Chemistry
A European
Journal

Chemistry Europe
European Chemical Societies Publishing

Accepted Article

Title: The Role of Lysozyme in the Formation of Bioinspired Silicon Dioxide

Authors: Marina Macchiagodena, Marco Fragai, Angelo Gallo, Marco Pagliai, and Enrico Ravera

To be cited as: *Chem. Eur. J.* **2024**, e202401249

Link to VoR: https://doi.org/10.1002/chem.202401249

WILEY VCH

10.1002/chem.202401249

WILEY VCH

RESEARCH ARTICLE

The Role of Lysozyme in the Formation of Bioinspired Silicon Dioxide

Marina Macchiagodena,[a] Marco Fragai,[a,b,c] Angelo Gallo,[d] Marco Pagliai,[a] Enrico Ravera,*[a,b,c,e]

[a] M. Macchiagodena, M. Fragai, M. Pagliai, E. Ravera* Department of Chemistry "Ugo Schiff", University of Florence, via della Lastruccia 3, 50019 Sesto Fiorentino, Italy
E-mail: enrico.ravera@unifi.it, ravera@cerm.unifi.it

[b] M. Fragai, E. Ravera
Magnetic Resonance Center (CERM), University of Florence, via L. Sacconi 6, 50019 Sesto Fiorentino, Italy

[c] M. Fragai, E. Ravera
Consorzio Interuniversitario Risonanze Magnetiche di Metallo- proteine (CIRMMP), via L. Sacconi 6, 50019 Sesto Fiorentino, Italy.

[d] A. Gallo
Department of Chemistry, University of Turin, Via P. Giuria 7, 10125 Torino, Italy.

[e] E. Ravera
Florence Data Science, University of Florence, Viale G.B. Morgagni 59, 50134 Florence, Italy.

Abstract: Several organisms are able to polycondensate tetraoxosilicic(IV) acid to form silicon(IV) dioxide using polycationic molecules. According to an earlier mechanistic proposal, these molecules undergo a phase separation and recent experimental evidence appears to confirm this model. At the same time, polycationic proteins like lysozyme can also promote polycondensation of silicon(IV) dioxide, and they do so under conditions that are not compatible with liquid-liquid phase separation. In this manuscript we investigate this conundrum by molecular simulations.

Introduction

It is well known that polycationic molecules can promote the polymerization of tetraoxosilicic(IV) (orthosilicic) acid to form highly condensed silicon(IV) dioxide (silica) without thermal processing [1]. This chemistry is used by organisms like diatoms, and represents as well a reasonably convenient and sustainable [2] route to the preparation of spherical silica, one of the most relevant contemporary industrial inorganic compounds, but also for other inorganic oxides [3]. The templating effect of the polycationic molecules is reasonably due to electrostatic interactions [4], yet the detailed molecular mechanism is not completely understood, on the one hand because of the elusive nature of orthosilicic acid itself, on the other hand because chemical intuition suggests that different molecules exert their templating effect differently. A very intriguing model was suggested by Lenoci and Camp [5,6] indicating that polyamines and polycationic peptides undergo liquid-liquid phase separation (LLPS) and then the polymerization occurs at the interface between the two liquid phases [7–9]. Two very recent papers by the group of Assaf Gal provided brilliant experimental verification of this model [10,11]. It is also well known that lysozyme and other polycationic proteins can perform the templating [4,12,13]. Also, in this case, the mechanism is unknown, and literature hosts competing views in which the protein is strongly distorted in the initial phases of the polymerization [14] and at the end the two components are completely independent from each other [15], or that no distortion occurs during the polymerization [16] and the two components remain in intimate contact [17]. In any case, no evidence of LLPS under the conditions in which polycondensation is performed has been found, mostly because lysozyme tends not to undergo LLPS at pH values around neutrality to slightly alkaline [18]. Thus, polycationic molecules on the one side and polycationic proteins on the other side appear to behave differently concerning the polycondensation. In this manuscript, we question this apparent difference through molecular simulations at different levels of theory.

Results and Discussion

Role of Charged Amino Acid Sidechains

At first, we directly interrogated the possible role of charged amino acid sidechains. We have performed Nudged Elastic Band calculations [19] to identify the transition state in the polycondensation reaction [20]. In line with previous literature, in implicit solvent, the DFT-calculated mechanism involves an intramolecular proton transfer, coupled to a nucleophilic substitution, an Anion Attack Mechanism, AAM, shown in Figure 1.

