Supplementary Information

Nucleation of mercury sulfide by dealkylation

Mironel Enescu, Kathryn L. Nagy & Alain Manceau

1. Evaluation of the Gibbs free energy

2. Energy profile for the Intermediate Product to Product Complex structural rearrangement

3. Effect of explicit water molecules on the dealkylation energy barrier

4. Correction of the contribution of translation-rotation solvation entropy to the energy barrier

5. Cleavage of the S-C bond by direct insertion of a Hg atom

6. References

7. Gaussian 09 input files and optimized Cartesian coordinates

1. Evaluation of the Gibbs free energy

Stationary points of the reaction pathway were identified by scanning the potential energy surface (PES) against the location of all atoms in the system without water molecules. The Gibbs free energy in aqueous solution (G_{aq}) for the molecular state (M) at each stationary point was evaluated according to:

 $G_{\text{aq}}(M) = E_{\text{CCSD(T)}}(M) + \Delta G_{\text{corr}}(M),$ (1)

where $E_{\text{CCSD(T)}}(M)$ is the total electronic energy (i.e., potential energy) including the solvent effects, and $\Delta G_{\text{corr}}(M)$ is the thermal and entropic contributions to *G* calculated at 298 K and 1 atm. The thermal and entropic contributions were obtained by performing frequency calculations. The *G*aq(M) values were corrected for the difference between the standard molar concentrations in solution and gas phase. Since the CPCM solvation energy does not include the energy of formation of solvent cavities, we evaluated this contribution separately for the final structures using the SMD version¹ of the polarizable continuum dielectric model (PCM). The correction for this effect on the calculated energy barriers was small, in the range of 1-2 kcal mol⁻¹. The conformational degeneracy of the transition state was taken into account by adding to *G* the term *-RT***ln(g)*, where *g* is the degeneracy, *R* the gas constant, and *T* the absolute temperature.

Energetic values in the presence of water molecules were calculated as follows. The geometry of the stationary points were reoptimized with explicit water molecules added in contact to the main atoms involved in the chemical reaction. The whole system ("real system"²) was optimized at the MP2 level of theory, and the "model system"² including all atoms except those belonging to water molecules was treated at the CCSD(T) level of theory. The ONIOM energy^{2,3} was calculated according to:

 $E_{ONIOM} = E_{CCSD(T)}(model) - E_{MP2}(model) + E_{MP2}(real).$ (2)

The three energy terms in eq. 2 were calculated in the presence of the bulk solvent (CPCM) and for the geometry optimized on the real system. Finally, the Gibbs free energy of the system was calculated from:

 $G_{aq} = E_{ONIOM} + \Delta G_{corr}(real)$ (1')

with Δ*G*_{corr}(real) obtained by a frequency calculation performed at the MP2 level of theory on the optimized real system (i.e., in the presence of the bulk solvent).

The choice of the pure continuum solvation model has only a limited impact on the calculated Gibbs free energy. For instance, we previously obtained⁴ $\Delta G(\text{FP}) = -3.9$ kcal mol⁻¹ for reaction (1) (main text) using the pure PCM model, compared to -6.4 kcal mol^{-1} with the pure CPCM model in the present study (Fig. 1).

The convergence criteria used in the geometry optimizations were the default settings in the Gaussian 09 program :

2. Energy profile for the IP to PC structural rearrangement

The energy profile was obtained by constraining the Hg_a-S_d distance to fixed values. The solvation model used here is the pure CPCM model (i.e., no explicit water molecules).

Figure S1. Energy profile for the rearrangement of the reaction product.

3. Effect of explicit water molecules on the dealkylation energy barrier

The transfer of the CH_3 ⁺ group is associated with a significant reorganization of the electronic density around S_d , S_a , and the C atom of the transferred methyl group. This process is expected to be sensitive to the presence of solute-solvent hydrogen bonds because of their weak but definite covalent character⁵. Therefore, a more accurate evaluation of the reaction energy barrier is obtained when explicit water molecules are placed nearby the main atoms involved in the reaction⁶. This type of calculation was performed for 2, 4 and 7 explicit solvent molecules. Obviously, in all cases, the atoms from the solute around which the explicit water molecules were placed were the same in the free reactant (FR) and transition state (TS) structures (Table S1). In a first approach (2H2O model in Fig. 1), two water molecules were placed in contact with the S_d atom of the two free reactants and of the TS complex. The length of the two S_d ... H hydrogen bonds in the optimized structures decreased from 2.37 Å in the free reactants to 2.24 Å in the TS complex, thus proving the importance of solute-solvent covalent interactions. In a second approach (4H2O model in Fig. 1), two additional solvent molecules were placed in contact with the transferred CH³ group because it is expected to become more electrophilic in the TS configuration. In a third approach (7H2O model in Fig. 1), the water molecules were placed as follows: three near Sd, two near Sa and two near the CH³ group. This arrangement approximately corresponds to the first solvation shell of the directly

