



PCCP

**Elucidating the Role of Catalytic Amino Acid Residues in the
Peptide-Mediated Silica Oligomerization Reaction
Mechanism**

Journal:	<i>Physical Chemistry Chemical Physics</i>
Manuscript ID	CP-ART-10-2021-004542.R1
Article Type:	Paper
Date Submitted by the Author:	04-Jan-2022
Complete List of Authors:	Hare, Stephanie R; University of Washington Pfaendtner, Jim; University of Washington, Chemical Engineering

SCHOLARONE™
Manuscripts

Elucidating the Role of Catalytic Amino Acid Residues in the Peptide-Mediated Silica
Oligomerization Reaction Mechanism

Stephanie R. Hare¹ and Jim Pfaendtner^{1,2}

¹*Dept. of Chemical Engineering, University of Washington, Seattle, Washington
98195*

²*Physical Science Division, Pacific Northwest National Laboratory, Richmond,
Washington 99354*

Abstract

Understanding the detailed mechanism by which the proteins of marine diatoms such as silaffins are able to control the morphology of silica oligomers has eluded synthetic chemists and materials scientists for decades. In this study, we use DFT calculations to determine how individual amino acid residues of silaffin catalyze silica dimerization. The reaction network for formation of a silica dimer was explored using several different small molecules, including water, guanidinium ion, and methylammonium ion; the latter two molecules representing analogs of arginine and lysine, both of which are known to play critical roles in enabling catalytic function of naturally occurring protein and synthetic analogs of silaffin. It was found that the lysine analog selectively lowers the energy of a direct water removal pathway for silicate dimerization. Comparing the energy landscapes and mechanisms for various catalysts for this reaction provides direct evidence for the role of lysine side chains of silaffins in the oligomerization of silica.

1. Introduction

Uncovering the underlying chemistry of a process that occurs in nature is a crucial step in developing biomimetic synthesis routes. Computational chemistry is uniquely poised to aid in answering mechanistic questions that identify these underlying chemistries. A particular system that has intrigued synthetic and computational scientists alike for the past several decades is the silica exoskeletons of marine diatoms, which can be seen in **Figure 1** from a sample of diatomaceous earth, a household product often used for pest control. A detailed understanding of the mechanism by which marine diatoms are able to sequester orthosilicic acid (**Figure 1**, left inset) from seawater to generate well-controlled exoskeleton morphologies has been a longstanding challenge in the field of biomimetic synthesis.

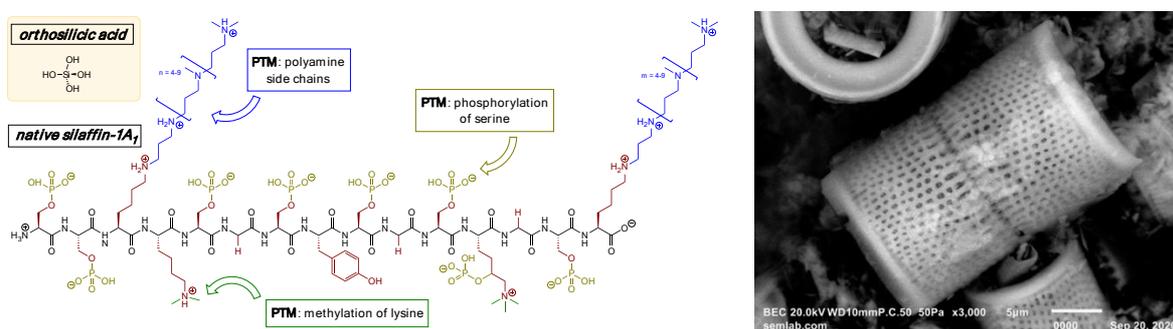


Figure 1: Left: Structure of native silaffin-1A₁ extracted from marine diatoms.¹ There have been various studies on modifications to the structure of native silaffin-1A₁ to investigate the role of particular amino acid residues on the oligomerization reaction to generate silica, namely modifications of the primary amino acid sequence and post-translational modifications (PTM) such as the addition of polyamine side chains to lysine residues (blue), phosphorylation of serine residues (gold)^{2,3}, and the methylation of lysine residues (green)³. Top left inset: Structure of orthosilicic acid. Right: A scanning electron microscopy (SEM) image of a silica nanostructure of a marine diatom from a sample of diatomaceous earth at 3,000x magnification.

Emulating nature's ability to efficiently control the synthesis of silica nanostructures would benefit a variety of applications in which silica is used, from materials chemistry to medicine.⁴ One example of such an application is the use of silica nanoparticles as drug delivery systems; orthosilicic acid undergoes a process known as Ostwald ripening in solution, and this process has been harnessed to generate molecularly imprinted nanoparticles of silica for selective and specific recognition of target molecules, analogous to the way antibodies are able to detect viruses in the body.⁵ It is therefore extremely valuable to have a detailed understanding of the mechanism in which orthosilicic acid molecules are able to oligomerize in different environments, an understanding that would allow researchers the ability to control both the morphology and thus the function of synthetic silica nanoparticles.

The polypeptides responsible for templating/catalyzing the synthesis of silica oligomers from orthosilicic acid building blocks within the cell wall of marine diatoms are called silaffins (a portmanteau of "silicon" and "affinity"). When isolated and added to a buffered solution of orthosilicic acid, silaffins spontaneously generate silica nanospheres. An example of a silaffin isolated from marine diatoms, specifically silaffin-1A₁, as it is referred to in the publication detailing the first time silaffins were isolated and identified⁶, can be seen in **Figure 1**, along with the structure of orthosilicic

acid. Native silaffin-1A1 exhibits several post-translational modifications (PTMs), including methylation of lysine residues, the addition of polyamine chains on other lysine residues, and phosphorylation of serine residues. Further, these polypeptides are composed of several repeating sequences of amino acids that, when fractionated, are also able to independently generate silica nanospheres in solution. Silaffins have inspired a number of synthetic analogs designed to achieve similar functionality. Recently, a synthetic elastin-like polypeptide (ELP) was found to have a specific activity 40 times that of native silaffin.⁷ Additionally, simple, free-standing films of alternating leucine and lysine residues at an air-water interface have been shown experimentally to generate nanoscale silica films.⁸ In order to ultimately understand the complex process occurring within marine diatoms, it is essential to understand the role of the proteins/peptides on the mechanism of orthosilicic acid oligomerization in solution.

In spite of the progress to date, there are still significant gaps in our understanding of the molecular scale mechanisms and driving forces that give rise to biosilicification reactions. Some hypotheses that have arisen to attempt to explain this behavior in various biomineralization systems include: peptide induced supersaturation of orthosilicic acid in its vicinity⁹, peptide mediated aggregation resulting in an electrostatic environment that effectively lowers the barrier for formation of silica oligomers¹⁰, or the peptide's side chains interacting directly with the orthosilicic acid molecules¹¹ to lower the barrier to oligomerization by lowering the relative energy of the rate-determining transition state structure. In this study, we have set out to investigate a component of the latter hypothesis; that is, whether particular small molecule catalysts are able to lower the barrier to dimerization of orthosilicic acid at an atomic-level resolution of the relevant reaction network.

Based on studies on the effects of peptide mutation on silica nanosphere morphology, it has been proposed that the mechanism for orthosilicic acid dimerization invokes two lysine sidechains to scaffold the reaction (**Figure 2**).¹¹ Though it has not been confirmed that the lysine serves the catalyzing role illustrated in **Figure 2**, it is a known requirement that protonated nitrogen-containing side chains (e.g., lysine or arginine) must be present in order for reaction to occur. This has led to several experimental studies looking specifically at the kinetics and resultant silica nanoparticle morphologies of systems using polylysines^{12,13} and polyarginines¹³ to catalyze the oligomerization reaction. Coradin et al.¹³ found that the number of NH_3^+ binding sites is a key factor in the rate of the oligomerization reaction. It should also be noted that it is unlikely that this reaction mechanism occurs in a single step as depicted in **Figure 2**, as this depiction shows deprotonated (negatively charged) and protonated (positively charged) orthosilicic acid molecules interacting, which are very unlikely to exist in the same reaction environment at the same pH with a long enough lifetime to come close enough to reaction. That said, an analogous reaction mechanism with proton transfers and condensation steps occurring in a step-wise fashion was investigated in this study (**Figure 3** and *vide infra*).