The activation energy calculated at the DFT level is 101 kJ/mol, in line with previous studies [20]. No structure showing the interaction between lysozyme and orthosilicic acid is available, therefore we used as a starting point the structure of lysozyme in interaction with a tetrahedral titanium(IV) precursor species (PDB: 7A70) [16]. In this structure, the guanidinium moiety from an

10.1002/chem.202401249

arginine (R14) sidechain interacts with the tetrahedral species through two intervening water molecules. When these components are added to the calculation, the activation energy of the AAM mechanism remains in a similar range (103 kJ/mol), well within the uncertainty of the methodology. It has been demonstrated earlier that protonated amines (that can represent lysine sidechains) have a comparable impact [20]. These results collectively suggest that the catalytic effect exerted by the protein is not related to lowering the activation energy. We thus questioned whether it can be related to the electrostatic pooling of the precursors as pointed out by Coradin et al. [4]. To check this hypothesis, we performed classical Molecular Dynamics (MD) simulations at different concentrations of orthosilicic acid, from the concentration used for the polymerization in the protocol described by Luckarift et al. [3], to much higher than what can be experimentally achieved. The choice of a non-reactive force field is motivated by the need to maintain the shape of the precursor molecules.

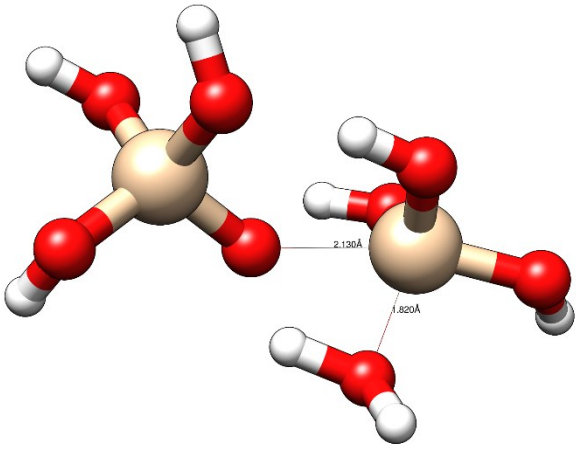


Figure 1. The transition state of the silicic acid condensation reac- tion in the AAM as obtained in implicit solvent.

Lysozyme-Orthosilicic Acid Interactions from MD Simulations

To represent the chemical system formed by lysozyme and orthosilicic acid in aqueous solution, we created three models with a number of H_4SiO_4 molecules equal to 8 (about 0.03 mol dm^{-3}), 83 (about 0.3 mol dm^{-3}), and 834 (about 3 mol dm^{-3}). Further details are given in Table 1. Each system is explicitly solvated with the TIP3P water model [21,22] at the standard density. We used the amber99sb-ildn [23] force field for the lysozyme and General Amber Force Field 2 (GAFF2) for H_4SiO_4. The GAFF2 has been obtained using the PrimaDORAC web interface [24,25] and the obtained harmonic force constants of O-Si-O angles are increased to obtain a stable H_4SiO_4 tetrahedral structure. Orthosilicic acid is expected to be fully protonated at the experimental working pH between 7 and 7.5.

Table 1. Simulated systems composition.

System	No. of H_4SiO_4 molecules	No. of H_2O molecules	Average box side (nm)
1	8	15448	7.86±0.01
2	83	13114	7.51±0.01
3	834	12541	7.85±0.01

The result that is immediately apparent is that orthosilicic acid tends to cluster around the protein, and the trend is more apparent when higher concentrations of silicic acid are considered (Figure 2). This accumulation is impossible to observe experimentally because of the reaction and can uniquely be appreciated thanks to the non-reactive simulation.

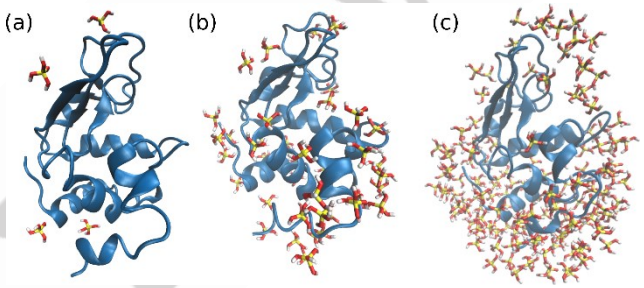


Figure 2. Snapshots of MD simulations of the three simulated systems, (a) system 1, (b) system 2, and (c) system 3. Only the H_4SiO_4 molecules present within 6 Å from lysozyme are shown.

