Electronic Supplementary Information:

Mechanochemical Disulfide Reduction Reveals Imprints of Noncovalent Sulfur···Oxygen Chalcogen Bonds in Protein-inspired Mimics in Aqueous Solution

Przemyslaw Dopieralski[(1,2),*] Martin E. Zoloff Michoff[(1,3)], and Dominik Marx[(1)]

[(1)] *Lehrstuhl für Theoretische Chemie,*

Ruhr–Universität Bochum, 44780 Bochum, Germany

[(2)] *Permanent Address: Faculty of Chemistry,*

University of Wroclaw, Joliot–Curie 14, 50-383 Wroclaw, Poland

[(3)] *Permanent Address: INFIQC–CONICET,*

Departamento de Química Teórica y Computacional, Facultad de Ciencias Químicas,

Univesidad Nacional de Córdoba, Córdoba, Argentina

Contents

I. AB INITIO MOLECULAR DYNAMICS SIMULATIONS

For the *ab initio* molecular dynamics (AIMD) simulations[1] we have made use of the efficient Car–Parrinello[2] propagation scheme as implemented in the CPMD program package.[3] The underlying electronic structure calculations have been carried out using the BLYP[4,5] exchange–correlation functional supplemented by Grimme's dispersion correction[6], together with a plane wave basis set at a kinetic energy cutoff of 70 Ry and Troullier–Martins[7] norm-conserving pseudopotentials. The supercell for all reported calculations was a simple cubic box of 16 Å in length containing one extended disulfide model (EDM) (cf. Figure 1 in the main text), one deprotonated 1,4–dithiothreitol molecule (DTT$^-$) and 112 water molecules to properly solvate both, the reactant and nucleophile molecules subject to periodic boundary conditions.

All AIMD simulations were performed in the canonical (*NVT*) ensemble at 300 K using Nosé–Hoover chain thermostats[8] in order to control the kinetic energy of the nuclei as well as the fictitious kinetic energy of the orbitals. An AIMD step of $\Delta t = 4$ a.u. (≈ 0.097 fs) was used for the integration of the Car–Parrinello equations of motion using a fictitious mass parameter for the orbitals of $\mu = 700$ a.u. together with the atomic masses of all hydrogens substituted by deuterium masses to allow for more efficient propagation.[1] Overall, this approach is identical to the well–tested basic protocol that we used in our previous work to study attack of the small diethyl disulfide model by hydroxide (OH$^-$) in aqueous solution[9,10].

The setup and computational protocol of the isotensional canonical AIMD simulations at 300 K (160 ps at each constant tensile force after equilibration) of only EDM in vacuum (dubbed "isolated" throughout the text), thus mimicking a "dry" gas phase environment for reference purposes, was identical to the one just described for the solvated case in explicit aqueous solution.

II. AB INITIO METADYNAMICS SAMPLING

We have extended the metadynamics machinery that has been introduced, validated and extensively discussed in our previous publications on disulfide mechanochemistry[9,10], where in stark contrast to the present work only a minimal molecular model, namely diethyl

disulfide, has been investigated to elucidate attack by the simplest possible nucleophile, OH^-(aq). Here, we use a much extended disulfide model (EDM) which is attacked by the rather large Cleland's reagent nucleophile, DTT^-(aq), as depicted schematically in Figure 1 of the main text, to compute and analyze the force–transformed free energy landscapes[11] for the associated S_N2 disulfide reduction process *via* S–S bond cleavage as a function of constant external tensile force.

Two collective variables (CVs) that span the reaction subspace were employed in the Lagrangian *ab initio* metadynamics simulations[1,12–15] using four multiple–walker replica[16,17] to accelerate the disulfide reduction process in terms of enhanced sampling. To drive the S_N2 disulfide bond cleavage by the solvated DTT anion, DTT^-(aq), we have employed the following set of CVs. The first one, CV1, is defined as $d_{S-S^\star} - d_{S^\star\cdots-S(DTT)}$ where $d_{S-S^\star}$ is the distance between the two sulfur atoms within the EDM disulfide and $d_{S^\star\cdots-S(DTT)}$ is the distance between the attacked sulfur site of EDM (labeled $S^\star$ throughout) and the attacking sulfur atom of the DTT anion, $^-S(DTT)$. The second collective variable, CV2, is the $S–S^\star–C^\alpha–C^\beta$ dihedral angle on the side of the reaction center, see Figure 1 in the main text for the labeling of the corresponding α- and β–C atoms. Additional repulsive walls have been applied to CV1 in order to prevent DTT^- from diffusing too far away from the reaction center, thus avoiding unproductive sampling of translational motion; the corresponding walls have been set at ±2.9 Å.