reacting atoms. In all model systems, the free reactant state is represented by two free supermolecules, each formed of a Hg(II)-methanethiolate complex (CH3S-Hg-SCH3) carrying several water molecules (0 and 2, 0 and 4, 2 and 5, respectively). The structure of the TS complex including the seven water molecules is shown in Fig. S2. Details of the reaction energetics are given in Tables S1-S3. *E* is the total electronic energy, *H* the enthalpy, *S*tot the total entropy, and *S*tr the translation-rotation entropy.

Table S2. Thermodynamic parameters (in kcal mol⁻¹) at 298 K and 1 atm for reaction (1) calculated using four explicit water molecules.

Structures		\mathbf{r} AП	$T\Delta S_{\rm tot}$	$T\Delta S_{\text{tr}}$	ΔG
FR					
TC 1 N	$\overline{\Omega} \cap \overline{\Omega}$ $\angle U.$	າະ ں. ر∠	-14.6	-19.5	34.7
FP	-8.0	-8.4	- .	-0.7	-10.5

Table S3. Single point energies (in Hartree) for pure CPCM calculations (*E*(ccsd(t)) and for supermolecule calculations with seven water molecules (*E*(ONION)).

Figure S2. Structure of the TS configuration optimized with seven explicit water molecules

4. Correction of the contribution of translation-rotation solvation entropy to the energy barrier

A general source of error attributed to the continuum solvent models is the improper evaluation of the solvation entropy⁷⁻¹⁰. When the translation-rotation entropy of the molecule in solution is calculated using the same method as in the gas phase, this entropy is largely overestimated. The result is a systematic error in the evaluation of the Gibbs free energy difference between molecular states involving a different number of molecules. For instance, this error occurs systematically in the calculation of the Gibbs free energy of molecular association and dissociation. In the case of an association, the reactants have six translational and six rotational degrees of freedom in total, which transform in the product to three translational and three rotational degrees of freedom plus six new vibrational degrees of freedom. To avoid this error, $Wertz^7$ proposed an alternative semi-empirical mode of calculation of the translation-rotation entropy for a molecule in solution which consists of attributing to the dissolved molecule an energy equivalent to one-half the gas phase value (see also ref. 10). In the present study, the contribution of the translation-rotation entropy to the reaction energy barrier is about 20 kcal mol⁻¹. According to Wertz's approach, it should be about half this value. In another approach, based on the analysis of data from the association of two weakly interacting molecules in aqueous solution, Yu et al. 9 concluded that the contribution of the translation-rotation entropy to the Gibbs free energy of association is less or equal to *9RT.* From this approach, our calculated reaction energy barrier should be corrected by at least -14.6 kcal mol⁻¹. By corroborating the two approaches, we conclude that a minimum correction of -10 kcal mol $^{-1}$ should be applied to the calculated energy barrier. The resulting predictive value for this barrier is 22 kcal $mol⁻¹$.

5. Cleavage of the S-C bond by direct insertion of a Hg atom

The alkyl group transfer between two thiolates could also occur indirectly through the formation of an intermediate product, in which mercury is inserted between the two atoms of the S-C bond. A first possibility is an insertion that occurs by an intramolecular rearrangement in a RS- Hg^* -S^{*}R^{*} complex: Hg^* is inserted between S^{*} and R^{*}, the R^{*} group is transferred to SR, and the resulting R*SR group is eliminated. However, the scan of the potential energy surface of this system with respect to the Hg*-C* (from R*) distance (results not shown) invalidates this reaction pathway because it yields a significant increase of energy. Besides, a direct consequence of an insertion reaction should be the formation of an intermediate product in which the C* and S* atoms are both bonded to Hg*. No energy minimum could be found for this molecular conformation.

