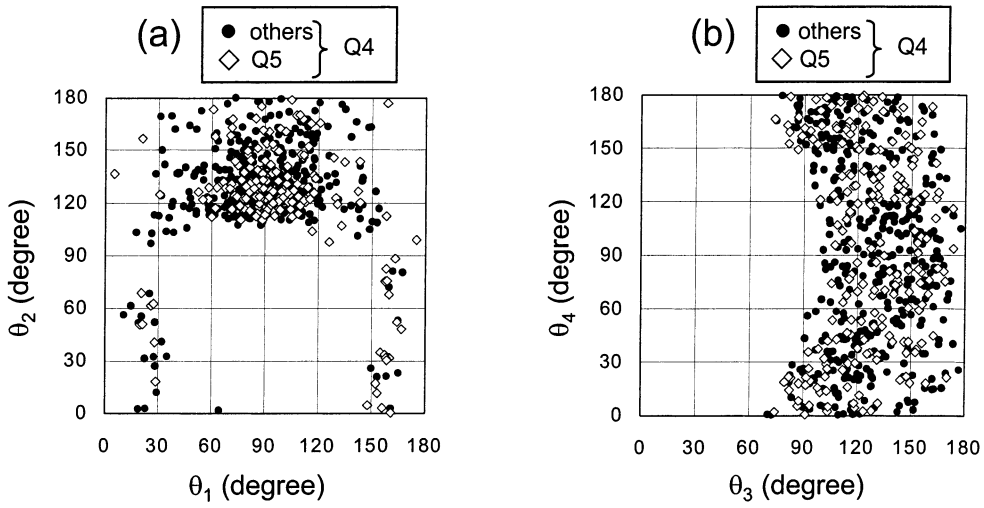

Figure 6. Directionality of intermolecular S···O═C interactions (**Q4**) with $d \leq 0.2$ Å. Open diamonds represent intermolecular S···O(amide) interactions (**Q5**). Filled circles represent the other types of intermolecular S···O interactions. (a) Spatial distribution of O relative to S determined by using angles θ_1 and θ_2. (b) Spatial distribution of S relative to O determined by using angles θ_3 and θ_4.

molecular crystals of organic sulfur compounds. The interactions may work in a wide range of distance ($-0.5 \leq d \leq 0.5$ Å). A similar broad peak was previously observed for the S···O interactions in proteins as well.[3]

Figure 6 shows structural features of intermolecular S···O interactions (**Q4**). Spatial locations of the carbonyl O atoms relative to the S atom (Figure 6a) made a cluster around $\theta_1 \sim 90°$ and $\theta_2 \sim 130°$, which is just the backside of the S−C or S−S bond. On the other hand, spatial locations of the S atoms relative to the carbonyl O atom (Figure 6b) did not make clusters: the plots spread randomly all over the area of $\theta_3 > 90°$, showing no clear directional preferences. The result was in sharp contrast not only to the directional preferences of intramolecular S···O interactions (**Q1**−**Q3**) shown in Figure 4b but also to the S···O interactions in proteins (Figure 1b).[3,4]

The observed intermolecular S···O═C interactions (**Q4**) were subsequently classified to S(CSC)···O and S(SSC)···O interactions by the Y atom. The ratios of these interactions to the total fragments ($N_{0.0}/N_Q$ and $N_{0.2}/N_Q$) listed in Table 1 clearly indicated that the S(SSC)···O interactions are more frequently detected in organic crystals. The observation is in accord with a higher reactivity of an S−S bond than that of an S−C bond.[19] A similar trend was observed for the S···O interactions in proteins (**Q5** in proteins).[3] However, most of the S(SSC)···O interactions observed in organic crystals involved an S−S bond with the dihedral angle less than 70°. The numbers of S(S*SC)···O interactions with a normal S−S dihedral angle were only three and 25 in the ranges of $d \leq 0.0$ and 0.2 Å, respectively. It was therefore suggested that the intermolecular S(SSC)···O interactions in organic crystals have different structural features from those in proteins. We found that the S(SSC)···O interactions of small organic sulfur compounds contain a large number of planar bifurcated interaction patterns shown in Figure 7.[20] Due to the presence of this type of S···O interactions, linear C−S···O atomic alignments were more frequently observed in organic crystals than linear S−S···O ones. The trend is indeed opposite to the S···O interactions in proteins.[3]

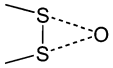

Figure 7. Planar bifurcated S(SSC)···O interaction.