The setup for the height of the Gaussian–shaped potential hills, the adaptive placing of the Gaussians and controlling the fictitious kinetic energy of the auxiliary degrees of freedom all follow our validated scheme.[9,10]. At variance with these previous simulations using OH^-(aq) as the attacking nucleophile, which can easily undergo Grotthuss–like structural diffusion[18,19] and thereby readily loose its nucleophilic character even being close to the disulfide bond, the DTT^-(aq) nucleophile does not need any external potential walls to prevent its protonation by water

III.   EXTENDED DISULFIDE MODEL ATTACKED BY DTT$^-$: COMPARISON TO MINIMAL DIETHYL DISULFIDE MODEL CLEAVED BY OH$^-$

In Figure 1, representative structures that characterize the reactants, the transition state structures and the final products are shown in thermal limit (corresponding to $F = 0.0$ nN)
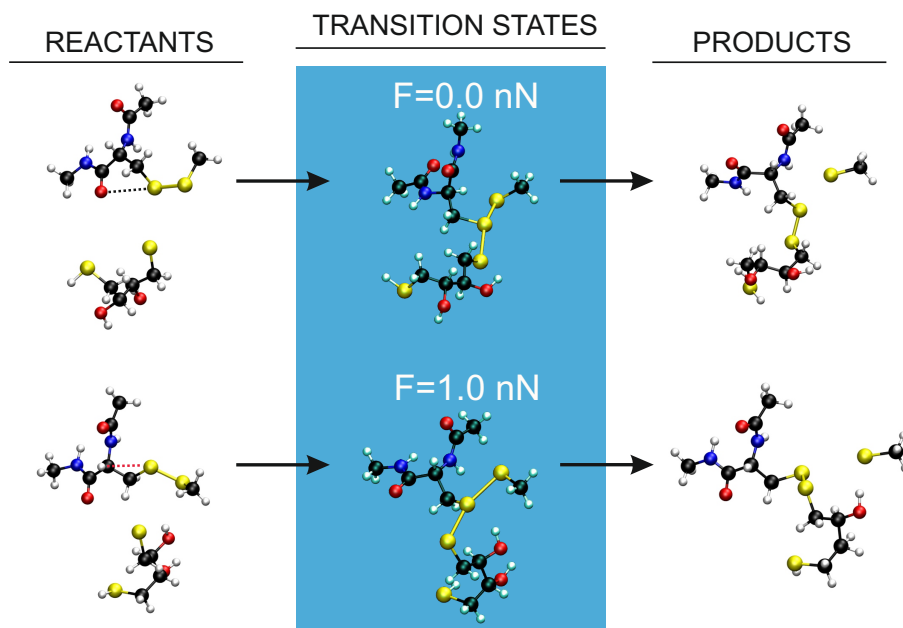
FIG. 1: Representative configuration snapshots of reactants, transition states and products along the reaction path for nucleophilic cleavage of the extended disulfide model (EDM) in bulk water by $DTT^-(aq)$ from isotensional *ab initio* metadynamics sampling at $T = 300$ K at zero force (top panels) and at a constant stretching force of $F = 1.0$ nN (bottom); note that none of the 112 explicit water molecules in the periodic supercell are depicted for clarity.

and subject to a constant external force of $F = 1.0$ nN as obtained from isotensional *ab initio* metadynamics simulations. For the reactant structure at zero force, the dotted line indicates schematically the quasi-linear chalcogen–chalcogen 1,5-type $S \cdots O$ bond that connects the γ–carbonyl oxygen of the protein mimic (EDM) to the nearby sulfur atom of the disulfide bridge. This type of noncovalent interaction vanishes upon applying even weak tensile forces and gets eventually replaced by a simple $S \cdots H–C$ contact caused by stretching EDM as shown for $F = 1.0$ nN in Figure 1. Additionally, it is demonstrated by the representative snapshot at $F = 1.0$ nN that the transition state structure features a H–bond, in which the donor is the closeby hydroxyl group of DTT^- and the acceptor is the attacked sulfur $S^\star$ within the disulfide bond of EDM.

The activation free energy as a function of the magnitude of the tensile force obtained for EDM reduction *via* DTT^- attack is compared in Figure 2a to that obtained earlier[9] for

the cleavage of the minimal diethyl disulfide model by hydroxide, $OH^-(aq)$. What is seen already at first glance is that $DTT^-(aq)$ is a far better nuclophile compared to $OH^-(aq)$ owing to the fact that displays much lower activation free energies at all applied forces, notably including also the zero force limit (see main text). The activation free energy is also more sensitive to mechanochemical activation of the reaction when using DTT^- as the reducing agent. In particular, already at a force of $F = 1.2$ nN, disulfide reduction with DTT^- starts to be an almost barrierless process, thus preventing its investigation by force–clamp single–molecule force spectroscopy experiments in this force regime, whereas OH^- attack remains to be subject to a sizable barrier even at 2 nN.

