Biomaterials 39 (2015) 59–66

Contents lists available at ScienceDirect

Biomaterials

journal homepage: www.elsevier.com/locate/biomaterials

Molecular mechanisms for intrafibrillar collagen mineralization in skeletal tissues

Zhijun Xu [a, *], Yang Yang [b], Weilong Zhao [a], Ziqiu Wang [a], William J. Landis [a], Qiang Cui [c], Nita Sahai [a, d, *]

[a] Department of Polymer Science, University of Akron, 170 University Ave, Akron, OH 44325, USA
[b] Department of Chemistry and Biochemistry, Rowan University, 201 Mullica Hill Rd, Glassboro, NJ 08028, USA
[c] Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin-Madison, 1101 University Ave, Madison, WI 53706, USA
[d] Integrated Bioscience Program, University of Akron, 170 University Ave, Akron, OH 44325, USA

ARTICLE INFO

Article history:
Received 15 August 2014
Accepted 16 October 2014
Available online

Keywords:
Bone
Hydroxyapatite
Collagen fibril
Simulation
Biomineralization

ABSTRACT

The critical role of the self-assembled structure of collagen in skeletal mineralization is long recognized, yet the angstrom to tens of nanometers length-scale nucleation mechanism of calcium phosphate mineral (Ca–P_i) remains unclear. Here, by constructing three-dimensional structure of collagen fibril, we report direct computational evidence of intrafibrillar Ca–P_i nucleation in the collagen matrix and illustrate the crucial role of charged amino acid sidechains of collagen molecules in nucleation. The all-atom Hamiltonian replica exchange molecular dynamics simulation shows that these charged sidechains are oriented toward the fibril "hole zones" and significantly template nucleation with amorphous Ca–Pi phase, ~1.3–1.6 nm in size, thus explaining the empirical observations that Ca–P_i nucleates principally in these regions. We also show that the low water density of about 0.70 g cm^{-3} in these zones may further benefit nucleation by lowering the enthalpic penalty for ion desolvation. This work provides insight, at the atomistic level, into the nucleation mechanism of bone crystals within a collagen matrix for understanding mineral deposition, interpreting mineralization experiments and guiding the design of new implantable materials.

1. Introduction

Bone, dentin, cementum and calcified cartilage tendon are examples of vertebrate mineralized connective tissues in the body. These tissues have a hierarchical architecture and are composite materials consisting of nanometer-sized, plate-shaped crystals of hydroxyapatite (HAP; with an idealized stoichiometry of $Ca_{10}(PO_4)_6(OH)_2$), collagen, non-collagenous proteins (NCPs), water, and small molecules, such as citrate, in the case of bone [1]. Collagen, the most abundant protein in the body, provides mechanical toughness, and HAP provides stiffness. The precise collagen-mineral architecture contributes to the unique physicochemical properties of such tissues. Type I collagen molecules self-assemble into a native, "quarter-stagger" structure forming microfibrils and higher order structures, including fibrils and fibers [2] (Supplementary Fig. S1). Transmission electron microscopy (TEM) studies show that only collagen in this native self-assembled structure supports intrafibrillar mineralization [3] and the Ca–P_i nucleus forms principally within specific "hole" zones between the collagen molecules [2,3]. Although tremendous amount of efforts have been invested by various research groups to understand the mechanisms for these phenomena [4–7], the interplay between developing mineral and the supramolecular assembly of collagen in biomineralization process is still poorly understood at the atomic to fibril scale (Angstroms to tens of nanometers).

These limitations exist primarily because of experimental difficulties in spatial and temporal resolution *in vitro* and *in vivo* [8–10] of the mechanisms of Ca^{2+} and P_i transport from bulk solution to the collagen interfibrillar region and Ca–P_i nucleation events. Broadly, ion transport models involve diffusion of individual ions and ion pairs or small oligomers versus those involving larger precursor solid Ca–P_i phases [4] or liquid phases promoted by anionic non-collagenous proteins in the bulk solution [11]. The localization of Ca–P_i nuclei within a fibril is proposed to involve non-collagenous proteins as either nucleation promoters or inhibitors [4,12–14]. In one version of the ion and small oligomer model, based on primary amino acid sequence analysis of the