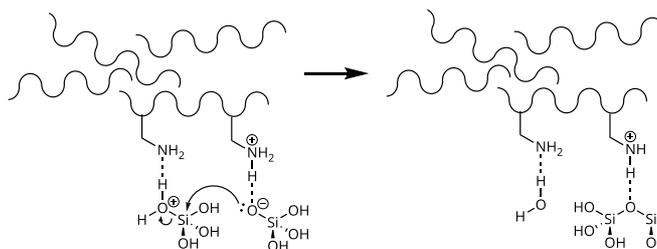


Figure 2: Proposed overall reaction mechanism of orthosilicic acid dimerization, involving two lysine sidechains of the catalyzing peptide (modified from Lechner and Becker¹¹).

Much of the groundwork for understanding the possible reaction pathways of this system originated from a study by Hu, et al.¹⁴, which used DFT to compute the energetics of the reaction network for two silica dimerization reaction types: one with no small molecule catalyst and one with a single water molecule that could act as either a catalyst or spectator at various points along the reaction pathway (**Figure 3**). To investigate the relative capability of lysine and arginine side chains to catalyze this reaction, all structures from Hu, et al. were re-optimized and the energetics of each reaction network was calculated with water replaced by a guanidinium ion, $C(NH_2)_3^+$ (as an arginine side chain analog) and methylammonium ion, $CH_3NH_3^+$ (as a lysine analog). The Hu, et al. reaction network is unique compared to other computational studies of the dimerization of orthosilicic acid because it recognizes the isomerization pathway from **intermediate 1** to **intermediate 2** as an energetically viable pathway to the product complex; while other studies acknowledge the possibility of both internal and external (via water or other protonated species in solution) proton transfers, this isomerization pathway has largely gone unexplored.^{15,16}

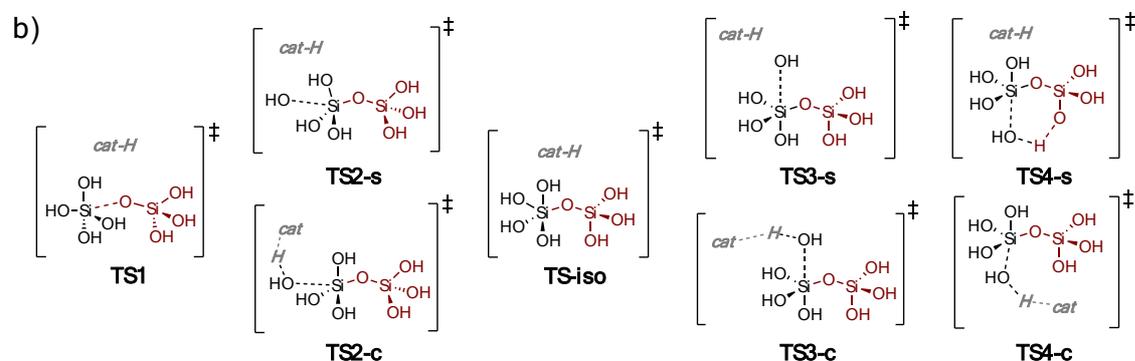
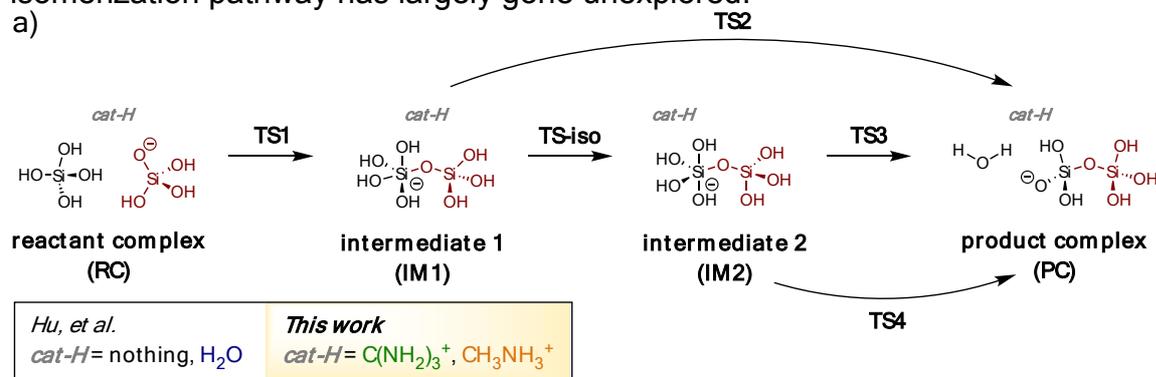


Figure 3: a) Reaction network under investigation in this study. b) Two-dimensional depictions of transition state structures. When a proton transfer occurs in the structure, a small molecule or ion [H_2O , $C(NH_2)_3^+$, or $CH_3NH_3^+$ in this study] can behave as either a spectator or proton donor, giving the structure the suffix **-s** or **-c**, respectively.

2. Results and Discussion

The energetics of the reaction network in **Figure 3** were calculated for four separate systems: 1) no catalyst, 2) H₂O as a catalyst or spectator, 3) C(NH₂)₃⁺ as a catalyst or spectator, and 4) CH₃NH₃⁺ as a catalyst or spectator. Results of the energies of these optimized structures is shown in **Figure 4**. Consistent with Hu, et al.,¹⁴ the naming convention for each stationary point for each system is as follows: the stationary point name in **Figures 3** and **4** is followed by **s** if the small molecule is behaving as a spectator (if this is relevant for the given structure) or **c** if the small molecule is behaving catalytically (e.g., donating or accepting a proton), and then the name is followed by the system abbreviation [**w** for H₂O, **g** for C(NH₂)₃⁺, and **ma** for CH₃NH₃⁺].

For each of these systems, many configurations and conformations of hydrogen-bonding networks are possible. To find the configurations that were most likely to be the lowest energy, Hu, et al. took the lowest energy configurations from the no catalyst system, added a water molecule to every possible location where the molecule could provide two additional hydrogen bonds and optimized each of those structures. In this work, every structure from the H₂O system had the H₂O molecule replaced with a C(NH₂)₃⁺ ion and a CH₃NH₃⁺ ion and each of those structures was optimized. Additionally, catalyst molecules were placed in arrangements that maximized the number of hydrogen bonds in the complex. The geometries of all of the optimized configurations for every system are included in the Supporting Information.