As explained in the main text, we were required to introduce a generalized molecular definition to distinguish disulfide configurations which are sterically blocked (called "closed") from those that are "open" to nucleophilic attack by the bulky DTT^- molecule. To this end, a molecular definition in terms of site classes is introduced: A specific atom X of the disulfide as represented by EDM (according to Figure 1 in the main text) is inside the reaction cone (thus blocking S_N2 attack at $S^\star$ by the nucleophile) if its $S^\star \cdots X$ distance does not exceed 4 Å and if its $S–S^\star \cdots X$ angle is in between 180 and 130 °. These EDM conformations are the "closed" ones, whereas all other cases correspond to "open" conformers. In panel b of Figure 2 we directly compare the simple proxy approach based on the S–S–C–C dihedral angle, as previously used[9,10] to study the attack of the minimal disulfide model by $OH^-(aq)$, to the molecularly detailed description of the reaction cone when $DTT^-(aq)$ attacks the extended disulfide model. The match is close to perfect, which implies that the molecular definition is able to describe whatever has been found before in case of the minimal model[9,10]. Transcending the simple proxy approach, the new molecular cone definition allows us to analyze the blocking atoms in full atomic resolution as unfolded in Figure 3, which allowed us to discover the new, third reactivity regime of disulfide reduction at low tensile forces as described in the main text.

IV. REACTION CONE ANALYSIS IN AQUEOUS SOLUTION *VERSUS* GAS PHASE

Having confirmed that the molecularly detailed description of the reaction cone is successful in reproducing previous conclusions obtained with the minimal disulfide model[9,10],
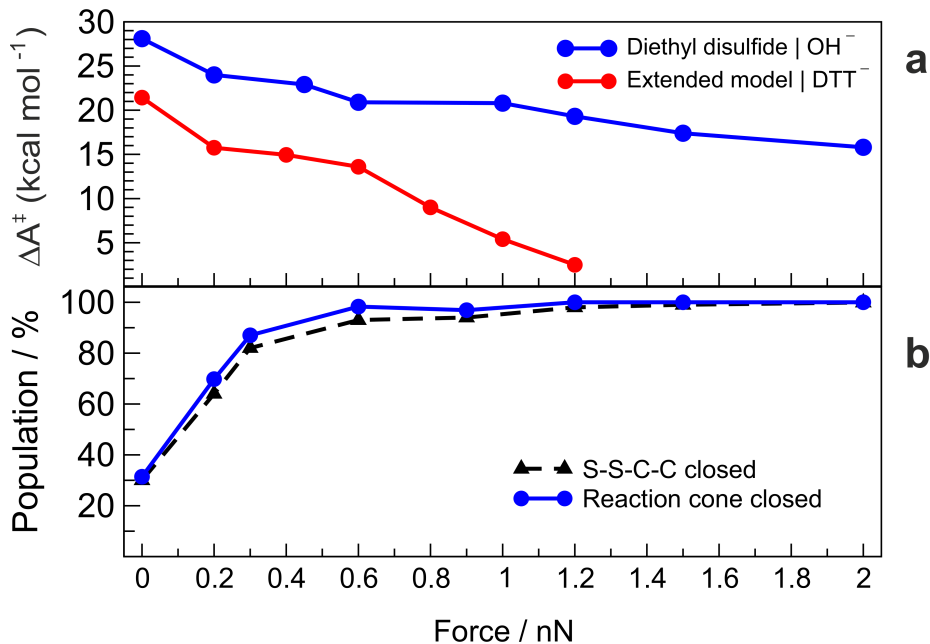
FIG. 2: (a) Activation free energies as a function of tensile force, $\Delta A^{\ddagger}(F)$, for nucleophilic cleavage of the extended disulfide model in bulk water by DTT^{-} nucleophile (red) compared to previous data[9] obtained from the minimal diethyl disulfide model cleaved by OH^{-} nucleophile (blue) in water all from isotensional *ab initio* metadynamics sampling at $T = 300$ K (solid lines connect the data linearly to guide the eye). (b) Force dependence of the relative population of the "closed" conformers (see text) in the reactant state for the minimal diethyl disulfide model[9] using (i) the simple proxy approach (see text; black triangles) based on the S–S–C–C dihedral angle as previously used[9] in direct comparison to (ii) the molecularly detailed new description of the reaction cone (see text; blue bullets) as applied to the same system for reference (see text). The procedure used to compute the populations according to reaction cone analysis is described in the last paragraph of Section III.