S···O(amide) Interactions (Q5). To directly compare the S···O interactions observed in organic crystals (fragment **Q4**) with those in protein, subsets of the interactions, i.e., S···O(amide) interactions (fragment **Q5**), were extracted. These interactions occupied about 35% of the total S···O interactions. As seen in Figures 5 and 6, fragment **Q5** did not show any difference from fragment **Q4** in terms of the dependence on the nonbonded S···O distance and the directions.

For comparison, directional preferences of S···O interactions in proteins (**Q5** in proteins) with $d \leq 0.2$ Å are shown in Figure 8 by using angles $\theta_1-\theta_4$. To eliminate the effects of the backbone constraints, the scattergrams contain only the S···O interactions that are formed between the amino acids separated by more than 10 residues. The S(CSC)···O and S(SSC)···O interactions in proteins showed similar directional preferences to each other. The spatial locations of the main-chain amide O atoms relative to the S atom (Figure 8a) clustered in the area around $\theta_1 \sim 90°$ and $\theta_2 \sim 130°$, like the cases of intramolecular 1,6-S···O interactions (**Q3**) and intermolecular S···O interactions (**Q4** and **Q5**). However, the distribution of the S atoms relative to the amide O atom (Figure 8b) showed a unique cluster around $\theta_3 \sim 90°$ and $\theta_4 \sim 0°$ (or 180°). The cluster corresponded to the direction above or below the amide plane, namely the direction of the π_O orbital. The observed directionality was significantly different from those obtained for **Q1**−**Q5** in organic crystals.

MP2 Calculations. To elucidate the reasons for the observed discrepancy between the structural features of S···O interactions in organic crystals (Figure 6) and those in protein (Figure 8), quantum chemical calculations were performed for several carbonyl compounds (**2**−**7**) and their molecular complexes with dimethyl disulfide (**1**). The results are summarized in Table 2.

(19) (a) Benson, S. W. *Chem. Rev.* **1978**, *78*, 23−35. (b) Maung, N. *J. Mol. Struct. (THEOCHEM)* **1999**, *460*, 159−166.

(20) (a) Mogensen, P. K.; Simonsen, O. *Acta Crystallogr. Sect. C* **1991**, *47*, 1905−1908. (b) Rees, W. C.; White, A. J. P.; Williams, D. J. *J. Org. Chem.* **1998**, *63*, 2189−2196. (c) Rees, W. C.; White, A. J. P.; Williams, D. J. *J. Org. Chem.* **1999**, *64*, 5010−5016. (d) Barriga, S.; Konstantinova, L. S.; Marcos, C. F.; Rakitin, O. A.; Rees, C. W.; Torroba, T.; White, A. J. P.; Williams, D. J. *J. Chem. Soc., Perkin Trans. 1* **1999**, 2237−2241.