To relate the increase in concentration to the chemistry of the protein, we have evaluated the concentration of silicic acid molecules around lysozyme residues, which we label as Local Tetraoxosilicic acid Concentration (LTC). This has been estimated computing the cumulative number of H_4SiO_4 and water molecules present within 6 Å from the Cαs of the single lysozyme residues [26,27]. The results are shown in Figure 3; increasing the number of H_4SiO_4 molecules, the LTC increases for almost all lysozyme residues, providing a quantitative indication of the trend we alluded to before of clustering of the acid around the protein. For the three simulated systems, the first 50 residues show a concentration of orthosilicic acid more uniform as compared to the other residues.

The trend observed at the lowest concentration - which we remind the reader is the concentration attained in the experiments - suggests a role for the pair of lysines 96-97 (one of the sequences in lysozyme that mimics the minimal silicification motifs of silaffins [28]), and for Arg-73. This can be quantitatively appreciated in the LTC values shown in Figure 4.
However, besides this observation, the distribution of silicic acid molecules is still not simply interpreted in terms of the primary

10.1002/chem.202401249

RESEARCH ARTICLE

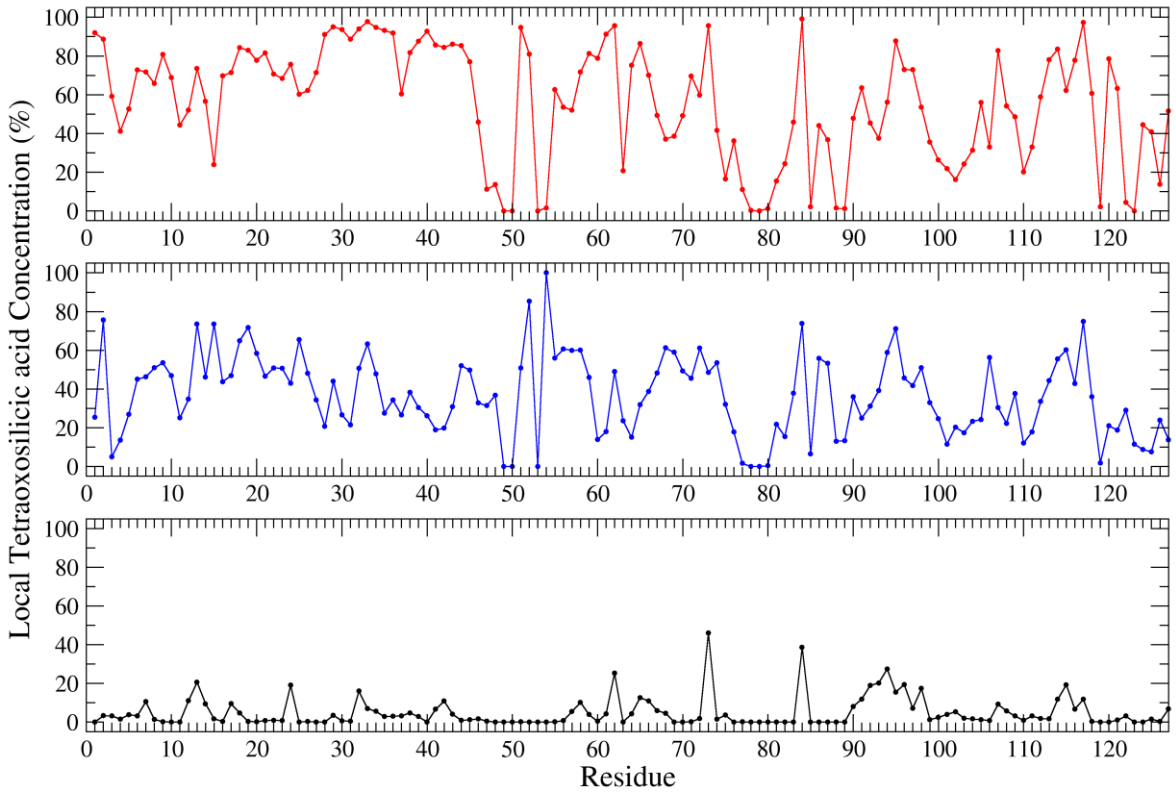


Figure 3. Local tetraoxosilicic acid concentration for the Cα atom of each residue of lysozyme. Bottom-top: System 1, black; System 2, blue; System 3, red.