we can now apply it to the much more complex current situation, i.e. DTT^{-} reducing the EDM protein mimic in aqueous solution. In the left panels of Figure 3, the total population of closed conformations as a function is force (panel a1) is decomposed in panels b1 and c1 to all possible specific atoms and atom groups that can possibly block the reaction cone, and thereby shield the disulfide bridge against nucleophilic attack and thus against reduction. This most detailed analysis supports the more coarse-grained presentation in Figure 2d of
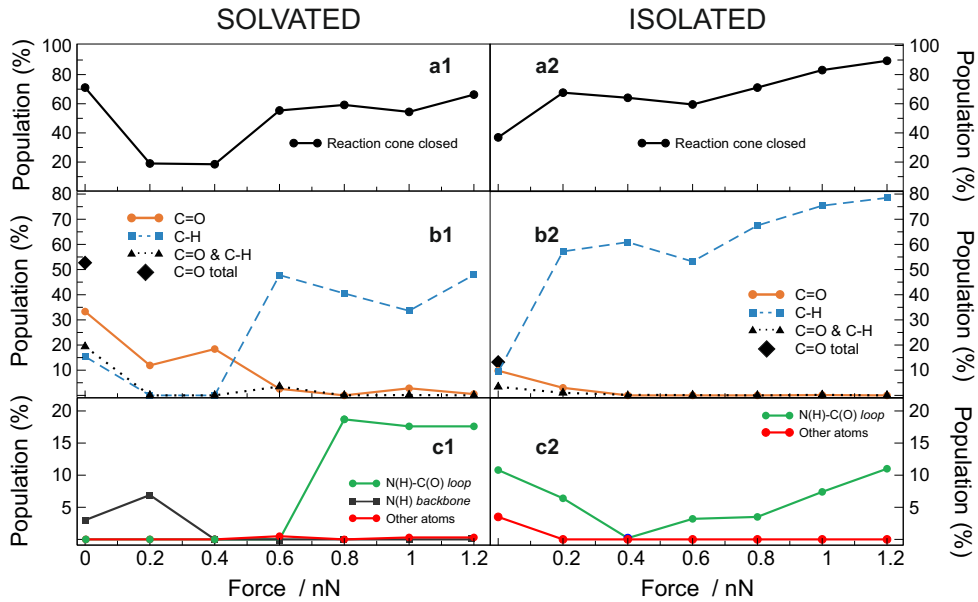
7

the main text.

FIG. 3: Analysis of overall (top panels) and atom-resolved (middle and bottom panels) blocking of the reaction cone in the reactant state (see caption of Figure 2 in the main text for details) as a function of tensile force using the molecular steric hindrance criterion (see text) applied to EDM conformations in aqueous solution (left panels) and in the gas phase reference case (right panels). Panels b depict blocking due to exclusively the C^γ=O group (orange circles and solid lines), exclusively the C^β–H group (blue squares and dashed lines), and both these groups at the same time (black triangles and dotted lines); the black diamond at $F = 0$ nN represents the total population of the C^γ=O group when blocking the cone. Panels c show blocking contributions due to the N(H)–C(O) group from the loop part of the protein mimic (green circles and solid lines), the N(H) group from the backbone part (black circles and solid lines), and any other atoms (red bullet and solid lines). See Figure 1 in the main text for atom and group labeling. The procedure used to compute the populations according to reaction cone analysis is described in the last paragraph of Section III.

More importantly, the same analysis of the relative contribution of closed conformations of EDM as a function of tensile force is also carried out in the dry, i.e. in the absence of solvation, as presented in the right panels of Figure 3 in one-to-one comparison to the solvated case to the left. As one can see in the top panels, the conformational force response

8

is drastically different in the gas phase compared to the liquid phase: compare panel a2 to a1 in Figure 3. Whereas the reaction cone is mostly closed in aqueous solution without applying force, i.e. in the thermal limit ($F = 0$ nN), it is mostly open without solvation by water molecules.

The atom-resolved analysis of cone blocking in Figure 3 discloses that at zero and low forces, it is the γ–carbonyl group, C^γ=O, that is mainly responsible to prevent nucleophilic attack of EDM by DTT$^-$ in aqueous solution, see panel b1. This is in stark contrast to the isolated reference case, where the C^γ=O contribution is essentially negligible not only at zero force, but also at all finite forces according to the data in panel b2. In the gas phase, instead, it is observed that that the C^β–H group is significantly blocking the reaction cone at forces of 0.2 nN and beyond. Refined analysis in panels c1 and c2 shows that contributions from other atoms and groups are below 20 % where the most pronounced effect is due to N(H)–C=O group from the loop of the protein mimic (highlighted in green in Figure 1 of the main text).

Overall, the conformational scenario of EDM in the absence of solvation water is characterized by (i) no steric hindrance of nucleophilic attack in the thermal limit and (ii) significant blocking of the respective reaction cone setting in already at small forces (of about 0.2 nN) which is entirely due to the C^β–H group. This behavior does not share any similarity with the scenario observed in aqueous solution, which is characterized by (i) significant steric blocking of nucleophilic attack in the thermal limit, (ii) opening of the respective reaction cone in the small force regime (of roughly 0.2 to 0.4 nN), and (iii) re-closing of the cone setting in at intermediate forces (of about 0.6 nN) and persisting up to the largest forces. The explanation of the unexpected blocking behavior in aqueous solution in the zero force limit is traced back to the existence of a chalcogen bond, being a 1,5-type S$\cdots$O noncovalent interaction.