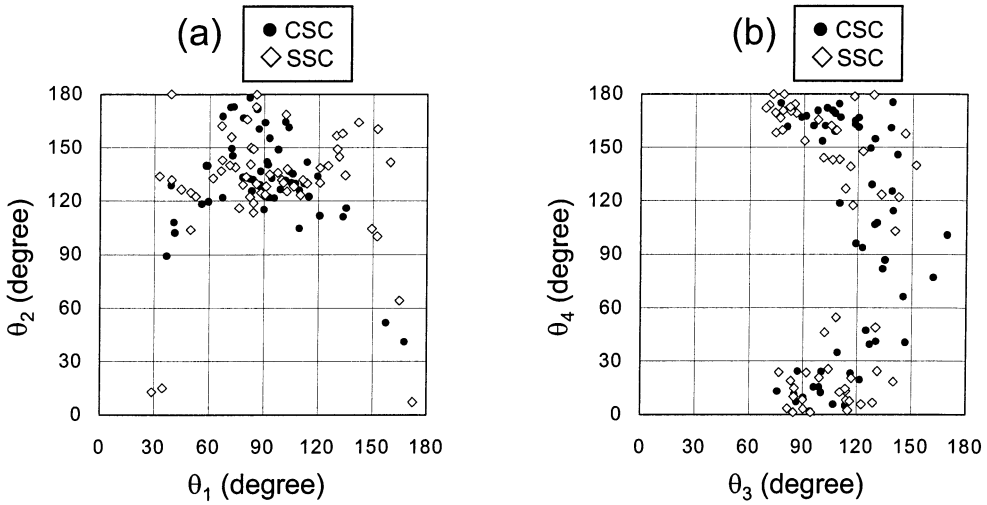

Figure 8. Directionality of S···O interactions in proteins (**Q5** in proteins) with $d \leq 0.2$ Å. Only the cases where the interactions are formed between amino acids separated by more than 10 residues are plotted. Filled circles represent the S(CSC)···O interactions. Open diamonds represent the S(SSC)···O interactions. (a) Spatial distribution of main-chain O relative to S determined by using angles θ_1 and θ_2. (b) Spatial distribution of S relative to main-chain O determined by using angles θ_3 and θ_4.

Table 2. Summary of the Quantum Chemical Calculations[a]

compounds	orbitals	energy levels (au)	total energies of the complexes with **1**[b] (au)	complexation energies[c] (kcal/mol)
HCHO (**2**)	n_O (HOMO)	−0.440	−988.927 28 (**P2**)	−1.82
	π_O (HOMO − 1)	−0.525	*d*	*d*
CH_3CHO (**3**)	n_O (HOMO)	−0.424	−1028.106 70 (**P3**)	−1.95
	π_O (HOMO − 1)	−0.495	−1028.106 48 (**V3**)	−2.05
CH_3COCH_3 (**4**)	n_O (HOMO)	−0.411	−1067.283 81 (**P4**)	−2.11
	π_O (HOMO − 1)	−0.479	−1067.284 61 (**V4**)	−2.52
$CH_3COCH_2CH_3$ (**5**)	n_O (HOMO)	−0.408	−1106.450 76 (**P5**)	−2.14
	π_O (HOMO − 1)	−0.476	−1106.452 15 (**V5**)	−2.73
CH_3COOCH_3 (**6**)	n_O (HOMO)	−0.444	−1142.333 02 (**P6**)	−2.14
	π_O (HOMO − 1)	−0.457	−1142.333 93 (**V6**)	−2.42
$CH_3CONHCH_3$ (**7**)	n_O (HOMO − 1)	−0.413	−1122.489 31 (**P7**)	−2.18
	π_O (HOMO)	−0.383	−1122.492 01 (**V7**)	−3.21

[a] Calculated at the MP2/6-31G(d) level. [b] Structures of the complexes (**P2**−**P7** and **V3**−**V7**) are shown in Figure 9. [c] Corrected with the basis set superposition errors (BSSE). [d] A stable structure was not found.

We employed disulfide **1**, instead of a simple sulfide (CH_3SCH_3), as a counterpart of the S···O interactions, because stable structures could not be located for some of the complexes with CH_3SCH_3, and also because the complexation energies were always smaller than those for the complexes with **1**.

It is clearly seen that the energy level of the carbonyl π orbital (π_O) of *N*-methylacetamide (**7**) is remarkably raised compared with that of formaldehyde (**2**), while the energy level of the O lone pair (n_O) remains almost unchanged. The reason for the elevation of the π_O orbital is due to the conjugation between the N lone pair and the carbonyl group. It is also important to note that the HOMO is assigned to n_O for **2**−**5**, whereas it is assigned to π_O for **7**. The energy levels of n_O and π_O are very close to each other for ester **6**. The inversion of the energy levels of n_O and π_O should affect directional preferences of S···O interactions.