sequence. Following [29], we have attributed the protein surface charges to 5 patches:
Patch 1 Arg-5, Lys-13, Arg-14, Lys-33, Arg-125 (Blue)
Patch 2 Arg-45, Arg-61, Arg-68, Arg-73 (Red)
Patch 3 Arg-112, Arg-114, Lys-116 (Orange)
Patch 4 Arg-21, Lys-96, Lys-97 (Green)
Patch 5 Lys-1 (Violet)
Patches 1 and 4 contain the minimal silicification motifs with two positive aminoacids followed by hydrophobic or aromatic residues (K13RHG and K96KIV) [28]. The center of mass of the patches is shown in Figure 5.

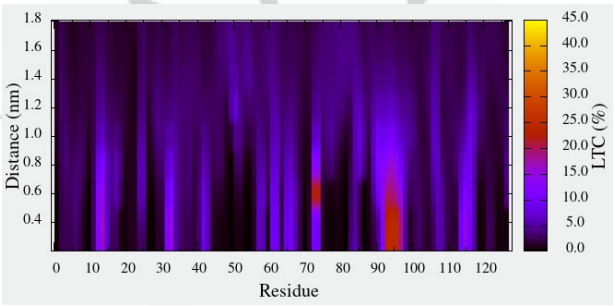


Figure 4. The LTC of silicic acid over the whole sequence of lysozyme for the system at the lowest concentration (System 1).

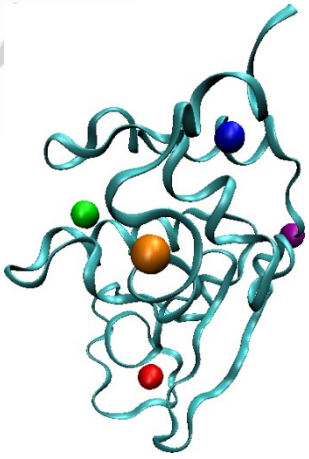


Figure 5. The position of the center of mass of the charge patches defined according to [29]. The colour code is given in the text.

On this basis, we have calculated the radial distribution function of silicic acid molecules with respect to the center of mass of the charge patches (Figure 6). From this analysis, several aspects can be deducted. In particular:

- the first peak of the radial distribution function for Patch 1 and, partly, for Patch 2 at the lowest concentration remains also at high concentration;
- Patch 2 and Patch 4 have the highest peak at the shortest distances at low concentrations, however, this peak is retained only for Patch 2 when the concentration is increased;

10.1002/chem.202401249

• Patch 5 (the N-terminus) and Patch 3 have the largest distances at the low concentration, but have the largest peaks at the lower distances when the concentration is increased.

Collectively, these observations indicate the relevance of the electrostatic pull even for the neutral silicic acid species (which is expected to be fully protonated at the working pH), and even in the presence of the electrostatic screening provided by water. This pull is particularly relevant at low concentration, and is apparently directed by Patches 1, 2, and 4. At the highest concentrations, silicic acid molecules tend to cluster around Patches 2 and 3, which are close in space, as also apparent in Figure 2. Finally, it seems that the N-terminus only contributes at the highest concentrations.