V.   ANALYSIS OF S$\cdots$O CHALCOGEN BONDING: AQUEOUS SOLUTION *VERSUS* GAS PHASE

The atom–resolved analysis of reaction cone in aqueous solution (see Figure 2d in the main text and left panels of Figure 3) has disclosed that the γ–carbonyl group, C^γ=O, blocks the disulfide bridge against nucleophilic attack in the liquid state at zero and small forces,

whereas this phenomenon is not at all observed in the gas phase (see Figure 3 in the main text and right panels of Figure 3).
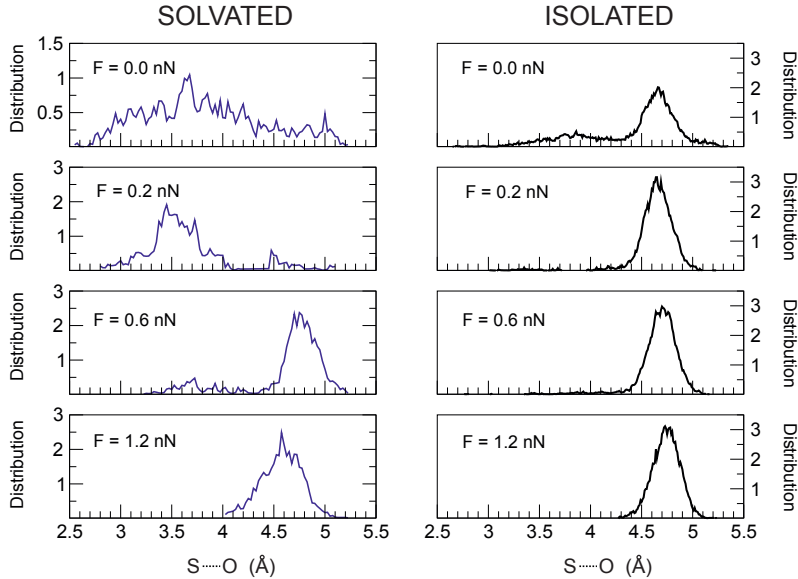

FIG. 4: Normalized probability distribution functions of the noncovalent $S^{\star}\cdots O$ distance within the $S–S^{\star}\cdots O=C^{\gamma}$ arrangement (cf. Figure 1 in the main text) as a function of tensile force. The data in aqueous solution (left panels) have been obtained from *ab initio* metadynamics simulations in the reactant state (see caption of Figure 2 in the main text for details) and those in the gas phase (right panels) are based on canonical isotensional *ab initio* molecular dynamics also at 300 K.

This suggests to investigate in more detail how the corresponding $S^{\star}\cdots O$ distance behaves in the liquid and gas phase as a function of force, which is depicted in Figure 4 in the left and right panels, respectively. At zero force, the solvated EDM is found to feature a broad peak which is centered at short $S^{\star}\cdots O$ distances of approximately 3.6 Å. The distribution of $S^{\star}\cdots O$ distances in the absence of solvation water is distinctly different: It is characterized by a pronounced peak at a much larger most probable $S^{\star}\cdots O$ distance of roughly 4.7 Å for EDM in the gas phase at $F = 0$ nN. Still, in the absence of force, the AIMD simulation is seen to clearly sample distances down to only 3 Å also in vacuum, which however contribute overall only to a negligible extent. In the gas phase, that picture is force-independent and gets even more pronounced in the sense that $S^{\star}\cdots O$ distances

smaller than about 4.5 Å are no longer found at a tensile force of 1.2 nN, thus sharpening the peak at a most probable $S^\star \cdots O$ distance as large as about 4.8 Å.

In aqueous solution, the picture of how the noncovalent $S^\star \cdots O$ interaction changes as a function of finite tensile force is very rich and distinctly different from the essentially force-independent behavior found in the gas phase according to Figure 4. At a small tensile force of 0.2 N, the distribution function becomes much more structured than at zero force and now features a pronounced peak at roughly 3.5 Å. However, upon isotensional stretching with an intermediate force of 0.6 nN, only a little bump remains at these small $S^\star \cdots O$ distances, whereas a pronounced peak is now found at about 4.7–4.8 Å much like in the gas phase. At the largest force, 1.2 nN, that peak shifts to slightly smaller distances, but there are no longer any $S^\star \cdots O$ contacts found below a distance of roughly 4 Å.