Indeed, the relative stabilities of the in-plane and vertical complexes of **2**−**7** with **1** (see also Figure 9) reasonably reflect the relative energy levels of the n_O and π_O orbitals of **2**−**7**. For formaldehyde (**2**), only an in-plane complex (**P2**) was obtained as the stable structure, probably because the energy level of π_O is too low for **2**. The complexation energy was calculated as −1.82 kcal/mol with the BSSE corrections at the MP2 level. On the other hand, both in-plane and vertical complexes could be located for **3**−**7**. The total energies and complexation energies obtained for the in-plane complexes of **3**−**6** (**P3**−**P6**) were similar to those obtained for the corresponding vertical complexes (**V3**−**V6**), while in the case of amide **7** the vertical complex (**V7**) was significantly more stable than the in-plane complex (**P7**). It should be noted that complexation energies in Table 2 include the contribution from a coexisting C−H···O hydrogen bond (not indicated in Figure 9) in all cases. Therefore, the stabilization energies due to S···O interactions alone would be smaller than the calculated complexation energies.

Structural parameters of the S···O interactions obtained for complexes **P2**−**P7** and **V3**−**V7** are listed in Table 3. The nonbonded S···O distances (r) range from 3.26 to 3.51 Å, which are approximately equal to the sum of van der Waals radii of S and O atoms ($-0.06 \leq d \leq 0.19$ Å). For all complexes, the carbonyl O atom of **2**−**7** approaches the S atom of **1** in the backside of the S−S bond ($\theta_1 = 78.1 \sim 88.4°$, $\theta_2 = 119.6 \sim 128.8°$). The linearity of an S−S···O atomic alignment is consistent with the directional preferences of S···O interactions observed in Figures 4a, 6a, and 8a. On the other hand, the S atom of **1** lies on the carbonyl plane of **2**−**7** for the in-plane complexes (**P2**−**P7**; $\theta_3 = 91.6-160.7°$, $\theta_4 = 66.6-105.6°$), while for the vertical complexes (**V3**−**V7**) the S atom comes over the carbonyl O atom ($\theta_3 = 80.6-88.3°$, $\theta_4 = 0.1-20.8°$).

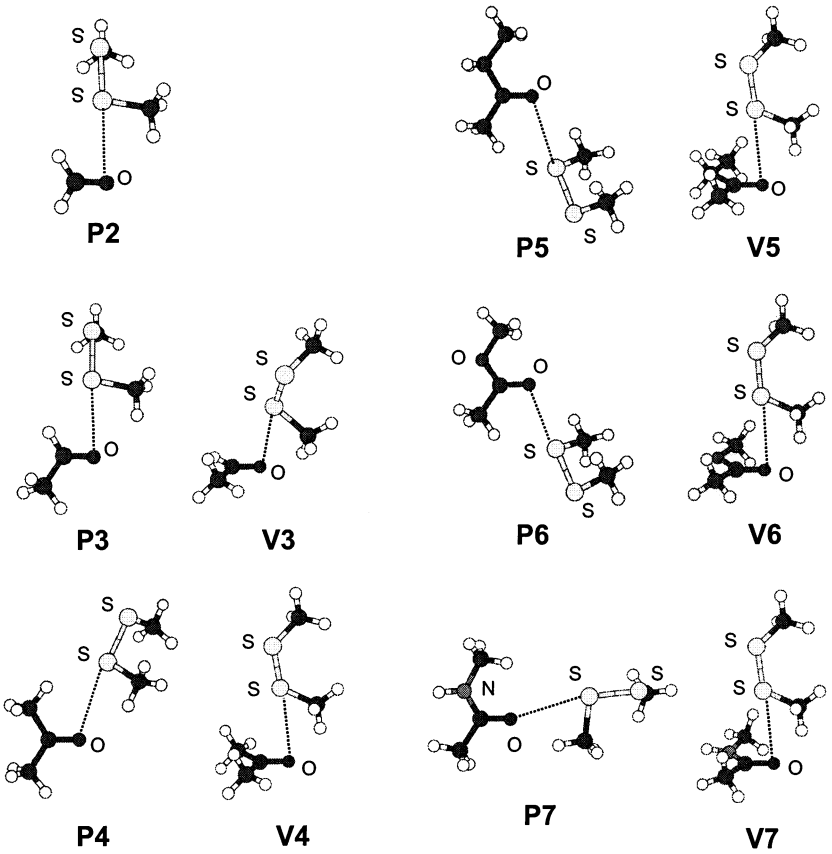

Figure 9. Structures of the molecular complexes between various carbonyl compounds (**2**−**7**) and dimethyl disulfide (**1**) optimized at the MP2/6-31G(d) level. **P** denotes the in-plane complex formed in the direction of the n_O orbital. **V** denotes the vertical complex formed in the direction of the π_O orbital. Complex **V2** could not be located as a stable structure.