* Corresponding authors. Department of Polymer Science, University of Akron, 170 University Ave, Akron, OH 44325, USA.
E-mail addresses: zxu@uakron.edu (Z. Xu), sahai@uakron.edu (N. Sahai).

http://dx.doi.org/10.1016/j.biomaterials.2014.10.048
0142-9612/

collagen molecules, it has been inferred, but never proven, that charged amino acids surround and are oriented towards the hole zones. These charged sidechains are then proposed to attract Ca^{2+} and P_i electrostatically, thus promoting Ca–P_i nucleation [15,16]. This model does not preclude the possibility of extrafibrillar inhibitors of nucleation.

Molecular modeling approaches can provide a detailed analysis of collagen structure at the atomic scale and could contribute significantly to distinguishing between the numerous proposed mineralization mechanisms [17,18]. Atomistic modeling, however, has been limited by the lack of a 3D fibril structure from the angstrom to tens of nanometers length-scale and, equally critically, by the difficulty of adequate statistical sampling of configurations in the rare nucleation events during the simulation. Nucleation of ionic biominerals, such as Ca–P_i and $CaCO_3$, probably involves strong and multiple electrostatic interactions between the ions and charged sidechains of associated proteins. In these conditions, sufficient relaxation and sampling of the ion activities are necessary to obtain unbiased simulation results. Additionally, the structure of the earliest nucleated phase has not been described by computational studies in a complete system of self-assembled, hierarchical protein structure and with complete solvation, although some relatively simple models of collagen-mimic molecules have been studied [19,20].

As a major advance in the field, the 3D structure of collagen fibrils was determined only recently by synchrotron X-ray diffraction showing that the molecules are arranged in a pseudo-hexagonal array [9], as originally hypothesized based on two-dimensional (2D) TEM analyses of bone or mineralized turkey tendon [21,22]. The 3D structure provides an excellent starting point for collagen fibril structure analysis, but its relatively low-resolution (5.13 Å) provides only the coordinates of the C_α atoms in the backbone of the molecule. High-resolution, atomic-level details are required for modeling intrafibrillar Ca–P_i nucleation mechanisms to determine if they involve the interactions of charged sidechain residues with Ca^{2+} and P_i ions.

Here we address all three aspects of intrafibrillar connective tissue mineralization, which have been difficult to investigate previously, namely, obtaining a high-resolution fibril structure, mechanism of intrafibrillar Ca–P_i nucleation in hole zones, and nucleus structure. To investigate the hypothesis that the locations and orientations of charged sidechain of amino acids play an important role in Ca–P_i nucleation at hole zones in the self-assembled structure of collagen fibrils [15,16], we employ molecular dynamics (MD) simulations to construct the high-resolution collagen structure at hierarchical levels from an atom to a fibril in a fully solvated system with 0.1 M NaCl as descripted in Supporting Information (SI) 1.1 (Supplementary Figs. S1–S4). Hamiltonian replica exchange MD (HREMD) [23] is an efficient computational approach for sampling ion distributions with multiple free energy barriers (SI 1.2–1.3; Supplementary Fig. S4) and computationally less expensive than standard, temperature-REMD (TREMD). HREMD is employed here to determine Ca–P_i cluster formation and structure, as well as water density profiles in the hole zones. Nucleus structure is obtained by analyzing HREMD simulation results and by applying Steinhardt bond order parameters (SI 2). Our study represents the first application of HREMD to explore mineral nucleation guided by the collagen fibril. Very recently, Wallace et al. used TREMD to investigate the structure and energetics of hydrated $CaCO_3$ clusters, but without the presence of any protein [18].