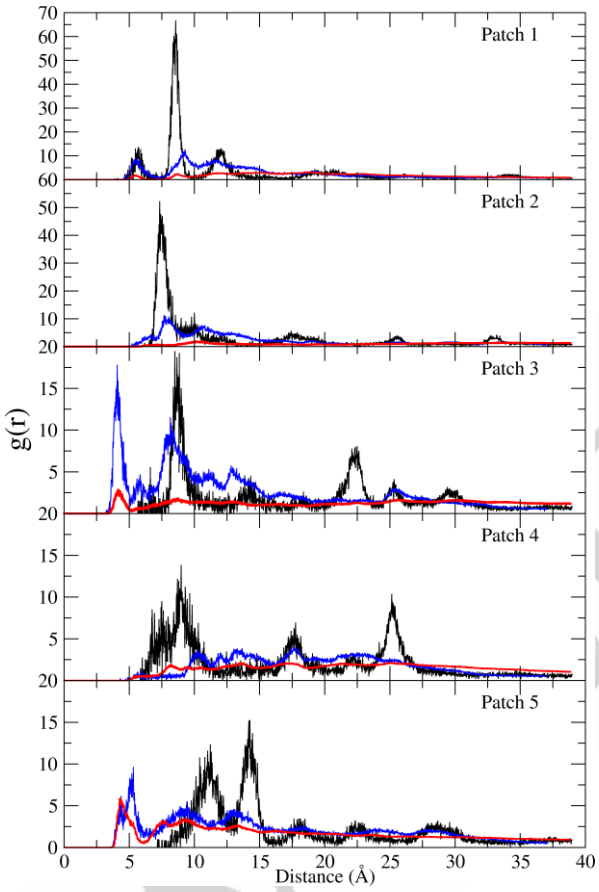


Figure 6. The radial distribution function of silicic acid with respect to 5 Patches; System 1, black; System 2, blue; System 3, red.

Conclusion

In conclusion, the molecular simulations confirm the model of the electrostatic pooling in the polycondensation of silica in the presence of lysozyme. On these grounds, we can postulate that the protein scaffold that harbors several charged sidechains in an orderly fashion exerts the same effect as a LLPS coacervate: in the protein, the charge is strongly concentrated and localized, as is in the coacervates of polycationic molecules, and thus contribute to concentrate the precursor molecules, in turn increasing the reaction rate. This picture thus reconciles nicely the experimental observations that LLPS is required for silicification peptides and not for polycationic proteins. As a last remark, this conclusion is reminiscent of the view of proteins as colloids [30]. The role of specific sidechains appears blurred in the present example. On these grounds, it will be possible to generate different arrangements of positive patches in a protein scaffold, to provide a more comprehensive representation of this fascinating phenomenon.

Experimental

Structures selection

We selected the structure 6F1O, [31] as it has been previously demonstrated that crystal structures belonging to crystal systems with higher Matthews coefficients [32] are in better agreement with solution NMR data, [33] hence they can be considered as a better proxy of the behavior of the protein in solution with respect to other crystallographic structures. The pose of tetrahedral precursor species on the protein surface was taken from the 7A70 PDB entry, [16] and optimized.

Structures optimization

The structures of orthosilicic acid and its mononegative anion have been optimized at the DFT level of theory, with the B3LYP functional [34–37], using Ahlrichs polarized basis set def2-SVP for all atoms and def2-TZVP for silicon [38], and Grimme's D4 dispersion correction [39]. CPCM implicit solvent (water) was used [40]. All calculations were carried out using ORCA 5.0.3 [41]. The structures of the oligomers of orthosilicic acid were optimized as discussed above, but def2-SVP was used for all atoms.

Molecular dynamics simulations

MD simulations are carried out in a cubic box with periodic boundary conditions using GROMACS 2021.2 software [42]. The systems have been initially minimized with the steepest descent procedure and subsequently heated to 298.15 K for 1 ns keeping constant temperature with a Nosé-Hoover approach, [43] with a time constant for coupling of 1 ps. Production run in the NPT ensemble (using the Parrinello-Rahman barostat [44] with a period of pressure fluctuations at the equilibrium setting at 2.0 ps) have been carried out for 10 ns imposing rigid constraints only on the X-H bonds (with X being any heavy atom) by means of the LINCS algorithm (δt=2.0 fs) [45]. Electrostatic interactions were treated by using particle-mesh Ewald (PME) [46] with a grid spacing of 1.2 Å and a spline interpolation of order 4. The total system charge (+7e) has been neutralized by a uniform background plasma. The cross interactions for Lennard-Jones terms were calculated using the Lorentz-Berthelot mixing rules [47,48] and we excluded intramolecular non-bonded interactions between atom pairs separated up to three bonds. The non-bonded interactions between 1-4 atoms involved in a proper torsion were scaled by the standard AMBER fudge factors (0.8333 and 0.5 for the Coulomb and Lennard-Jones, respectively). LTC has been evaluated from the cumulative number of water molecules [$nH_2O(r)$] and orthosilicic acid

molecules [$nH_4SiO_4(r)$] present within a distance r from the Cαs of the single residues using the following relation (eq. 1):

$$LTC(r) = \frac{100 \cdot V_m^{H_4SiO_4} \cdot n_{H_4SiO_4}}{V_m^{H_4SiO_4} \cdot n_{H_4SiO_4} + V_m^{H_2O} \cdot n_{H_2O}} \quad (1)$$

where $\boldsymbol{V_m^{H_4SiO_4}} = \mathbf{0.053}$ dm^3/mol and $\boldsymbol{V_m^{H_2O}} = \mathbf{0.019}$ dm^3/mol are the average excluded volumes for silicic acid and water, respectively.

