Biochemical and Biophysical Research Communications 491 (2017) 236–240

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Substrate binding interferes with active site conformational dynamics in endoglucanase Cel5A from *Thermobifida fusca*

Xukai Jiang [a], Yuying Wang [a], Limei Xu [b], Guanjun Chen [a, *], Lushan Wang [a, **]

[a] State Key Laboratory of Microbial Technology, School of Life Science, Shandong University, Jinan, 250100, PR China
[b] Department of Molecular Medicine and Genetics, School of Medicine, Shandong University, Jinan, 250012, PR China

ARTICLE INFO

Article history:
Received 7 July 2017
Accepted 14 July 2017
Available online 15 July 2017

Keywords:
Molecular dynamics
Catalysis
Conformational dynamics
Cellulase

ABSTRACT

The role of protein dynamics in enzyme catalysis is one of the most active areas in current enzymological research. Here, using endoglucanase Cel5A from *Thermobifida fusca* (*Tf*Cel5A) as a model, we applied molecular dynamics simulations to explore the dynamic behavior of the enzyme upon substrate binding. The collective motions of the active site revealed that the mechanism of *Tf*Cel5A substrate binding can likely be described by the conformational-selection model; however, we observed that the conformations of active site residues changed differently along with substrate binding. Although most active site residues retained their native conformational ensemble, some (Tyr163 and Glu355) generated newly induced conformations, whereas others (Phe162 and Tyr189) exhibited shifts in the equilibration of their conformational distributions. These results showed that *Tf*Cel5A substrate binding relied on a hybrid mechanism involving induced fit and conformational selection. Interestingly, we found that *Tf*Cel5A active site could only partly rebalance its conformational dynamics upon substrate dissociation within the same simulation time, which implies that the conformational rebalance upon substrate dissociation is likely more difficult than the conformational selection upon substrate binding at least in the view of the time required. Our findings offer new insight into enzyme catalysis and potential applications for future protein engineering.

1. Introduction

The biological importance of enzymes cannot be overstated. Enzymes are critical to a broad range of functions, including metabolism, gene regulation, cell survival, and intracellular communication [1–3]. Additionally, they catalyze specific biochemical reactions, increasing reaction rates by many orders of magnitude to more biologically relevant timescales. Therefore, mechanistic exploration of enzyme catalysis has broad implications for both fundamental and applied research.

Cellulose consisting of β-1,4-linked glucose units is the most abundant and sustainable resource on Earth [4]. Biological conversion of biomass can reduce dependence on fossil fuels and offers opportunities to resolve environmental problems, such as global warming and air pollution. Cellulases can be divided into two types according to their different functional roles: endoglucanase (EG; EC 3.2.1.4) hydrolyzes bonds internally in cellulose chains; cellobiohydrolase (EC 3.2.1.91/176) acts preferentially on chain ends, successively cleaving cellobiose as the main product. Diverse cellulases synergistically hydrolyze cellulosic biomass into soluble sugars that are subsequently converted into biofuels and biochemicals in biorefinery industries. High catalytic activity is the most important feature of industrial enzymes; however, previous studies revealed that the reaction rate of cellulose hydrolysis quickly decreases seconds or minutes after reaction onset [5–7]. For example, the activity of EGs retain only 20% of their maximal rate after 5 min, with the gradual reduction in hydrolytic rate often persisting for days [8]. The slowdown of the catalytic rate in cellulases is closely related to the interplay between enzyme and cellulose substrate, with Murphy et al. [8] proposing that the slowdown is linked to transient inactivation of the enzyme caused by the cellulose substrate. However, explanations for transient inactivation and the role of the substrate from a structural standpoint remain unclear.

In this study, we selected EG Cel5A from *Thermobifida fusca* (*Tf*Cel5A) as the model and performed multiple molecular

* Corresponding author. Room 416, State Key Laboratory of Microbial Technology, Jinan, 250100, China.
** Corresponding author. Room 416, State Key Laboratory of Microbial Technology, Jinan, 250100, China.
E-mail addresses: guanjun@sdu.edu.cn (G. Chen), lswang@sdu.edu.cn (L. Wang).

http://dx.doi.org/10.1016/j.bbrc.2017.07.086
0006-291X/

dynamics (MD) simulations under a tandem framework. We analyzed the collective motions of the active site, the conformational distribution of active site residues, and overall structural flexibility to illustrate the change in protein dynamics of the enzyme upon substrate binding. Our results provided insight into the EG catalytic mechanism and provided an explanation for the attenuated catalytic activity observed during cellulose hydrolysis.

2. System and methods

2.1. System preparation

*Tf*Cel5A (PDB: 2CKR) used in this study belongs to glycoside hydrolase family 5 consisting of the enzymes sharing similar $(\beta/\alpha)_8$-barrel topology. To investigate the dynamic behavior of the enzyme upon substrate binding, we constructed three MD-simulation systems (Fig. S1), wherein the action of substrate binding was the only variable. *Tf*Cel5A remains in a dissociative state in substrate-free system, but binds cellopentose in a substrate-bound system. After simulations in the substrate-bound system, the substrate was removed from the complex, and only the enzyme was used for subsequent MD simulations. This continuous system enabled exploration of protein dynamics during substrate binding using a contrastive strategy.

