J Mol Model (2011) 17:289–300
DOI 10.1007/s00894-010-0724-6

ORIGINAL PAPER

The effect of calciums on molecular motions of proteinase K

Shu-Qun Liu · Yan Tao · Zhao-Hui Meng · Yun-Xin Fu · Ke-Qin Zhang

Received: 18 January 2010 / Accepted: 19 April 2010 / Published online: 6 May 2010

Abstract The native serine protease proteinase K binds two calcium cations. It has been reported that Ca^{2+} removal decreased the enzyme's thermal stability and to some extent the substrate affinity, but has discrepant effects on catalytic activity of the enzyme. Molecular dynamics simulations were performed on the Ca^{2+}-bound and Ca^{2+}-free proteases to investigate the mechanism by which the calciums affect the structural stability, molecular motions, and catalytic activity of proteinase K. Very similar structural properties were observed between these two forms of proteinase K during simulations; and several long-lived hydrogen bonds and salt bridges common to both forms of proteinase K were found to be crucial in maintaining the local conformations around these two Ca^{2+} sites. Although Ca^{2+} removal enhanced the overall flexibility of proteinase K, the flexibility in a limited number of segments surrounding the substrate-binding pockets decreased. The largest differences in the equilibrium structures of the two simulations indicate that, upon the removal of Ca^{2+}, the large concerted motion originating from the Ca1 site can transmit to the substrate-binding regions but not to the catalytic triad residues. In conjunction with the large overlap of the essential subspaces between the two simulations, these results not only provide insight into the dynamics of the underlying molecular mechanism responsible for the unchanged enzymatic activity as well as the decreased thermal stability and substrate affinity of proteinase K upon Ca^{2+} removal, but also complement the experimentally determined structural and biochemical data.

Keywords Calcium · Catalytic activity · Essential dynamics · Large concerted motions · Molecular dynamics · Proteinase K · Structural stability

Abbreviations

MD	Molecular dynamics
ED	Essential dynamics
SPC	Single point charge
GRF	Generalized reaction field
NHB	Number of hydrogen bonds
NNC	Number of native contacts
SSE	Number of residues in the secondary structure elements
Rg	Radius of gyration
SASA	Solvent accessible surface area
RMSD	Root mean square deviation
ENE	Potential energy
MSD	Mean square displacement
SD	Standard deviations

Shu-Qun Liu and Yan Tao contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00894-010-0724-6) contains supplementary material, which is available to authorized users.

S.-Q. Liu · Y. Tao · Y.-X. Fu · K.-Q. Zhang (✉)
Laboratory for Conservation and Utilization of Bio-Resources,
Yunnan University,
Kunming 650091 Yunnan, People's Republic of China
e-mail: kqzhang1@yahoo.com.cn

Y. Tao
Library of Yunnan University,
Kunming 650091 Yunnan, People's Republic of China

Z.-H. Meng
Department of Cardiology, No. 1 Affiliated Hospital,
Kunming Medical College,
Kunming 650032 Yunnan, People's Republic of China

Y.-X. Fu (✉)
Human Genetics Center, School of Public Health,
The University of Texas Health Science Center,
Houston, TX 77030, USA
e-mail: shuqunliu@ynu.edu.cn

RMSF Root mean square fluctuation
PME Particle mesh ewald
TMSF Total mean square fluctuation

Introduction

Proteinase K (EC 3.4.21.64) from the fungus *Tritirachium album* limber is a member of the subtilisin family [1, 2]. The predominant site of cleavage by proteinase K is the peptide bond adjacent to the carboxyl group of aliphatic and aromatic amino acids with bulky side chains [3]. The very broad peptide cleavage activity of proteinase K has led to a wide application in academic, agricultural, and industrial fields [4–6]. Structures of proteinase K both alone and in complex with peptide chloromethyl ketone inhibitors have been solved to 0.98–2.2 Å in a series of X-ray crystallographic investigations [7–10]. These structures present a well-defined global fold comprising 15 β strands, six α helices, and one 3/10 helices (Fig. 1). The catalytic triad consists of D39, H69 and S224; the oxyanion hole is primarily formed by N161; and the substrate recognition site is primarily formed by two segments, G100-Y104 and S132-G136, which can form a three-stranded antiparallel β sheet with the substrate [10]. The native proteinase K contains two calcium cations (Fig. 1). The Ca1 is tightly bound with high affinity by $O_{\delta1}$ and $O_{\delta2}$ of D200 and carbonyl oxygen atoms of P175 and V177, whereas the Ca2 is bound with lower affinity by $O_{\delta1}$ and $O_{\delta2}$ of D260 and carbonyl oxygen atom of T16 [8, 9, 11].