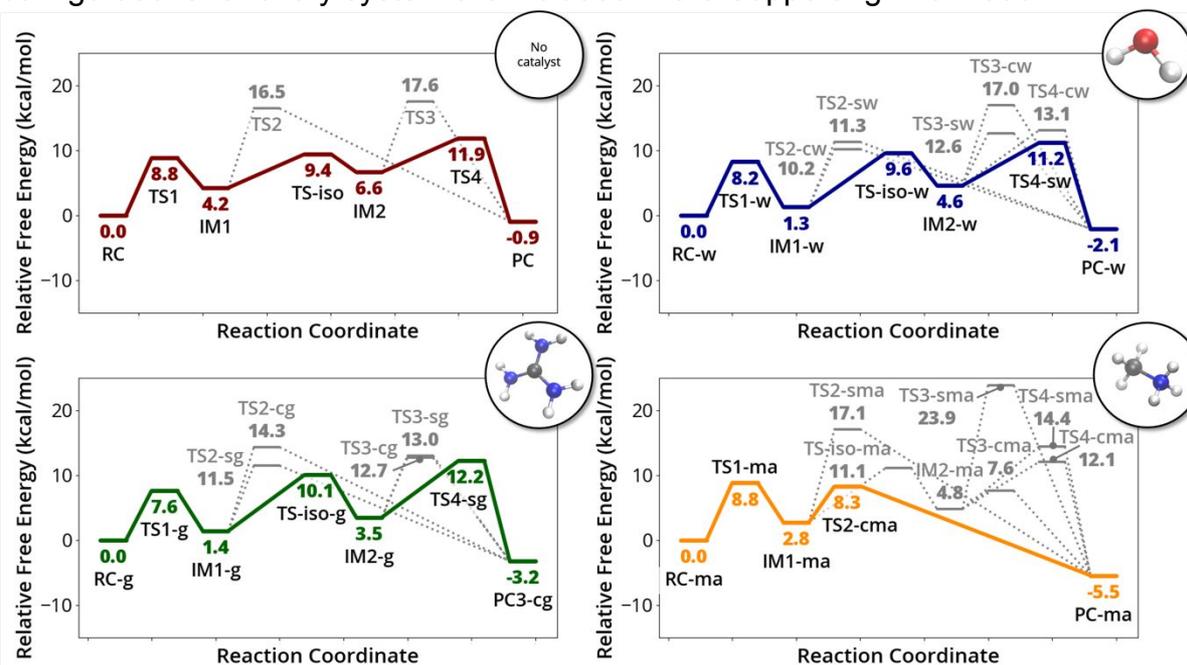


Figure 4: Energetics of the reaction networks shown in **Figure 3** for orthosilicic acid dimerization with a) no catalyst (red), b) one H₂O molecule as a catalyst or spectator (blue), c) one C(NH₂)₃⁺ ion as a catalyst or spectator (green), and d) one CH₃NH₃⁺ ion as a catalyst or spectator (orange). The minimum (**bold**) energies of configurations for each stationary point are labelled and the lowest energy pathway from reactant complex (RC) to product complex (PC) is highlighted in the respective color assigned to that system. Energies shown are free energies relative to the reactant complex at the CAM-B3LYP/6-311++G(2d,2p) level of theory.

The lowest energy pathways for each system are highlighted in color, where the lowest energy pathway is defined as the complete pathway from reactant complex to product

complex with the lowest energy rate-determining step (i.e., the step in the mechanism with the highest free energy barrier). All systems' lowest energy pathways are compared in the plot in **Figure 5**.

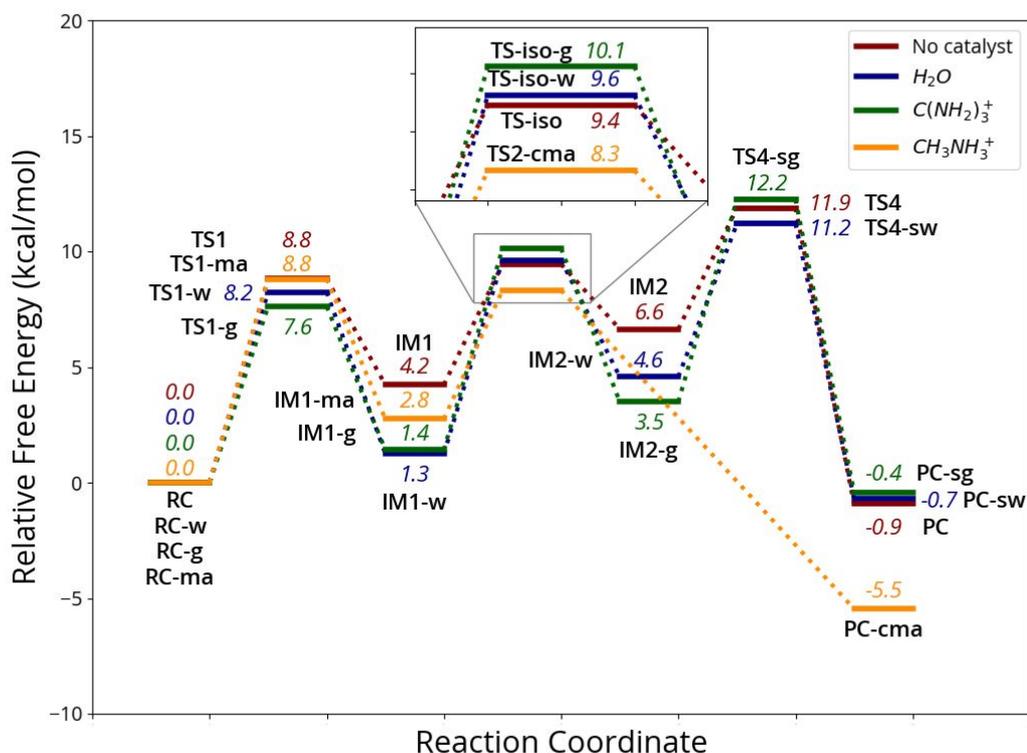


Figure 5: The lowest energy pathways from **Figure 3** for each system, overlaid for comparison. Every system's lowest energy pathway, aside from the CH₃NH₃⁺-catalyzed system, is the isomerization pathway. In contrast, CH₃NH₃⁺ selectively lowers the barrier for water removal of **IM1** enough to make direct water removal the preferred pathway. Note: Because CH₃NH₃⁺ protonates the hydroxyl group on **IM1** to remove water, the product complex **PC-cma** contains (OH)₃SiOSi(OH)₃, H₂O, and CH₃NH₂, whereas (OH)₃SiOSi(OH)₂O⁻ is deprotonated in the other three systems.

2.1 No Catalyst System

In the case of the system with no catalyst, the lowest energy pathway is definitively the isomerization pathway, where instead of **IM1** losing water directly, **IM1** isomerizes to **IM2** before losing water via **TS4**, where a proton is being transferred from a hydroxyl group located on the opposing silicon center to the leaving hydroxyl group (**Figure 6**, **TS4**). **TS2** and **TS3** are both higher in energy than their competing transition state structures, **TS-iso** and **TS4** respectively, because they both require a proton transfer to occur between the leaving hydroxyl group and a hydroxyl group on the same silicon center from which the hydroxyl is leaving. The proximity of the hydroxyl groups makes a non-ideal transfer angle of that proton in the four-atom-centered transition state; meaning, the orbital overlap between the lone pair on the oxygen of the leaving hydroxyl group and the σ^* -antibonding orbital of the breaking O-H bond is not optimal. Comparatively, **TS4** has a six-membered ring motif in its structure, allowing that proton transfer to occur with minimal strain.

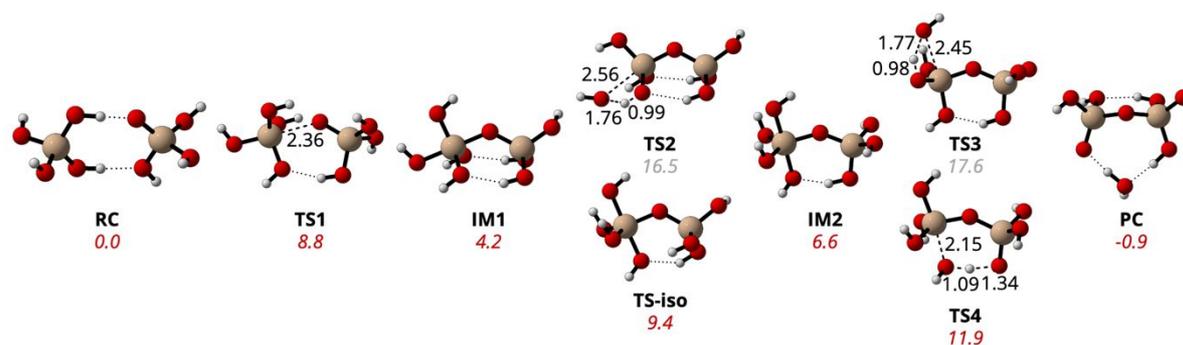


Figure 6: Three-dimensional ball-and-stick representations of the lowest free energy optimized structures for each stationary point in the reaction network for the no catalyst system. Energies shown below each structure's name are free energies relative to the reactant complex (**RC**). Where two transition state structures follow a common intermediate, the free energy of the lower energy structure is highlighted in red. Forming/breaking bond distances in the transition state structures are shown in Angstroms (Å).