Acknowledgements

The financial support provided by the MUR – Dipartimenti di Eccellenza 2023-2027 (DICUS 2.0) to the Department of Chemistry "Ugo Schiff" of the University of Florence is acknowledged.
M.M. and M.P. acknowledge the National Recovery and Resilience Plan, Mission 4 Component 2 - Investment 1.4 - NATIONAL CENTER FOR HPC, BIG DATA AND QUANTUM COMPUTING - funded by the European Union - NextGenerationEU - CUP B83C22002830001.
E.R. and A.G. acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 1409 published on 14.9.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU– Project Title Mechanism of bioinspired inorganic oxide formation (MInO) – CUP D53D23016850001- Grant Assignment Decree No. 1386 adopted on 01/09/2023 by the MUR.
We acknowledge the CINECA under the ISCRA initiative (project ISCRA-C MInO), for the availability of high-performance computing resources and support. The authors acknowledge the support and the use of resources of Instruct-ERIC, a landmark ESFRI project, and specifically the CERM/CIRMMP Italy centre.

Conflict of Interest

The authors declare no competing financial interest.

Keywords: Silicon dioxide • Lysozyme • Polycondensation • Molecular dynamics simulations • Biosilica

References

[1] N. Kröger, R. Deutzmann, M. Sumper, Science 1999, 286, 1129.
[2] R. Curley, J. D. Holmes, E. J. Flynn, Appl. Nanosci. 2021, 11, 1777–1804.
[3] H. Luckarift, M. Dickerson, K. Sandhage, J. Spain, Small 2006, 2, 640.
[4] T. Coradin, O. Durupthy, J. Livage, Langmuir 2002, 18, 2331.
[5] L. Lenoci, P. J. Camp, J. Am. Chem. Soc. 2006, 128, 10111.
[6] L. Lenoci, P. J. Camp, Langmuir 2008, 24, 217.
[7] M. Sumper, Angew. Chem. Int. Ed. 2004, 43, 2251.
[8] M. Sumper, E. Brunner, Adv. Funct. Mater. 2006, 16, 17.
[9] N. Nassif, J. Livage, Chem. Soc. Rev. 2011, 40, 849.
[10] H. Zhai, T. Bendikov, A. Gal, Angew. Chem. Int. Ed. 2022, 61, e202115930.
[11] H. Zhai, Y. Fan, W. Zhang, N. Varsano, A. Gal, ACS Biomater. Sci. Eng. 2023, 9, 601.
[12] H. R. Luckarift, J. C. Spain, R. N. Rajesh, O. S. Morley, Nat. Biotechnol. 2004, 22, 211.
[13] T. Coradin, J. Livage, Acc. Chem. Res. 2007, 40, 819.
[14] T. M. Stawski, D. B. van den Heuvel, R. Besselink, D. J. Tobler, L. G. Benning, Beilstein J. Nanotechnol. 2019, 10, 182.
[15] D. B. van den Heuvel, T. M. Stawski, D. J. Tobler, R. Wirth, C. L. Peacock, L. G. Benning, Front. Mater. 2018, 5.
[16] L. Gigli, E. Ravera, V. Calderone, C. Luchinat, Biomolecules 2021, 11.
[17] F. Bruno, L. Gigli, G. Ferraro, A. Cavallo, V. K. Michaelis, G. Goobes, E. Fratini, E. Ravera, Langmuir 2022, 38, 8030.
[18] P. Pyne, R. K. Mitra, J. Phys. Chem. Lett. 2022, 13, 931.
[19] V. Ásgeirsson, B. O. Birgisson, R. Bjornsson, U. Becker, F. Neese, C. Riplinger, H. Jónsson, J. Chem. Theory Comput. 2021, 17, 4929.
[20] X. Liu, C. Liu, C. Meng, Int. J. Mol. Sci. 2019, 20.
[21] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, J. Chem. Phys. 1983, 79, 926.
[22] M. Pagliai, M. Macchiagodena, P. Procacci, G. Cardini, J. Phys. Chem. Lett. 2019, 10, 6414.
[23] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, Proteins 2010, 78, 1950.
[24] P. Procacci, J. Chem. Inf. Model. 2017, 57, 1240.
[25] S. Marsili, G. F. Signorini, R. Chelli, M. Marchi, P. Procacci, J. Comput. Chem. 2010, 31, 1106.
[26] D. Roccatano, M. Fioroni, M. Zacharias, G. Colombo, Protein Sci. 2005, 14, 2582.
[27] M. Casoria, M. Macchiagodena, P. Rovero, C. Andreini, A. M. Papini, G. Cardini, M. Pagliai, J. Pept. Sci. 2024, 30, e3543.
[28] C. C. Lechner, C. F. Becker, J. Pept. Sci. 2014, 20, 152.
[29] A. Steudle, J. Pleiss, Biophysical journal 2011, 100, 3016.
[30] J. Loeb, J. Gen. Physiol. 1921, 3, 557.
[31] M. Plaza-Garrido, M. C. Salinas-Garcia, A. Camara- Artigas, Acta Crystallogr. D 2018, 74, 480.
[32] B. Matthews, J. Mol. Biol. 1968, 33, 491.
[33] A. Schirò, A. Carlon, G. Parigi, G. Murshudov, V. Calderone, E. Ravera, C. Luchinat, J. Struct. Biol.: X 2020, 4, 100019.
[34] A. D. Becke, J. Chem. Phys. 1993, 98, 1372.
[35] C. Lee, W. Yang, R. G. Parr, Phys. Rev. B 1988, 37, 785.
[36] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, J. Phys. Chem 1994, 98, 11623.
[37] S. H. Vosko, L. Wilk, M. Nusair, Can. J. Phys. 1980, 58, 1200.
[38] F. Weigend, R. Ahlrichs, Phys. Chem. Chem. Phys. 2005, 7, 3297.
[39] E. Caldeweyher, C. Bannwarth, S. Grimme, J. Chem. Phys. 2017, 147, 034112.
[40] V. Barone, M. Cossi, J. Phys. Chem. A 1998, 102, 1995.
[41] F. Neese, Wiley Interdiscip. Rev. Comput. Mol. Sci. 2018, 8, e1327.