A second possibility is an insertion reaction involving two RS-Hg-SR complexes. In this case, the Hg atom from one complex interacts with a thiolate group from the other complex. The corresponding insertion product was identified and its optimized structure shown in Figure S3. In a second reaction step, the CH³ group bonded to Hg could be transferred to an adjacent S atom. The CH3-S-CH³ group thus formed could dissociate as in the case of the direct alkyl group transfer channel shown in Figure 1. The global reaction also is described by reaction (1) (main text). The difficulty with this model is that the Gibbs free energy of the insertion product is too high relative to that of the free reactants. A calculation performed at the ONIOM (CCSD(T)/MP2) level of theory using four explicit water molecules gives a positive difference of 49.0 kcal mol⁻¹. This is about 15 kcal mol⁻¹ higher than the energy barrier for the direct alkyl group transfer calculated under the same conditions $(34.7 \text{ kcal mol}^{-1})$. Obviously, the transition state leading to the formation of this insertion product should be even higher in energy. We conclude that this pathway should be disregarded.

Figure S3. Formation of a Hg(II)-sulfide dimer by an insertion reaction; shown here is the intermediate product. Atoms directly involved in the reaction are marked with a star.

6. References

- 1 Marenich, A. V., Cramer, C. J. & Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **113**, 6378-6396 (2009).
- 2 Svensson, M. *et al.* ONIOM: A multilayered integrated MO+MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and Pt(P(*t*-Bu)3)2 + H² oxidative addition. *J. Phys. Chem.* **100**, 19357-19363 (1996).
- 3 Dapprich, S., Komaromi, I., Byun, K. S., Morokuma, K. & Frisch, M. J. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Mol. Str. - Theochem.* **461**, 1-21 (1999).
- 4 Manceau, A. *et al.* Formation of mercury sulfide from Hg(II)-thiolate complexes in natural organic matter. *Environ. Sci. Technol.* **49**, 9787−9796 (2015).
- 5 Reed, A. E., Curtiss, L. A. & Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **88**, 899-926 (1988).
- 6 Afaneh, A. T., Schreckenbach, G. & Wang, F. Y. Theoretical study of the formation of mercury $(Hg²⁺)$ complexes in solution using an explicit solvation shell in implicit solvent calculations. *J. Phys. Chem. B* **118**, 11271-11283 (2014).
- 7 Wertz, D. H. Relationship between the gas-phase entropies of molecules and their entropies of

solvation in water and 1-octanol. *J. Am. Chem. Soc.* **102**, 5316-5322 (1980).

- 8 Giesen, D. J., Cramer, C. J. & Truhlar, D. G. Entropic contributions to free-energies of solvation. *J. Phys. Chem.* **98**, 4141-4147 (1994).
- 9 Yu, Y. B., Privalov, P. L. & Hodges, R. S. Contribution of translational and rotational motions to molecular association in aqueous solution. *Biophys. J.* **81**, 1632-1642 (2001).
- 10 Lau, J. K. C. & Deubel, D. V. Hydrolysis of the anticancer drug cisplatin: Pitfalls in the interpretation of quantum chemical calculations. *J. Chem. Theor. Comput.* **2**, 103-106 (2006).