How can all these puzzling observations be possibly understood and rationalized? Attractive interactions between divalent sulfur atoms and nearby chalcogen atoms, including notably oxygen atoms, have been studied since many decades and are often denoted as chalcogen bonds / bonding.[21–23] Chalcogen bonds are noncovalent interactions, much like the better known hydrogen bonds or also halogen bonds, and are increasingly used in synthetic and supramolecular chemistry.[21–23] Moreover, they are known to play important functional roles in protein stabilization and biomolecular recognition[24–27].
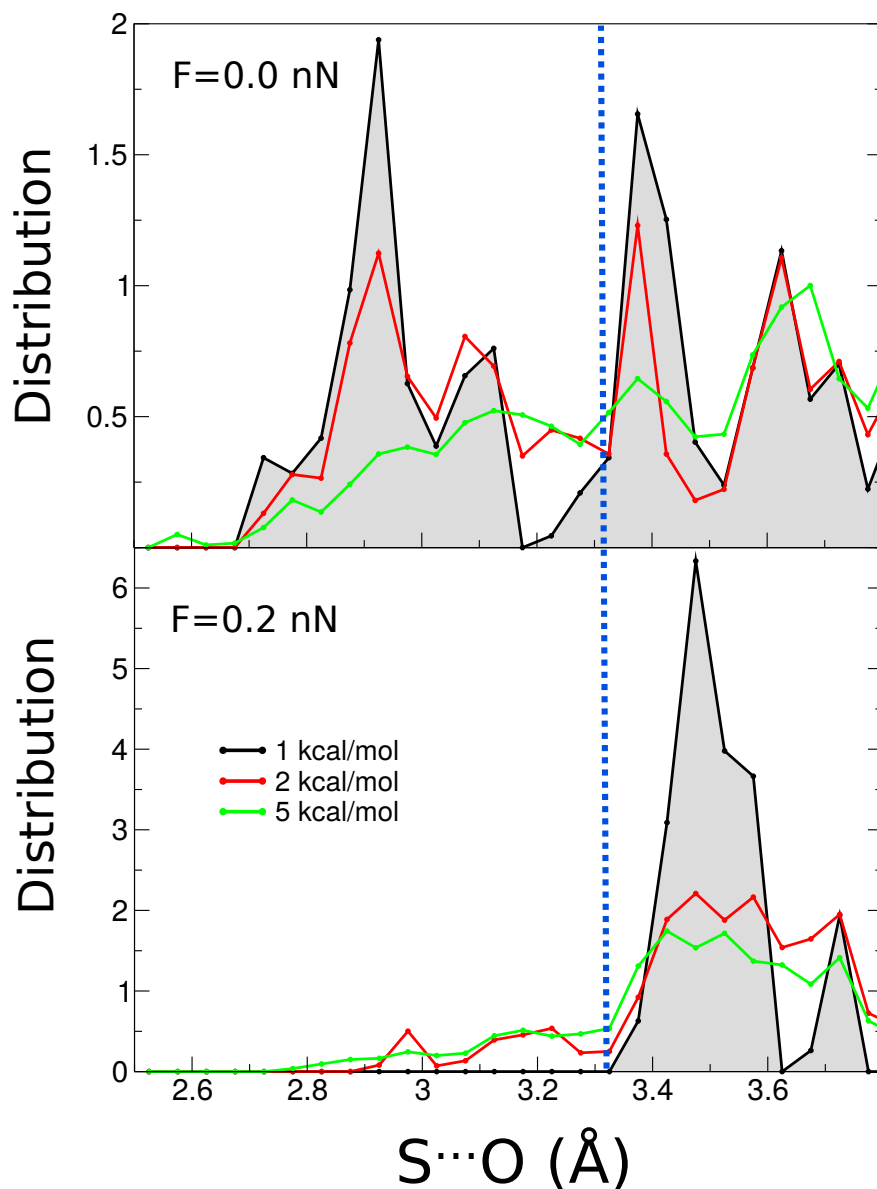
FIG. 5: Normalized probability distribution functions of the noncovalent $S^\star \cdots O$ distance within the $S\text{--}S^\star \cdots O\text{=}C^\gamma$ arrangement (cf. Figure 1 in the main text) in the absence of force (top panel) and with an applied force of $F = 0.2$ nN (bottom). The vertical dotted line marks the most conservative (i.e. smallest) sum of S and O van der Waals radii as discussed in the text, implying that $S^\star \cdots O$ chalcogen bonding is observed to the left of that separatrix according to the IUPAC definition[20] from 2019. The distributions have been computed using the configurations generated by ab initio metadynamics of the full EDM/DTT$^-$/water system deep in the free energy basin corresponding to the reactant state (according to the collective variable subspace) by considering cartesian structures therein up to 1 kcal/mol (black), 2 kcal/mol (red) and 5 kcal/mol (green) above the free energy minimum.

According to its IUPAC definition[20] proposed in 2019, one of the features of a typical chalcogen bond (ChB) complex is that the intermolecular distance between the ChB donor atom (Ch) and the nucleophilic site in the acceptor (A) tends to be less than the sum of the van der Waals radii. The problem with this definition is the fundamentally ill-defined nature of van der Waals (vdW) radii combined with their variability within a few of tenths of an Angstrom depending on various structural factors[28]. A comprehensive table of vdW radii was compiled by Bondi[29] in 1964 and remains heavily used up to the present day. Our atoms of interest are sulfur and oxygen and using Bondi's vdW radii we arrive at a value of $(1.80 + 1.52 = 3.32)$ Å. Using instead the recommended crystallographic values yields $(1.80 + 1.55 = 3.35)$ Å, whereas the isolated atom sum provides a value of $(2.06 + 1.71 = 3.77)$ Å.[30] Thus, the conclusion is that it is impossible to define unequivocally a unique value for sum of the two vdW radii and, by inference, the chalcogen bonding according to IUPAC. In order to circumvent that dilemma in the present case where thermal fluctuations in the aqueous phase moreover introduce pronounced intermolecular distance variations in thermal equilibrium, we computed the $S^\star \cdots O$ distance distribution functions depicted in Figure 5. They are obtained based on sampling protein mimic (EDM) molecular structures from the free energy basin corresponding to the reactant state (of the disulfide cleavage reaction) by considering energy ranges up to 1 kcal/mol (black), 2 kcal/mol (red) and 5 kcal/mol (green) above the free energy minimum. Clearly, the lower that energy range taken into account in the analysis is, the closer that structural ensemble is to something that could be called the "representative protein mimic structure" in the aqueous solution at 300 K (whereas the structure at the absolute free energy minimum would correspond to the "most probable protein mimic structure" in aqueous solution at 300 K). In the absence of force (top panel), there is a significant population of $S^\star \cdots O$ distances within the protein mimic that are well below the most conservative estimate for chalcogen bonding which is marked by the vertical dotted line; integration of the respective probability distributions provides corresponding populations of 33, 32, and 21 % when using energy thresholds of 1, 2, and 5 kcal/mol. Evidently, that scenario significantly changes when applying even a tiny tensile force (bottom panel) which close to eliminates $S^\star \cdots O$ chalcogen bonded protein mimic structures; the corresponding populations are 0, 12, and 15 % considering relative energies up to 1, 2, and 5 kcal/mol. This structural analysis, thus, clearly supports the discussion in the main text as to the presence of significant chalcogen bonding (according