2.2 H₂O System

In the case of an H₂O molecule being involved in the reaction as a catalyst or spectator, the lowest energy pathway remains the isomerization pathway; however, the presence of an H₂O molecule lowers the barrier for H₂O loss significantly by intervening in the H₂O loss transition state structures. Specifically, in the case of **TS2**, when the H₂O molecule is acting as a spectator (**TS2-sw**), the H₂O molecule can be a hydrogen bond donor to the leaving OH group, stabilizing its impending negative charge, contributing stabilization energy to counteract the strain energy of the original four-membered transition state structure. When the H₂O molecule behaves as a catalyst (**TS2-cw**), the relevant atoms are arranged in a six-membered array, with a proton being transferred to the H₂O molecule as the H₂O molecule transfers a proton to the leaving hydroxyl, fully relieving the strain associated with this transition state structure. The opposite situation is seen with **TS3**, where the water-catalyzed mechanism (that is, when water donates a proton to the leaving hydroxyl group, rather than spectating during an intramolecular proton transfer) is actually 4.4 kcal/mol higher in free energy than the transition state structure where the H₂O molecule is behaving as a spectator. This difference between catalytic transition state structures is likely because one of the proton transfers occurring in **TS2-cw** is taking place on the opposite side of a strong hydrogen bond between the oxygen losing a proton and a hydroxyl group on the opposing silicon center. This hydrogen bond is not present in either **TS3-cw** or **TS4-cw**, and since two proton transfers are occurring (i.e., two bonds are breaking), the relative energies of these TSSs is higher than **TS3-sw** or **TS4-sw**, where only one proton transfer occurs. With **TS4**, the H₂O molecule behaving as a spectator adds two additional hydrogen bonds to the structure by linking two hydroxyl groups in a hydrogen bonding network away from the leaving hydroxyl, making it the lowest energy transition state structure in the group of four competing pathways. When the H₂O behaves as a catalyst (**TS4-cw**), the cyclic part of the transition state structure expands to eight total atoms, but this does not relieve any additional strain compared to when the leaving hydroxyl group is a part of a six-membered array.

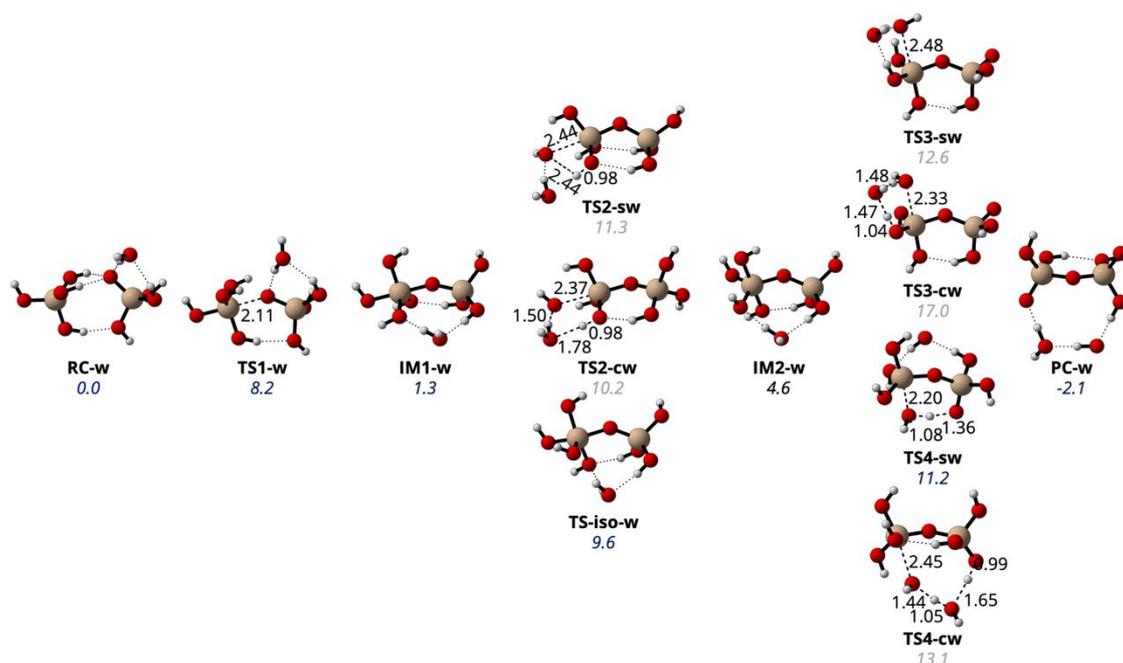


Figure 7: Three-dimensional ball-and-stick representations of the lowest free energy optimized structures for each stationary point in the reaction network for the one H_2O molecule system. Energies shown below each structure's name are free energies relative to the reactant complex (**RC-w**). Where two transition state structures follow a common intermediate, the free energy of the lower energy structure is highlighted in dark blue. Forming/breaking bond distances in the transition state structures are shown in Angstroms (Å).

2.3 $\text{C}(\text{NH}_2)_3^+$ System

When using a single $\text{C}(\text{NH}_2)_3^+$ ion, an arginine side chain analog, to catalyze or scaffold the reaction as a spectator, the picture is quite similar to that of the one H_2O molecule system (**Figure 8**). The lowest energy pathway is the isomerization pathway, and the **TS2** structures show a similar energy difference compared to the H_2O system, except that when the $\text{C}(\text{NH}_2)_3^+$ ion acts as a catalyst (**TS2-cg**, **Figure 8**), the relative free energy is about 4 kcal/mol higher than the equivalent structure in the one H_2O molecule system (**TS2-cw**, **Figure 7**). The similarities between H_2O and $\text{C}(\text{NH}_2)_3^+$ ion are not particularly surprising, as water and guanidinium ion have very similar pK_a s (the pK_a of guanidinium ion is about 13.6¹⁷ and the pK_a of water is 14¹⁸), despite the fact that guanidinium ion bears a positive charge, which is delocalized over all of the hydrogens in the structure. The reason **TS2-cg** is higher in relative free energy in its reaction network than **TS2-cw** is because guanidinium ion has quite a different shape from a water molecule and can only behave as a hydrogen bond donor, whereas water can be both a donor and acceptor. This causes slight geometric changes of the stationary points in the reaction network and in the transition state structures where $\text{C}(\text{NH}_2)_3^+$ ion acts as a catalyst, the $\text{C}(\text{NH}_2)_3^+$ ion is just donating a proton to the leaving hydroxyl group because there is no available lone pair in the $\text{C}(\text{NH}_2)_3^+$ ion structure to accept a proton from the silica dimer (see **TS2-cg** and **TS3-cg** in **Figure 8**). A transition state structure corresponding to **TS4-cg** could not be optimized, most likely because the guanidinium ion is not acidic enough to donate a proton in this conformation, and an intramolecular proton transfer from a hydroxyl group on the adjacent silicon center is always preferred (i.e., **TS4-sg** in **Figure 8**).