7. Gaussian 09 input files and optimized Cartesian coordinates

```
7.1. Optimization of the TS state with the pure cpcm solvation model
```

```
%nprocshared=8
%mem=16GB
%chk=ts4c.chk
# opt(ts, readfc, maxstep=4, noeigentest) mp2/genecp scrf=cpcm
Title Card Required
0 1
C 2.07720800 3.04341800 1.22928400
S 1.82988400 2.39719100 -0.46252900
C 4.30532800 -2.28022600 0.09052300
S 2.48251000 -2.16510300 0.28647500
Hg  2.15686200  0.10166200  -0.08937000
H 1.40098300 2.53401800 1.93050600
H 3.12043400 2.88627200 1.53349300
H 4.60970700 -1.96452200 -0.91610400
H 4.82062700 -1.67279600 0.84604100
H 1.85427000 4.11962400 1.20617200
H 4.56884200 -3.33862400 0.23341300
C = -1.17498900 -2.77124600 -1.05433800S -2.16838200 -2.40543800 0.45058300
C -0.43217700 2.34743200 -0.39315100
S -2.86573900 2.11796300 -0.26130700
Hg -2.57088800 -0.12888900 0.10125600
H -1.74005700 -2.51785200 -1.96203500
H -0.22806100 -2.21264000 -1.04001300
H -0.53722500 3.42896500 -0.45264500
H -0.40721700 1.86197400 0.57608000
H -0.95805100 -3.84966300 -1.05261900
H -0.48674300 1.75331700 -1.29862200
H C 0
aug-cc-pvdz
****
S \capaug-cc-pvtz
****
H\sigma 0
sdd
****
Hg 0
F 1 1.00
    0.54500000 1.0000000
F 1 1.00
    1.5800000 1.0000000
****
HG 0
sdd
```

```
7.2. Optimization of the IP state with the pure cpcm solvation model
```
%nprocshared=8 %mem=16GB %chk=ts4c_int1.chk # opt mp2/genecp scrf=cpcm Title Card Required 0 1 C 0.17980900 2.84791100 -1.20214200 S -0.93138800 2.63841200 0.21830400 C -4.49933200 -1.16467600 -0.06311800 S -2.74973700 -1.66895500 -0.33068700 Hg -1.69904100 0.36940200 -0.02968400 H 0.97657400 2.09196500 -1.15405400 H -0.43793300 2.73621300 -2.10346200 H -4.64315400 -0.77901300 0.95523700 H -4.80879500 -0.40723400 -0.79560400 H 0.59586300 3.86366900 -1.14762700 H -5.11037000 -2.06976700 -0.19654700 C 0.18292000 -2.60644700 1.50827200 S 0.70851900 -2.45214800 -0.24441800 C 0.25775500 2.57738700 1.58676900 S 3.45318900 1.27025600 0.01295500 Hg 2.04151800 -0.51607000 -0.12203100 H 1.05558500 -2.66526300 2.17380100 H -0.45389000 -1.75866400 1.79678400 H 0.63005000 3.59990500 1.74182900 H 1.07893400 1.89574200 1.32287200 H -0.40267000 -3.53393300 1.59542300 H -0.29461500 2.23368800 2.47142800 H C 0 aug-cc-pvdz **** $S \qquad 0$ aug-cc-pvtz **** $H\sigma$ 0 sdd **** Hg 0 F 1 1.00 0.54500000 1.0000000 F 1 1.00 1.5800000 1.0000000 **** HG 0 sdd

7.3. Optimization of the PC state with the pure cpcm solvation model

```
%nprocshared=8
%mem=16GB
%chk=pc4b.chk
# opt mp2/genecp scrf=cpcm 
Title Card Required
0 1
 C -1.23334900 -1.06068600 2.58002000
 S -2.72630800 -1.39188500 1.60891200
C -0.60489600 3.43738400 -0.00182500
```

S	-1.88522600	2.21149500	0.49138600
Hg	-1.09706500	0.36485700	-0.71071500
Η	-0.35698600	-1.01946900	1.91490400
Η	-1.36739200	-0.08158700	3.06104700
Η	0.40385800	3.04474100	0.19312600
Η	-0.69701700	3.69107700	-1.06697800
Η	-1.09311500	-1.83885300	3.34556500
Η	-0.76577400	4.34230100	0.60261100
C	1.79532400	1.16188300	2.50166200
$\rm S$	3.04327500	0.52267000	1.31048200
C	-2.20527000	-2.95319600	0.84822300
S	0.09343400	-1.31165900	-1.83681900
Hg	1.62995300	-0.44296200	-0.27817700
Η	0.99319300	1.70360100	1.98087000
Η	1.35999500	0.33915400	3.08530400
Η	-2.00372300	-3.70807000	1.62332200
Η	-1.30941700	-2.77758400	0.23216600
Η	2.32089700	1.84897700	3.18110500
Η	-3.02923200	-3.29813900	0.20646700
H C O **** $S \t 0$ $***$ * Hq 0 sdd $***$ * * * Hq 0 F 1 1.00	aug-cc-pvdz aug-cc-pvtz		
F	0.54500000 1.00 $\mathbf{1}$	1.0000000	
	1.5800000	1.0000000	
$***$ *			
HG sdd	0		

7.4. Optimization of the TS state with the cpcm solvation model and 4 explicit water molecules %nprocshared=12

```
%mem=44GB
%chk=ts4c 4w.chk
# opt(ts, modredundant, noeigentest, maxstep=4) mp2/genecp scrf=cpcm
Title Card Required
0<sub>1</sub>\mathsf{C}2.12227700
                 -3.16317800 -0.870315002.18309800
                 -2.289062000.73496000
\rm S0.93857900
\mathsf C2.67192900
                    3.02234400
\mathbf S2.65507700
                    2.15417900
                                -0.68201500Hq
      2.44573200
                   -0.050295000.01917100
\rm H1.36250400
                   -2.70540400-1.51940500-1.35003300\rm H3.10870800
                  -3.12149300\rm H1.75225900
                    2.81650600
                                  1,50201000
\rm H3.54629500
                    2.72069600
                                  1.53044700
H1.85565500
                  -4.20849900-0.65709100H2.73664100
                    4.09761200
                                  0.71674400
    -1.033106003.12885600
                                 -0.25966400\mathsf{C}-1.367190002.15371900
S-2.128282000.90551100
                 -2.12341100\mathcal{C}-0.09046400-1.88538700S-2.500215000.89778000
                 0.16227100 - 0.15954200Hq -2.39404900
```