to the current IUPAC definition) within the protein mimic (EDM) in the thermal limit as seen for $F = 0.0$ nN in panel (a).

In addition to such bare interatomic Ch$\cdots$B distances, an important property of chalcogen bonds is the peculiar directionality of such noncovalent S$\cdots$O interactions with respect to the local covalently bonded environment. Detailed analyses of crystal structures[31] revealed that this interaction is characterized by different preferred relative orientations with respect to the O=C–C– plane in small organic sulfur compounds compared to proteins as sketched in the top and bottom inset, respectively, of Figure 6. In proteins, the S$\cdots$O=C–C– dihedral angle typically adopts values around $\pm 90°$ and, thereby, leads to a preferred perpendicular orientation of the noncovalent S$\cdots$O bonds w.r.t. the carbonyl groups (bottom inset and dotted green lines); note that this dihedral corresponds to the S$^\star\cdots$O=C$^\gamma$–C$^\beta$– grouping in case of the EDM protein mimic according to the labeling introduced in Figure 1 of the main text. In contrast, small organic compounds prefer this dihedral angle to be around $0°$, thus stabilizing coplanar S$\cdots$O=C–C– structural motifs (top inset and dotted red lines in Figure 6).

According to our simulation results of EDM in aqueous solution, there is, first of all, a clear propensity to stabilize such noncovalent S$^\star\cdots$O bonds exclusively in the limit of zero and small forces with bond lengths of roughly 3.5 Å as demonstrated by the $F = 0.0$ and 0.2 nN data in the left panels of Figure 4 as well as at a more refined level in Figure 5 (note the differences in the respective S$^\star\cdots$O distance scales). Intermediate and, even more for larger tensile forces, of 0.6 and 1.2 nN, respectively, counteract this attractive interaction and thus break these noncovalent bonds as seen by favoring large disulfide sulfur to carbonyl oxygen separations that exceed 4.5 Å in solution. The orientational analysis in Figure 6 discloses that the S$^\star\cdots$O=C$^\gamma$–C$^\beta$– dihedral angle vividly fluctuates between the two limiting values, i.e. ± 90 and $0°$, even in the limit of short such chalcogen–chalcogen 1,5-type S$\cdots$O bonds in aqueous solution. This might not be too surprising since the S$\cdots$O bond in EDM is established between a sulfur atom of the disulfide bridge and the oxygen atom of the nearby γ–carbonyl group, whereas this needs not to be so in general for proteins where any far–distant divalent sulfur atom can form a noncovalent S$\cdots$O bond with exposed carbonyl groups thus providing more local orientational flexibility.

Yet, independently from the particular orientation of the noncovalent S$^\star\cdots$O bond with respect to the γ–carbonyl group of EDM in water, it is clear that the resulting S–S$^\star\cdots$O
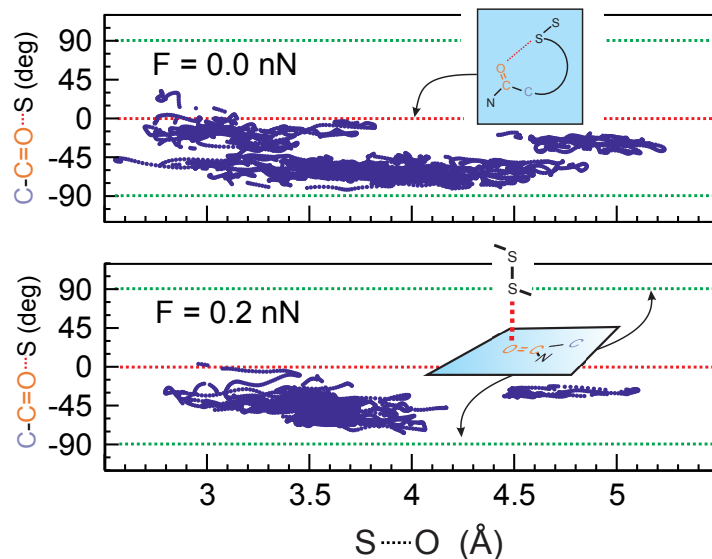
14