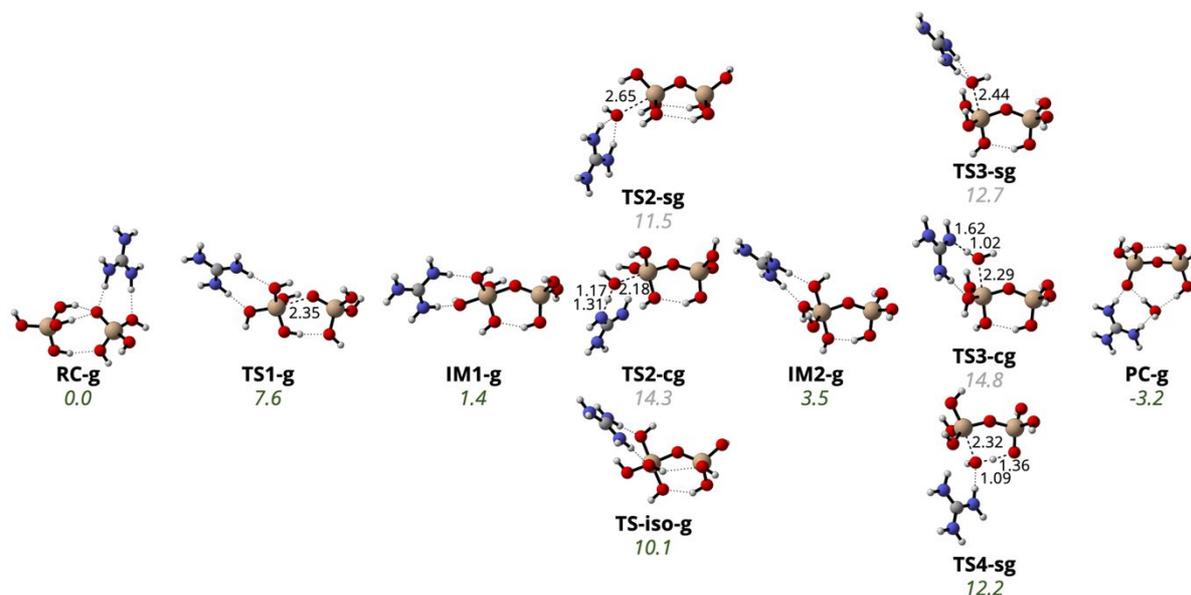


Figure 8: Three-dimensional ball-and-stick representations of the lowest free energy optimized structures for each stationary point in the reaction network for the one $\text{C}(\text{NH}_2)_3^+$ ion system. Energies shown below each structure's name are free energies relative to the reactant complex (**RC-g**). Where two transition state structures follow a common intermediate, the free energy of the lower energy structure is highlighted in dark green. Forming/breaking bond distances in the transition state structures are shown in Angstroms (Å). No **TS4-cg** structure could be optimized.

2.4 CH_3NH_3^+ System

Only the one CH_3NH_3^+ ion system shows a different mechanism for the lowest energy pathway, where the direct water loss mechanism via **TS2-cma**, with the CH_3NH_3^+ ion behaving as a catalyst, is lower in energy than the isomerization pathway. It should be noted that, like the $\text{C}(\text{NH}_2)_3^+$ ion case and unlike H_2O , CH_3NH_3^+ ion does not have a free lone pair available to synchronously accept a proton from the silica dimer as it donates a proton to a leaving hydroxyl group (see **TS2-cma**, **TS3-cma**, **TS4-cma** in **Figure 9**), and this also leads to the lowest energy product complex being one where all substituent molecules are neutrally charged (**PC-ma** in **Figure 9**). The acidity of CH_3NH_3^+ ion allows for direct water loss via **TS2-cma** to have a lower barrier than the isomerization transition state structure, **TS-iso-ma**. The pK_a of CH_3NH_3^+ ion is about 10.7, which is similar to the measured pK_a of the lysine ammonium group at about 10.5.¹⁹ Interestingly, the lysine ammonium group's pK_a can be lowered significantly dependent on the environment around the side chain.¹⁹ A future research direction could be to investigate the functional pK_a of the lysine groups of R5 or native silaffin, which could lend credence to the fact that these lysine side chains are behaving as a Brønsted acid in this reaction. It is worth nothing that pK_a of orthosilicic acid is about 9.8,²⁰ comparable to a lysine side chain's pK_a . It is possible that this means the reaction environment needs to be basic enough for there to be $\text{Si}(\text{OH})_3\text{O}^-$ ions present, but the molecule catalyzing the water removal step needs to be *acidic* enough to facilitate protonation of a hydroxyl group of **intermediate 1** or **2**.

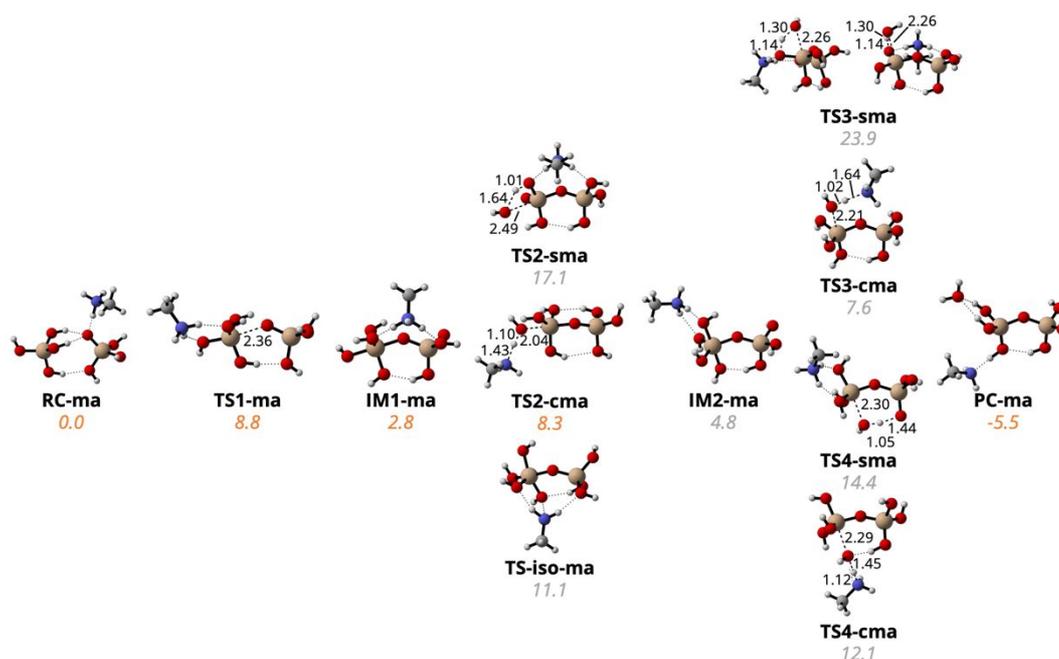


Figure 9: Three-dimensional ball-and-stick representations of the lowest free energy optimized structures for each stationary point in the reaction network for the one CH_3NH_3^+ ion system. Energies shown below each structure's name are free energies relative to the reactant complex (**RC-ma**). Where two transition state structures follow a common intermediate, the free energy of the lower energy structure is highlighted in dark orange. Forming/breaking bond distances in the transition state structures are shown in Angstroms (\AA). Two views of **TS3-sma** are shown for clarity.

2.5 Distortion-Interaction Analysis

We used a distortion-interaction analysis²¹ (also sometimes called an activation strain model²²) to understand the origins of the relative differences in energy between transition state structures for the second step of the reaction (direct water removal, **TS2**, or isomerization, **TS-iso**) for each of the systems involving a small molecule catalyst (**Figure 10**). The assumption of a distortion-interaction analysis is that the barrier of a reaction step is due to a combination of two main factors: the energy penalty required to bring the components of the transition state structure from their minimum energy structures to their geometries in the transition state structure (distortion energy) and the energy benefit due to the components of the transition state structure interacting with one another (interaction energy).