10.1002/chem.202401249

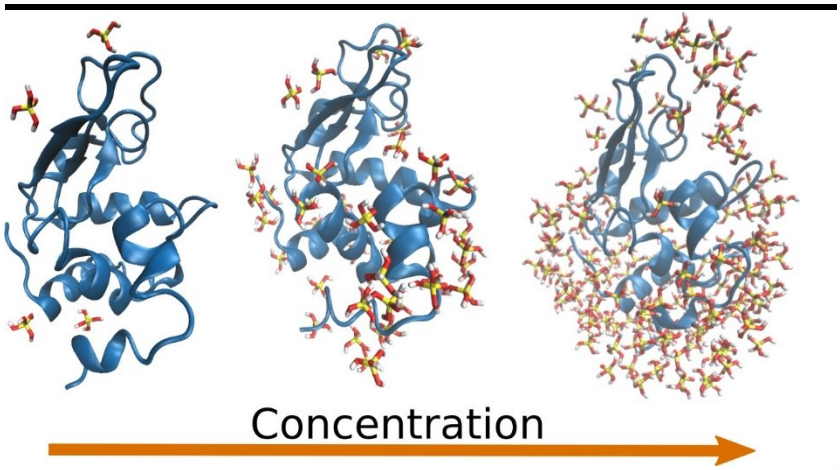
[42] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, SoftwareX 2015, 1, 19.
[43] S. Nosé, Mol. Phys. 1984, 52, 255.
[44] M. Parrinello, A. Rahman, J. Appl. Phys. 1981, 52, 7182.
[45] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, J. Comput. Chem. 1997, 18, 1463.
[46] T. Darden, D. York, L. Pedersen, J. Chem. Phys. 1993, 98, 10089.
[47] L. H. Antoon, Annals of Physics 1881, 248, 127.
[48] B. Marcellin, Comptes rendus de l'Académie des sci- ences 1898, 126, 1703.

10.1002/chem.202401249

WILEY VCH

RESEARCH ARTICLE

Entry for the Table of Contents



The mechanism of formation of silica templated by lysozyme is here interrogated using molecular dynamics simulations. The results strongly suggest that the - densely charged - protein electrostatically attracts the precursor and can thus accelerate the polycondensation reaction by increasing its concentration.

Institute and/or researcher Twitter usernames: @UNI_FIRENZE @CERM_CIRMMP @MarcoPagliai @AngeloGallo83