FIG. 6: Structural correlation of the $S^\star \cdots O$ distance and the $S^\star \cdots O=C^\gamma-C^\beta-$ dihedral angle extracted from *ab initio* metadynamics sampling in the reactant state (see caption of Figure 2 in the main text for details) in the zero force limit (top) and at a small force, $F = 0.2$ nN (bottom). The ideal coplanar and perpendicular orientations of such chalcogen bonds (see text) are shown schematically as top and bottom insets, where the dihedral angle corresponds to $0°$ and $\pm 90°$ (marked using red and green dotted lines, respectively).

arrangement (as depicted by the top–left snapshot of Figure 2 in the main text) strongly hinders collinear attack of $S^\star$ by the nucleophile, being S^- of DTT^-(aq), which is the required transition state for disulfide reduction and thus disulfide bond cleavage as shown in the inset of Figure 2b in the main text. In other words: The 1,5-type $S \cdots O$ chalcogen bond blocks the reaction cone at zero and small tensile forces in aqueous solution, whereas this noncovalent bond and thus the induced steric hindrance is no longer effective already when reaching the intermediate force regime.

* Electronic address: `przemyslaw.dopieralski@chem.uni.wroc.pl`

[1] D. Marx and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*, Cambridge University Press, Cambridge, 2009.

[2] R. Car and M. Parrinello, *Phys. Rev. Lett.*, 1985, **55**, 2471–2474.

[3] J. Hutter and *et al.*, CPMD *Program Package*.

[4] A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.

[5] C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

[6] S. Grimme, *J. Comput. Chem.*, **27**, 1787–1799.

[7] N. Troullier and J. L. Martins, *Phys. Rev. B*, 1991, **43**, 1993–2006.

[8] G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635–2643.

[9] P. Dopieralski, J. Ribas-Arino, P. Anjukandi, M. Krupicka, J. Kiss and D. Marx, *Nat. Chem.*, 2013, **5**, 685–691.

[10] P. Dopieralski, J. Ribas-Arino, P. Anjukandi, M. Krupicka and D. Marx, *Nat. Chem.*, 2017, **9**, 164–170.

[11] P. Dopieralski, J. Ribas-Arino and D. Marx, *Angew. Chem. Int. Ed.*, 2011, **50**, 7105–7108.

[12] A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 12562–12566.

[13] M. Iannuzzi, A. Laio and M. Parrinello, *Phys. Rev. Lett.*, 2003, **90**, 238302–1–4.

[14] A. Laio and M. Parrinello, Computer simulations in condensed matter: From materials to chemical biology, Berlin Heidelberg, 2006, pp. 315–347.

[15] B. Ensing, A. Laio, M. Parrinello and M. L. Klein, *J. Phys. Chem. B*, 2005, **109**, 6676–6687.

[16] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti and M. Parrinello, *J. Phys. Chem. B*, 2006, **110**, 3533–3539.

[17] N. N. Nair, E. Schreiner and D. Marx, *inSiDE*, 2008, **6**, 30–35.

[18] M. E. Tuckerman, D. Marx and M. Parrinello, *Nature*, 2002, **417**, 925–929.

[19] D. Marx, A. Chandra and M. E. Tuckerman, *Chem. Rev.*, 2010, **110**, 2174–2216.

[20] C. B. Aakeroy, D. L. Bryce, G. R. Desiraju, A. Frontera, A. C. Legon, F. Nicotra, K. Rissanen, S. Scheiner, G. Terraneo, P. Metrangolo and G. Resnati, *Pure Appl. Chem.*, 2019, **91**, 1889–1892.

21 S. Benz, J. López-Andarias, J. Mareda, N. Sakai and S. Matile, *Angew. Chem. Int. Ed.*, 2017, **56**, 812–815.

22 R. Gleiter, G. Haberhauer, D. B. Werz, F. Rominger and C. Bleiholder, *Chem. Rev.*, 2018, **118**, 2010–2041.

23 L. Vogel, P. Wonner and S. M. Huber, *Angew. Chem. Int. Ed.*, 2019, **58**, 1880–1891.

24 M. Iwaoka, S. Takemoto, M. Okada and S. Tomoda, *Chem. Lett.*, 2001, **30**, 132–133.

25 R. J. Fick, G. M. Kroner, B. Nepal, R. Magnani, S. Horowitz, R. L. Houtz, S. Scheiner and R. C. Trievel, *ACS Chem. Biol.*, 2016, **11**, 748–754.

26 M. O. Mitchell, *J. Mol. Model.*, 2017, **23**, 287.

27 K. Kriz, J. Fanfrlik and M. Lepsik, *ChemPhysChem*, 2018, **19**, 2540–2548.

28 S. S. Batsanov, *Inorg. Mater.*, 2001, **37**, 871–885.

29 A. Bondi, *J. Phys. Chem.*, 1964, **68**, 441–451.

30 N. L. Allinger, X. Zhou and J. Bergsma, *J. Molec. Struc. THEOCHEM*, 1994, **312**, 69 – 83.

31 M. Iwaoka and N. Isozumi, *Molecules*, 2012, **17**, 7266–7283.