To isolate specifically how the different small molecule catalysts affect the relative energetics of the transition state structures, each transition state structure was split into two fragments: 1) the catalyst [H_2O , $\text{C}(\text{NH}_2)_3^+$, or CH_3NH_3^+] and 2) the silica moiety. A single point energy calculation of each of these fragments was conducted to get their electronic energies in the geometry at the transition state structure. Each of these fragments was then optimized to a minimum. The difference in energy between the geometry of the fragment in the transition state structure and the energy of the optimized structure is referred to as the “distortion” energy. The distortion energy of each of the fragments is then summed to get the total distortion energy required to create the geometry of the transition state structure (fragment 1 distortion, red, and fragment 2 distortion, light red, in **Figure 10**). The “interaction energy” (light blue in **Figure 10**) is then the difference between the barrier for this step of the reaction (white with black border in **Figure 10**) and the total distortion energy.

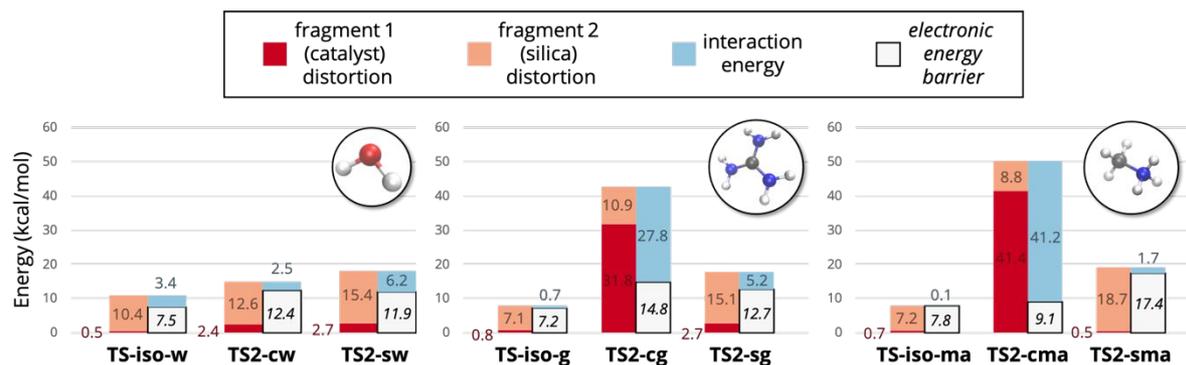


Figure 10: Results of the distortion-interaction analysis used to determine the difference in factors affecting the size of the barrier for the second step of the silica dimerization reaction (direct water removal, **TS2**, or isomerization, **TS-iso**) for the three systems involving a small molecule catalyst/spectator.

An important aspect of this analysis to note is that the barriers reported are *electronic* energies, whereas previous energies reported were *free* energies. Calculating relative free energies is done by conducting a vibrational frequency analysis, which is only possible at stationary points on a potential energy surface (which the distorted fragment structures are not). This is necessary to note because, in terms of electronic energy, the CH₃NH₃⁺ system shows a different relative ordering of transition state structures compared to free energies, with the isomerization transition state (**TS-iso-ma**) being about 1.3 kcal/mol lower in electronic energy than the direct water removal transition state with the CH₃NH₃⁺ donating a proton to the leaving hydroxyl group (**TS2-cma**). Still, this energy difference is small in the CH₃NH₃⁺ case, but more significant in the H₂O and C(NH₂)₃⁺ cases (4.9 and 7.6 kcal/mol, respectively), so it can be confidently said that the isomerization pathway is preferred in the latter cases, but not necessarily in the CH₃NH₃⁺ case.

Comparing isomerization transition state structures between systems, in each case, the distortion energy of the small molecule catalyst is minimal, and the total distortion energy is dominated by distortion of the silica moiety. This distortion is highest in the H₂O case, though this system also exhibits the highest interaction energy, and so the electronic energy barriers between all three systems are very similar. Looking at the direct water removal transition state structures with the small molecule donating a proton to the leaving hydroxyl group, the distortion energies of the silica moiety are similar, but the distortion energies of the small molecules in the C(NH₂)₃⁺ and CH₃NH₃⁺ systems are significantly higher than the H₂O case. This is because **TS2-cg** and **TS2-cma** are “later” transition state structures than **TS2-cw**; as in, the proton transfer from the small molecule to the silica moiety is more advanced in the optimized **TS2-cg** and **TS2-cma** structures compared to **TS2-cw**, causing the single point electronic energy of the distorted catalyst structure to be quite high. The advanced nature of the proton transfers in **TS2-cg** and **TS2-cma** also lead to high interaction energies between the catalyst and silica moieties. The interaction energy of **TS2-cma** ends up higher relative to the distortion energy than in **TS2-cg**, once again likely because of the relative acidity of methylammonium ion compared to guanidinium ion. Finally, looking at the direct water removal transition state structures with the small molecule behaving as a spectator in the reaction, **TS2-sw** and **TS2-sg** have very similar barriers that can be attributed to similar distortion and interaction energies, but **TS2-sma** has a higher barrier due to a higher distortion energy of the silica moiety and a much smaller

interaction energy between the catalyst and silica fragments. Ultimately, it can be concluded that the interaction energy between the CH_3NH_3^+ and silica fragments in **TS2-cma** is the biggest factor determining the difference in preference in pathways between small molecule catalyst systems.

2.6 Two CH_3NH_3^+ System

As a final step in our analysis, in order to verify the validity of Lechner and Becker's proposed mechanism¹¹ in **Figure 2**, a system involving two CH_3NH_3^+ ions was investigated (**Figures 11** and **12**). In this system, the only stationary points investigated were those where one CH_3NH_3^+ ion behaved as a spectator and the other as a catalyst/proton donor (where relevant). As can be seen in **Figure 11**, the relative energies of complexes in the reaction network with two CH_3NH_3^+ ions are similar to the relative energies of complexes involving one CH_3NH_3^+ ion, with **TS1** increasing in relative energy by about 1.2 kcal/mol and **TS2** decreasing in relative energy by about 1.7 kcal/mol. The relative free energy decrease of **TS2** implies that the addition of a second protonated amine to coordinate to the silica dimer system could increase the rate of dimerization of orthosilicic acid. The KXXK domain (i.e., two lysine residues separated by two other residues) is commonly observed in the primary amino acid sequences of silaffins,¹¹ so it is possible that two lysine residues behave in tandem in this manner to catalyze this reaction. More studies need to be conducted in order to determine whether or not this is the reason that two lysine side chains are able to catalyze the formation of nanospheres of silica *in situ*. Lowest energy optimized structures can be seen in **Figure 12**.

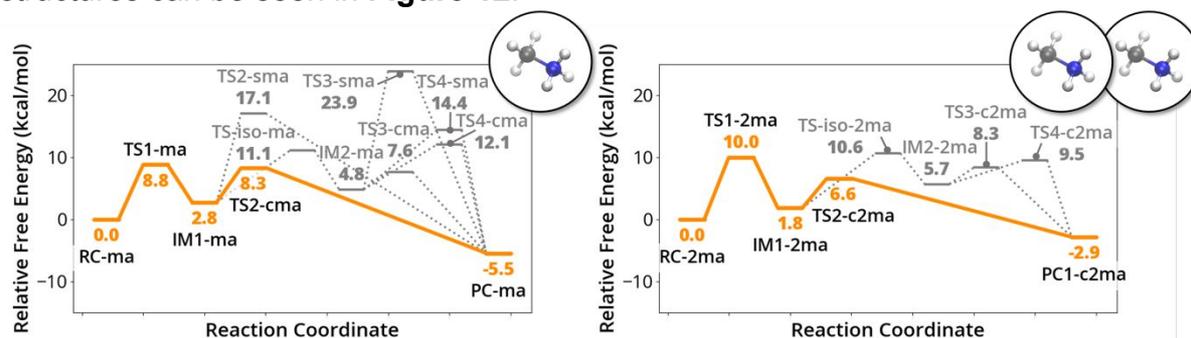


Figure 11: A comparison of the reaction networks employing one CH_3NH_3^+ ion (left) and two CH_3NH_3^+ ions (right) as catalysts or spectators. In the case of two CH_3NH_3^+ ions, in the relevant transition state structures, one CH_3NH_3^+ ion always behaved as a catalyst and the other a spectator.