Table 3. Structural Parameters of the S···O Interactions for Complexes **P2**−**P7** and **V3**−**V7**[a]

compounds	complexes	r (Å)	d (Å)	θ_1 (deg)	θ_2 (deg)	θ_3 (deg)	θ_4 (deg)
HCHO (**2**)	**P2**	3.26	−0.06	82.6	128.8	91.6	87.2
CH_3CHO (**3**)	**P3**	3.28	−0.04	81.3	127.4	91.8	88.8
	V3	3.47	0.15	82.4	122.2	88.3	20.8
CH_3COCH_3 (**4**)	**P4**	3.50	0.18	86.4	121.3	110.3	105.6
	V4	3.38	0.06	78.1	123.1	80.6	0.1
$CH_3COCH_2CH_3$ (**5**)	**P5**	3.51	0.19	86.2	121.2	109.0	103.4
	V5	3.35	0.03	78.9	123.5	80.6	2.5
CH_3COOCH_3 (**6**)	**P6**	3.47	0.15	85.0	121.9	107.9	103.6
	V6	3.28	−0.04	81.8	125.7	84.1	2.3
$CH_3CONHCH_3$ (**7**)	**P7**	3.38	0.06	88.4	119.6	160.7	66.6
	V7	3.34	0.02	81.7	122.5	81.5	2.3

[a] Calculated at the MP2/6-31G(d) level. Definitions of the structural parameters are shown in Figure 3.

It is obvious that θ_3 becomes larger for the in-plane complexes (**P2**−**P7**) as the steric congestion around the carbonyl group increases.

Directional Properties of S···O Interactions. On the basis of the results from MP2 calculations, the structural features of S···O═C interactions observed in organic crystals (Figures 4 and 6) and in proteins (Figure 8) can be rationalized as follows.

First, the linearity of a C−S···O or S−S···O atomic alignment for fragments **Q1**−**Q5** should arise from a strong intrinsic propensity of the S···O interactions, as seen in the calculated stable structures (**P2**−**P7** and **V3**−**V7**). For 1,4- and 1,5-S···O interactions, the linearity was slightly disturbed by the structural constraints of fragments **Q1** and **Q2**. However, 1,6- and intermolecular S···O interactions in organic crystals and those in proteins commonly showed fine linearity: a cluster appeared in the area around $\theta_1 \sim 90°$ and $\theta_2 \sim 130°$ on the θ_1 vs θ_2 scattergrams (Figures 4a, 6a, and 8a).

Second, the variable relative directional properties of S···O interactions to the carbonyl O atom seen in Figures 4b, 6b, and 8b should result from several structural factors: (1) For fragments **Q1** and **Q2**, a vertical attack of the S atom to the carbonyl plane is prohibited by the structural constraints, making only in-plane formation of intramolecular S···O interactions possible. (2) In the cases of in-plane S···O interactions, steric congestion around the carbonyl group would push the S atom away from the bulky substituents, as observed in **P2**−**P7**. Such steric repulsion may be of importance for the observed difference in the lower limit values of θ_3 at $\theta_4 = 90°$ ($\theta_3 \geq \sim 60°$ for **Q1**, $\sim 90°$ for **Q2**, $\sim 100°$ for **Q3**−**Q5**, and $\sim 120°$ for **Q5** in proteins). (3) A directional preference of S···O interactions in proteins, i.e., a vertical attack of S to O (Figures 1b and 8b), can be ascribed to a strong intrinsic propensity of S···O(amide) interactions. (4) Involvement of various types of carbonyl groups, such as ketones and esters, in intermolecular S···O interactions (**Q4**) would dilute the directionality to some extent, as seen in Figure 6b. (5) Intermolecular S···O(amide) interactions (**Q5**) in organic crystals did not show any directionality despite the uniformity. This may be due to the effects of a significant crystal packing force.