7.5. Optimization of the IP state with the cpcm solvation model and 4 explicit water molecules

```
%nprocshared=12
%mem=24GB
%chk=ts4c int1 4w opt.chk
# opt mp2/genecp scrf=cpcm
Title Card Required
0<sub>1</sub>-1.35447800\mathsf C-0.549708002.53311000
                                0.11181500
\mathbf S-1.527151002.09573700
\mathsf C-3.99131900-2.508564000.07000900
\mathbf S-2.17924900-2.55258800-0.24812000Hg -1.68580600
                  -0.30219900-0.04414600\rm H0.41708900
                   2.01131500
                                 -1.31495500\rm H-1.142846002.23245500
                                -2,22869300H-4.20110100-2.136066001.08182900
                               -0.67136100H-4.50780000-1.88434600-1.34399400-0.410401003.62325600
H-3.54514500-0.01512700-4.34987600H0.94779900
                  -2.631441001.50574100
\mathcal{C}-2.41090600-0.268942001.36476600
S\mathcal{C}-0.324184002.39459800
                               1.43608500
```


7.6. Optimization of the PC state with the cpcm solvation model and 4 explicit water molecules %nprocshared=8

```
%mem=16GB
%chk=pc4b_opt.chk
# opt mp2/genecp scrf=cpcm 
Title Card Required
0 1
C 2.10184000 1.21304400 1.58397700
S 3.10356800 0.74845600 0.14724200
C -0.68569800 -0.79815900 3.20597100
S 0.27871900 -2.13523000 2.38931700
Hg  0.02417000 -1.46280900  0.16772500
H 1.03830000 1.14513900 1.31699300
H 2.32477300 0.49869100 2.38889000
H -0.21390700 0.18071800 3.04054800
H -1.71795300 -0.77588600 2.83000500
H 2.34839100 2.23680000 1.90297400
H -0.69236100 -1.01904900 4.28340800
C -0.92209400 3.35222900 1.56762400
S -2.41006000 2.54661000 0.85423800
C 2.41338600 1.91059400 -1.06014300
S -0.47783600 -0.64527600 -1.99319400
```


7.7. Optimization of the TS state with the cpcm solvation model and 7 explicit water molecules %nprocshared=12

```
%mem=24GB
%chk=ts4c 7w.chk
# opt=(ts, modredundant, maxstep=1, noeigentest) mp2/genecp scrf=cpcm
Title Card Required
0<sub>1</sub>\mathsf C-1.663060002.05262400 -2.63993700
    -1.83842400\rm S2.05674800
                                -0.81874400\mathsf C-2.71009500-2.528476001.79178600
\mathbf S-2.52447200-2.50756700-0.03757700Hq
    -2.17387400-0.24074500-0.41054200\rm H-0.874674001.34509000
                                  -2.93190300\rm H-2.618838001.76512300
                                 -3.09724800\rm H-1.82121800-2.110417002.28258300
\rm H-3.59994700-1.962054002.09705400
H-1.394276003.07449800
                                 -2.94306200\rm H-2.82906400-3.582280002.08319700
                   -3.120771001.11833700
                                  1.17390300
\mathsf{C}-2.71095200S
     2.16288000
                                  -0.282156002.05312400
                                -0.37425400\mathcal{C}0.40489000
                                0.04132800<br>0.03111800
                   1.87982600
S2.78281400
                 -0.42704700Hq 2.54062000
```


38 39 1.0 40 1.0 39 40 41 42 1.0 43 1.0 42 43 B 2 14 D B 14 15 D B 15 25 B B 15 28 B B 15 43 B B 2 36 B B 2 30 B H C O 0 aug-cc-pvdz **** S 0 aug-cc-pvtz **** Hg 0 sdd **** Hg 0 F 1 1.00 0.54500000 1.0000000 F 1 1.00 1.5800000 1.0000000 **** HG 0 sdd