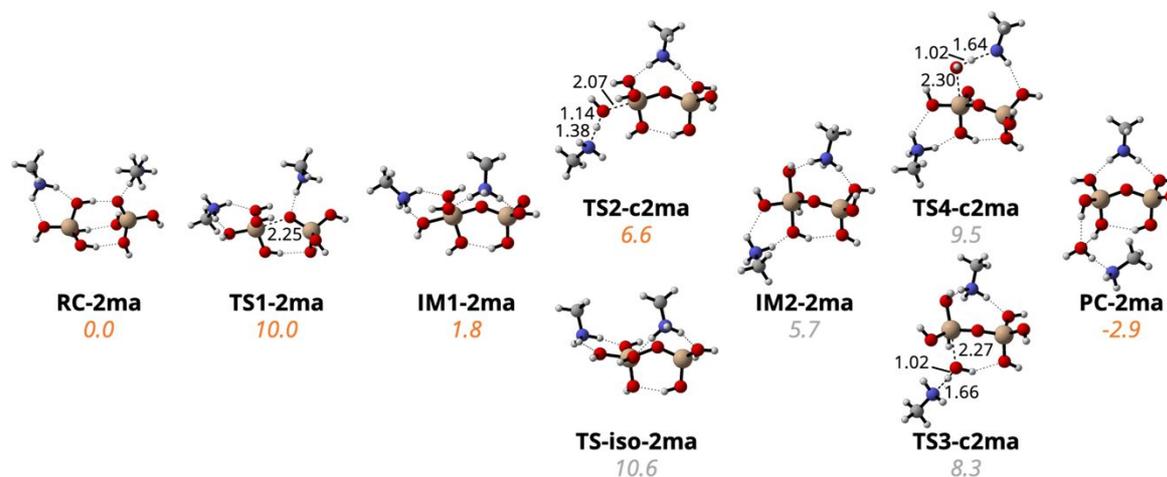


Figure 12: Three-dimensional ball-and-stick representations of the lowest free energy optimized structures for each stationary point in the reaction network for the two CH_3NH_3^+ ion system. Energies shown below each structure's name are free energies relative to the reactant complex (**RC-2ma**). Where two transition state structures follow a common intermediate, the free energy of the lower energy structure is highlighted in dark orange. Forming/breaking bond distances in the transition state structures are shown in Angstroms (\AA).

While including a second methylammonium ion appears to lower the barrier for the water removal step of the reaction, the step that would be rate-determining in this case is the generation of **IM1**, the initial complex of two orthosilicic acid molecules containing a pentacoordinate silicon center. That barrier appears to be ~ 1.2 kcal/mol higher in energy, which is likely due to a lowering in relative energy of the reactant complex, **RC-2ma**. A more thorough investigation of different hydrogen-bonding arrangements of these species would need to be conducted to draw definitive conclusions about the relative energies of stationary points in this reaction network, but the purpose of this study was to determine whether the initial reaction mechanism proposed by Lechner and Becker¹¹ would be reasonable. Indeed, the energetics of the reaction mechanism proposed (split into two steps, rather than a single step) are reasonable for this reaction to occur. Additionally, we are able to conclude that a direct water removal mechanism, as drawn in the original proposed mechanism, is more likely than the isomerization mechanism found to be the lowest energy pathway when using water molecule as a catalyst in the DFT studies of silica dimerization by Hu, et al.¹⁴

3. Conclusions

The results of this study indicate the possible ways in which lysine residues on silaffin peptides may be able to selectively catalyze the oligomerization reaction of orthosilicic acid compared to water and arginine. Namely, the acidity and shape of the protonated amine portion of the lysine sidechain lowers the barrier for water removal at the rate-determining step of the reaction mechanism. A methylammonium ion catalyst was able to change the preferential pathway in the reaction network from an isomerization pathway (preferred when a water molecule or guanidinium ion is used as a catalyst/spectator) to a direct water removal pathway. This preference was interrogated using a distortion-interaction analysis, which revealed that the strong interaction between methylammonium ion and the silica dimer in the transition state structure is to blame for the difference in preference of the second step of the reaction mechanism in the reaction network (i.e., direct water removal or isomerization).

Additionally, the direct water removal pathway remained a viable mechanism at ambient conditions when two methylammonium ions scaffolded the reaction, which supports the proposal by Lechner and Becker¹¹ that two lysine side chains are involved in the reaction when conducted in the presence of a catalyzing peptide. Ultimately, this study reveals the mechanistic origins of synthetic studies that find lysine residues present within a catalyzing peptide are necessary to generate silica nanospheres in solution.

A next step in this analysis would be using molecular dynamics simulations to systematically sample different hydrogen-bonding configurations of the small molecule catalysts investigated. When doing such investigations using chemical intuition, there is always the possibility of not sampling a wide enough breadth of chemical space. Classical molecular dynamics simulations will be conducted on the silica dimer in pure water, as well as solutions of protonated amines, to determine the dominant hydrogen-bonding configurations of these systems in solution.

4. Methods

Structure optimizations and energy calculations were conducted using *Gaussian 16*²³ software. Structures were characterized by a frequency calculation revealing zero imaginary frequencies for potential energy minima and exactly one imaginary frequency for transition state structures. The connection of each transition state structure to its flanking minima was verified by intrinsic reaction coordinate (IRC) calculation.²⁴ Unless otherwise noted, structures were optimized using the CAM-B3LYP²⁵/6-311++G(2d,2p) level of theory using the IEFPCM²⁶ implicit solvation model with a water solvent. The geometrical and energetic dependence on computational method was examined by comparing CAM-B3LYP, M06-2X, and MP2 calculations by Hu, et al.¹⁴ CAM-B3LYP was in good agreement with MP2 calculations at a lower computational cost and the 6-311++G(2d,2p) basis set was found to be large enough for accuracy, but small enough to converge in a reasonable amount of time. Thus, the CAM-B3LYP/6-311++G(2d,2p) level of theory was chosen for the study described here.

3-dimensional ball-and-stick images of molecular structures were generated using CYLView20 software.²⁷

Scanning electron microscope (SEM) image in **Figure 1** was captured using a JEOL JSM-IT-100 high resolution SEM with EDS for elemental analysis and variable (low) pressure control for imaging without evaporated coating.

5. Acknowledgments

This research was supported by the U.S. Department of Energy Condensed Phase and Interfacial Molecular Science (CPIMS) Program under Award DE-SC0019483. Computational time and resources were provided by the Hyak supercomputing cluster at the University of Washington and the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under contract no. DE-AC02-05CH11231. SRH would like to acknowledge SEM Lab, Inc. (sem lab.com) for allowing use of their equipment to capture SEM images of diatomaceous earth. SRH would also like to thank Dr. Nadia

Intan for her helpful discussions and constructive comments in the preparation of this manuscript.