According to the above considerations, it is assumed that the linearity of Y−S···O (Y = C and S) interactions is not affected by crystal packing force but is slightly disturbed by structural constraints of the interacting fragments. Similarly, the vertical

nature of S···O(amide) interactions is not affected by weak noncovalent interactions in protein structures but by crystal packing force. Therefore, the strengths of the external structural factors and intrinsic directional propensities of S···O interactions would decrease in the following order: structural constraints > Y−S···O linearity > packing force > verticality of S···O-(amide) > protein structures. The order is quite informative for molecular design and protein engineering, as discussed below. The conclusions are in accord with rather deep potential surfaces of the S···O(amide) interaction with respect to θ_1 and θ_2 angles and the shallow ones with respect to θ_3 and θ_4 angles calculated previously.[3]

Implications for Molecular Design and Protein Engineering. The linearity of C−S···O and S−S···O atomic alignments must serve as useful tools in the fields of molecular design and crystal engineering. It has been a current topic to build huge particles with particular functions by self-assembling small molecular units.[9] To combine the unit molecules in a desired direction, hydrogen bonds, ionic interactions, and coordination of ligands to a metal center have been usually employed. S···O interactions can be applied as such chemical adhesives because they have strong linear directional propensities, which may overcome other intermolecular interactions such as crystal packing force. On the other hand, the structural preference of S···O interactions relative to the O atom would follow either an in-plane or vertical direction. This feature is in contrast to a previous view of the interactions illustrated in Figure 1a. The directional feature, however, may be easily affected by a crystal packing force.

In protein structures, S···O(amide) interactions maintain their intrinsic directional propensities (i.e., linearity and verticality as shown in Figure 1b), suggesting that the S···O(amide) interactions can be important elements in determining the stability and structure of proteins. Methionine (Met) and cysteine (Cys) residues in proteins have been considered to be merely hydrophobic: no significant interaction with other amino acid residues had been known, except for some specific weak interactions, such as S···C(π) interactions[21] and S−H···X (X = O and N) hydrogen bonds.[22] However, the S atoms in proteins should not be merely hydrophobic moieties. This concept may impose some impact on the field of protein engineering.[10] The mutation from Cys to Ala would lose not only the S−S linkage but also the S···O interactions if they exist originally. Similarly, the mutation from Met to other hydrophobic amino acids such as Leu would not only change the hydrophobicity. Moreover, S···O interactions must be useful for designing drugs containing an S atom and small cofactors interacting specifically to an S−S bond in enzymes. These applications of S···O interactions may be promising in light of recent studies showing the importance of S···O interactions in drug design[1] and enzymatic functions.[2]

Conclusions

According to the present statistical database analyses of various types of S···O═C interactions as well as quantum chemical calculations on the model systems, intrinsic structural preferences of S···O interactions have been characterized as follows. 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. In addition to these structural features, it has also been revealed that the linearity of S···O interactions would overcome crystal packing force, whereas the vertical nature of S···O(amide) interactions may be overcome by packing force. The verticality, however, would survive in protein structures. The features of S···O interactions revealed herein will be useful in the fields of molecular design and protein engineering.

Acknowledgment. We thank the Data Processing Center of Kyoto University for the use of the Cambridge Structural Database (CSD version 5.21). This work was supported by Grant-in-Aid for Scientific Research on Priority Areas (C) "Genomic Information Science" from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

JA026472Q

(21) (a) Morgan, R. S.; Tatsch, C. E.; Gushard, R. H.; McAdon, J. M.; Warme, P. K. *Int. J. Pept. Protein Res.* **1978**, *11*, 209−217. (b) Reid, K. S. C.; Lindley, P. F.; Thornton, J. M. *FEBS Lett.* **1985**, *190*, 209−213. (c) Zauhar, R. J.; Colbert, C. L.; Morgan, R. S.; Welsh, W. J. *Biopolymers* **2000**, *53*, 233−248.

(22) (a) Ippolito, J. A.; Alexander, R. S.; Christianson, D. W. *J. Mol. Biol.* **1990**, *215*, 457−471. (b) Stickle, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D. *J. Mol. Biol.* **1992**, *226*, 1143−1159. (c) Pal, D.; Charkrabarti, P. *J. Biomol. Struct. Dynam.* **1998**, *15*, 1059−1072.