2.2. MD simulations

All MD simulations were performed using Gromacs 4.5 software (http://www.gromacs.org/) under periodic boundary conditions [9]. The protein was solvated using the SPC model in a cubic box [10]. Water molecules that overlapped with protein heavy atoms were removed, and the total number of atoms in different the systems was >40,000. Metal ions (34 Na^+ and 25 Cl^-) were added to produce a neutral system exhibiting 0.1 mol/l ionic concentration by randomly replacing water molecules with ions. CHARMM27 and C35 carbohydrate force fields were used to describe the protein and the cellopentose substrate, respectively [11]. Energy minimization and system equilibration were performed for each simulation as described in our previous studies [12,13]. Production simulations lasting 50 ns were performed in isothermal–isobaric (NPT) ensemble. The simulation in each system was repeated three times to ensure the reliability of simulation results. The LINCS algorithm was used to constrain all bonds to hydrogen atoms in the protein, and the SETTLE algorithm was used to constrain water molecules [14,15]. The Particle Mesh Ewald method was used to evaluate long-range electrostatic interactions [16]. The cut-off for non-bonded pairs was 10.0 Å, and the pair list was updated every 10 fs.

2.3. Principal component analysis (PCA)

PCA was used to display the collective motions of the protein [13,17]. To characterize the effect of the substrate on dynamic protein behavior, we analyzed and compared collective motion in the active site based on the simulations performed in different systems. The positional covariance matrix, C, of atomic coordinates and their eigenvectors were used, and the elements of C were calculated according to the following equation:

$$C_i = \left\langle (q_i - \langle q_i \rangle)\left(q_i - \left\langle q_j \right\rangle\right)\right\rangle (i, j = 1, 2\cdots, 3N)$$

where q_i represents the Cartesian coordinates of the ith C_α atom, and N represents the number of C_α atoms in the *Tf*Cel5A active site. The average was calculated over the equilibrated trajectory after superimposition on the initial structure to remove overall translations and rotations using a least-square fit procedure. The matrix, C, was diagonalized using an orthogonal coordinate-transformation matrix, T, which transformed C into a diagonal matrix, Λ, of eigenvalues, λ_i:

$$\Lambda = T^T C_{ij} T$$

where the columns represent the eigenvectors indicating the motion directions relative to $\langle q_i \rangle$, and each eigenvector has an eigenvalue representing the total mean square fluctuation of the analytical system along the corresponding eigenvector.

On the basis of PCA results, the trajectory projection on the first two eigenvectors was calculated to display the conformational ensemble of the active site. According to Schlitter's method [18], the values of conformational entropy in different systems were estimated to quantitate the change in conformational dynamics of the active site.

2.4. Estimating side-chain dihedral angles of the active site residues

The dynamics of amino acid residues in proteins can be described by side-chain orientations [19]. Therefore, we calculated the side-chain dihedral angles for each active site residue to investigate their conformational changes. The active site residues in *Tf*Cel5A include His158, Trp162, Phe163, Tyr189, Glu192, Leu227, Glu263, Trp299, Ser305, Tyr330, His334, Tyr338, Glu355, Tyr361, Trp389, Asp394, Phe395, and Arg396, all of which are located $\leq$4 Å away from the substrate and form the binding cleft in *Tf*Cel5A. For residues lacking an δ-carbon (such as Ser), only one dihedral angle was defined by the atoms N-C_α-C_β-C_γ. For residues containing an δ-carbon, two dihedral angles were defined by the atoms N-C_α-C_β-C_γ and C_α-C_β-C_γ-C_δ, respectively. If the residue contains an equivalent δ-carbon (such as Leu or Phe), only the Cδ1 atom was used to define the dihedral angle.

3. Results

3.1. PCA of the TfCel5A active site

The *Tf*Cel5A active site is depicted by a cupped cleft that forms complex interactions with the cellulose substrate (Fig. 1A). To characterize variations in *Tf*Cel5A conformational dynamics upon substrate binding, we performed PCA for the *Tf*Cel5A active site based on MD simulation results under different systems. PCA allows projection of the conformational ensemble sampled by the active site using a reduced set of degrees of freedom. Although all principal components (PCs) contribute the collective motions, only the first few eigenvectors are related to the concerted motions and quickly decrease in amplitude to define a number of constrained and localized fluctuations. Therefore, the first two eigenvectors were used to describe the conformational ensemble of the active site.

Fig. 1 shows the trajectory projections of the *Tf*Cel5A active site on PC1 and PC2 in three different systems. The active site sampled a large and expanded conformational ensemble in the substrate-free system (Fig. 1B), with a conformational entropy of 622.3 cal/mol·K according to Schlitter's method; however, substrate binding transformed the collective motion of the active site into a constrained and well-clustered mode (Fig. 1C). At this point, *Tf*Cel5A ceased to present the large number of conformational microstates that had been present in the substrate-free system. Consistent with this observation, the entropy value decreased to 540.1 cal/mol·K. In the substrate-removed system (Fig. 1D), the overall profile of the active site conformational ensemble became similar to that observed in the substrate-free system along with slight increases in