6. Supporting Information

Supporting information includes Cartesian coordinates and electronic energies, enthalpies, free energies, and free energies using the quasi-RRHO approximation of all optimized structures referenced in this manuscript, as well as a folder containing all of these structures in MOL2 file formats, and is available free of charge [at \[...\]](#)

7. References

- (1) Kröger, N.; Lorenz, S.; Brunner, E.; Sumper, M. Self-Assembly of Highly Phosphorylated Silaffins and Their Function in Biosilica Morphogenesis. *Science* **2002**, *298* (5593), 584–586. <https://doi.org/10.1126/science.1076221>.
- (2) Daus, F.; Pfeifer, E.; Seipp, K.; Hampp, N.; Geyer, A. The Role of Phosphopeptides in the Mineralisation of Silica. *Org. Biomol. Chem.* **2020**, *18* (4), 700–706. <https://doi.org/10.1039/c9ob02438g>.
- (3) Wallace, A. K.; Chanut, N.; Voigt, C. A. Silica Nanostructures Produced Using Diatom Peptides with Designed Post-Translational Modifications. *Adv. Funct. Mater.* **2020**, *30* (30), 1–20. <https://doi.org/10.1002/adfm.202000849>.
- (4) Jeelani, P. G.; Mulay, P.; Venkat, R.; Ramalingam, C. Multifaceted Application of Silica Nanoparticles. A Review. *Silicon* **2020**, *12* (6), 1337–1354. <https://doi.org/10.1007/s12633-019-00229-y>.
- (5) Piletska, E.; Yawer, H.; Canfarotta, F.; Moczko, E.; Smolinska-Kempisty, K.; Piletsky, S. S.; Guerreiro, A.; Whitcombe, M. J.; Piletsky, S. A. Biomimetic Silica Nanoparticles Prepared by a Combination of Solid-Phase Imprinting and Ostwald Ripening. *Sci. Rep.* **2017**, *7* (1), 1–9. <https://doi.org/10.1038/s41598-017-12007-0>.
- (6) Kröger, N.; Deutzmann, R.; Sumper, M. Polycationic Peptides from Diatom Biosilica That Direct Silica Nanosphere Formation. *Science* **1999**, *286* (5442), 1129–1132. <https://doi.org/10.1126/science.286.5442.1129>.
- (7) Lin, Y.; Jin, W.; Qiu, Y.; Zhang, G. Programmable Stimuli-Responsive Polypeptides for Biomimetic Synthesis of Silica Nanocomposites and Enzyme Self-Immobilization. *Int. J. Biol. Macromol.* **2019**, *134*, 1156–1169. <https://doi.org/10.1016/j.ijbiomac.2019.05.159>.
- (8) Lutz, H.; Jaeger, V.; Berger, R.; Bonn, M.; Pfaendtner, J.; Weidner, T. Biomimetic Growth of Ultrathin Silica Sheets Using Artificial Amphiphilic Peptides. *Adv. Mater. Interfaces* **2015**, *2* (17), 1500282. <https://doi.org/10.1002/admi.201500282>.
- (9) Coradin, T.; Livage, J. Effect of Some Amino Acids and Peptides on Silicic Acid Polymerization. *Colloids Surfaces B Biointerfaces* **2001**, *21* (4), 329–336. [https://doi.org/10.1016/S0927-7765\(01\)00143-6](https://doi.org/10.1016/S0927-7765(01)00143-6).
- (10) Jain, A.; Jochum, M.; Peter, C. Molecular Dynamics Simulations of Peptides at the Air-Water Interface: Influencing Factors on Peptide-Templated Mineralization. *Langmuir* **2014**, *30* (51), 15486–15495. <https://doi.org/10.1021/la503549q>.
- (11) Lechner, C. C.; Becker, C. F. W. A Sequence-Function Analysis of the Silica Precipitating Silaffin R5 Peptide. *J. Pept. Sci.* **2014**, *20* (2), 152–158. <https://doi.org/10.1002/psc.2577>.
- (12) Belton, D.; Paine, G.; Patwardhan, S. V.; Perry, C. C. Towards an Understanding of (Bio)Silicification: The Role of Amino Acids and Lysine Oligomers in Silicification. *J. Mater. Chem.* **2004**, *14* (14), 2231–2241. <https://doi.org/10.1039/b401882f>.
- (13) Coradin, T.; Durupthy, O.; Livage, J. Interactions of Amino-Containing Peptides with Sodium Silicate and Colloidal Silica: A Biomimetic Approach of Silicification. *Langmuir* **2002**, *18* (6), 2331–2336. <https://doi.org/10.1021/la011106q>.
- (14) Hu, H.; Hou, H.; He, Z.; Wang, B. Theoretical Characterizations of the Mechanism for the Dimerization of Monosilicic Acid in Basic Solution. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15027–15032. <https://doi.org/10.1039/c3cp52117f>.
- (15) Trinh, T. T.; Jansen, A. P. J.; Van Santen, R. A.; Jan Meijer, E. The Role of Water in Silicate Oligomerization Reaction. *Phys. Chem. Chem. Phys.* **2009**, *11* (25), 5092–5099.

- <https://doi.org/10.1039/b819817a>.
- (16) Moqadam, M.; Riccardi, E.; Trinh, T. T.; Lervik, A.; Van Erp, T. S. Rare Event Simulations Reveal Subtle Key Steps in Aqueous Silicate Condensation. *Phys. Chem. Chem. Phys.* **2017**, *19* (20), 13361–13371. <https://doi.org/10.1039/c7cp01268c>.
- (17) Xu, B.; Jacobs, M. I.; Kostko, O.; Ahmed, M. Guanidinium Group Remains Protonated in a Strongly Basic Arginine Solution. *ChemPhysChem* **2017**, *18* (12), 1503–1506. <https://doi.org/10.1002/cphc.201700197>.
- (18) Silverstein, T. P.; Heller, S. T. PKa Values in the Undergraduate Curriculum: What Is the Real PKa of Water? *J. Chem. Educ.* **2017**, *94* (6), 690–695. <https://doi.org/10.1021/acs.jchemed.6b00623>.
- (19) Isom, D. G.; Castañed, C. A.; Cannon, B. R.; García-Moreno, B. E. Large Shifts in PKa Values of Lysine Residues Buried inside a Protein. *Proc. Natl. Acad. Sci.* **2011**, *108* (13), 5260–5265. <https://doi.org/10.1073/pnas.1010750108>.
- (20) Belton, D. J.; Deschaume, O.; Perry, C. C. An Overview of the Fundamentals of the Chemistry of Silica with Relevance to Biosilicification and Technological Advances. *FEBS J.* **2013**, *279* (10), 1710–1720. <https://doi.org/10.1111/j.1742-4658.2012.08531.x>.An.
- (21) Bickelhaupt, F. M.; Houk, K. N. Analyzing Reaction Rates with the Distortion/Interaction-Activation Strain Model. *Angew. Chem. Int. Ed.* **2017**, *56* (34), 10070–10086. <https://doi.org/10.1002/anie.201701486>.
- (22) Vermeeren, P.; Hamlin, T. A.; Bickelhaupt, F. M. Chemical Reactivity from an Activation Strain Perspective. *Chem. Comm.* **2021**, 5880–5896. <https://doi.org/10.1039/d1cc02042k>.
- (23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 16 Revision C.01. 2016.
- (24) Maeda, S.; Harabuchi, Y.; Ono, Y.; Taketsugu, T.; Morokuma, K. Intrinsic Reaction Coordinate: Calculation, Bifurcation, and Automated Search. *Int. J. Quantum Chem.* **2015**, *115* (5), 258–269. <https://doi.org/10.1002/qua.24757>.
- (25) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange-Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393* (1–3), 51–57. <https://doi.org/10.1016/j.cplett.2004.06.011>.
- (26) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2093–2999.
- (27) Legault, C. Y. CYLview20. Université de Sherbrooke 2020.