2. Model and methods

We employed targeted MD implemented in GROMACS to build the initial structure of the entire collagen molecule on the basis of a low-resolution (5.16 Å) experimental structure [9], which provides only the C_α coordinates. After the targeted MD simulation, a high-resolution, atomic-level collagen model structure was obtained with a root-mean-square deviation (RMSD) of C_α atoms of ~0.01 Å in comparison to the low-resolution experimental structure. The entire process of constructing the full, optimized structure of the collagen molecules in the fibril is shown schematically in Fig. 1 and more details are given in Supporting Information 1.1. Subsequently, the structure was equilibrated in the presence of water and 0.1 M NaCl. Considering the computational cost and sampling difficulty, the complete collagen molecule (3360 amino acid residues) was not used in our simulations. Instead, the e1 and e2 bands [24,25] (Supplementary Fig. S1) in the hole zone of the fibril were selected for investigations of the structural characteristics of the assembled collagen molecules. Water molecules were initially distributed randomly over the simulation box. Four independent simulations with different initial structures of the e1 and the e2 bands were performed for 20 ns using conventional MD, respectively. The trajectories from all four simulations (20 ns each) in the final 5 ns were used for analyzing the fibril structure.

HREMD simulation [26,27] was used to investigate cluster formation to explore whether the packed collagen molecules in the fibril can act as a nucleating template for HAP formation. A total of 28 replicas was used and an exchange attempt was made every 1 ps, for all even and odd replica pairs in an alternating fashion (20 ns for each replica), with exchange acceptance ratios between 17% and 25% (SI 1.3). A detailed description on the implement of HREMD is provided in SI 1.3.

The initial collagen structure for HREMD simulation was obtained from the simulation of collagen fibril self-assembled structure. Two different concentrations of Ca^{2+} and P_i were designed to investigate the dependence of cluster formation and structure on ion concentrations. For each concentration, we prepared two different initial ion distributions to explore the robustness of cluster formation: Ca^{2+} and P_i were initially placed randomly either in the entire system or only in the hole zone (SI 1.2.2).

For both collagen structure optimization and nucleation simulations, the classic CHARMM22 force field [28] and TIP3P [29] models are used for the amino acids and the explicit water molecules, respectively. These force fields have been applied successfully and widely in studying biological systems in the condensed phase. It is well-established that some features of the force field for ions, such as their interaction with the protein/water atoms [30,31], are less certain than the robust force fields for the amino acids and water. We anticipate, however, that the qualitative trends observed in this study are not sensitive to the ion parameters [30,31]. Periodic Boundary Condition is applied in all directions. Particle Mesh Ewald (PME) [32] summation is used to treat the long-range electrostatic interactions with a real-space cutoff of 12.0 Å. For van der Waals interactions, a switching scheme is applied for interatomic distances between 10.0 and 11.0 Å. The isothermal-isobaric (NPT) ensemble simulations (310 K and 1 atm) are carried out. Parrinello-Rahman [33] and Nosé-Hoover [34,35] methods are used as the pressure and temperature coupling schemes, respectively, and 1 fs is applied as the time step. See SI for the details about the simulation setup and method.

3. Results and discussions

The 2D collagen fibril structure has been described in terms of “bands” observed by TEM [2]. “Hole” and “overlap” zones are formed by the association of adjacent collagen molecules in the fibril [2]. Twelve bands are identified in the hole and overlap zones, each containing electron dense, dark bands corresponding to mineralized regions alternating with electron light bands corresponding to collagen molecules (Supplementary Fig. S1) [15]. *In vivo*, the e1 and e2 bands are among the earliest to mineralize [11] and, based on the primary amino acid sequence of collagen, these bands have been shown to have high charge density [2,15].

To explore the hypothesis that charged amino acids located in the vicinity of the fibril hole zone have their sidechains oriented toward the hole zones, we optimize and analyze the e1 and e2 band structures around the hole zone in the complete self-assembled fibril (Fig. 2; only the configurations of the charged sidechains are shown for clarity, with red color for Glu and Asp residues and blue color for Lys and Arg). The hole zone is surrounded by six collagen molecules, and significantly, has charged residues from each molecule located around it (Fig. 2a, d). Most of these charged sidechains are located at one or two positions along the collagen molecule backbones of the e1 and e2 bands (Fig. 2b, e). A few charged sidechains are scattered elsewhere along the backbone. The charged sidechains in the e2 band are even more closely associated than those in the e1 band. Oppositely charged sidechains interact electrostatically to form multiple salt-bridge nests (Fig. 3a, c). Analysis of