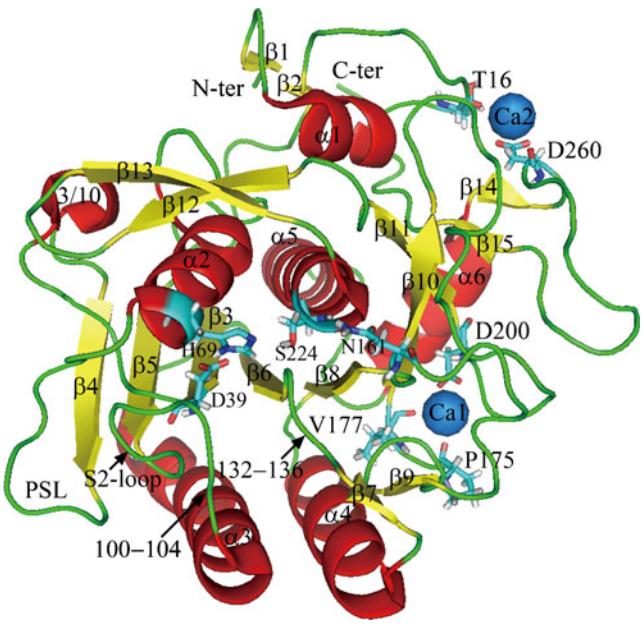

Fig. 1 Ribbon representation of proteinase K. The structure was obtained from PDB with PDB code 1IC6. The α helices, β strands and loops are colored red, yellow and green, respectively. The catalytic triad residues (D39, H69 and S224), oxyanion hole (N161), and calcium binding sites (Ca1 site: P175, V177 and D200; Ca2 site: T16 and D260) are shown as stick models. The two Ca^{2+}, Ca1 and Ca2, are shown as blue spheres

The catalytic activity of the native proteinase K remains stable over a wide range of temperatures (25–80°C) and pH values (6–10) with substantially increased activity at higher temperature (50–60°C) and pH (8–10). If proteinase K was depleted of Ca^{2+}, inconsistent functional consequences have been reported. Bajorath and colleagues reported that the catalytic activity of proteinase K toward the synthetic substrate succinyl-A-A-A-p-nitroanilide decreased by 80% within 6 hours in the absence of Ca^{2+} [11]. Interestingly, however, the activity of the Ca^{2+}-free proteinase K toward the same substrate was reported as being almost unchanged under the optimum pH and temperature, because no apparent difference in k_{cat}/K_m values was found between the Ca^{2+}-bound and Ca^{2+}-free proteinases K [12]. In addition, the transition temperature T_m upon thermal denaturation was observed to drop by ~ 10°C when Ca^{2+} was removed [12]. This has led to the proposition that the Ca^{2+} binding enhances the thermal stability of the enzyme [12, 13], but that the Ca^{2+} removal may have marked [11] or minor effect on catalytic activity [12].

Although the stabilizing effect of Ca^{2+} on enzyme structure has been known for decades, the detailed aspects of the underlying mechanism remain obscure when only biochemical data and static structures are available. Additionally, how to interpret the inconsistent effects of Ca^{2+} removal, as described above, on catalytic activity of proteinase K? Relying on the static crystallographic structures is certainly not enough. Proteases are dynamic entities so that any attempt to understand the mechanism of their action requires an analysis of their dynamic behavior while they are performing the catalytic function. Here in this study, we have performed molecular dynamics (MD) simulations of proteinases K with and without calciums to investigate the mechanism by which Ca^{2+} binding stabilizes enzyme structure and how Ca^{2+} removal affects the substrate affinity and catalytic activity. Conventional structural/geometrical analyses were used to assess the stability of the structures during simulations; and the essential dynamics (ED) technique was utilized to investigate the influence of Ca^{2+} on the structural stability and molecular motions of proteinase K. The results in this paper would greatly facilitate the interpretation of the previously published experimental results.

Materials and methods

Preparation of the starting models

The crystal structure of proteinase K at the highest resolution of 0.98 Å (PDB code 1IC6) [9] was taken from

PDB protein structure bank. To prepare the starting model for the Ca^{2+}-bound proteinase K, all hetero atoms such as NO3 and crystallographic waters were removed but the two calcium cations, Ca1 and Ca2, were retained. Here in the case of proteinase K, we considered that, after a series of position-restrained MD simulations, the latterly added explicit waters could diffuse to occupy the hydrate sites acting as crystallographic waters to maintain the integrity of the protein structure. This is verified by visual examination of the protein-solvent systems described below. The Ca^{2+}-bound model was referred to as the native proteinase K in the rest of the text. The model containing no Ca^{2+} was referred to as the Ca^{2+}-free proteinase K in the rest of the text.

Molecular dynamics simulation setup

The GROMACS package [14, 15] was used for molecular dynamics simulations using the GROMOS96 43a1 force field. The native and Ca^{2+}-free structural models were individually solvated using the single point charge (SPC) water molecules [16] in rectangular periodic boxes with a 1.4 nm solute-wall minimum distance. After a first steepest descent energy minimization with positional restraints on the solute, 5 and 1 chloride ions were introduced respectively to these two systems by replacing water molecules at the highest electrostatic potential to compensate for the net positive charges. This added up to a total of 52,070 and 52,079 atoms for the systems of the native and Ca^{2+}-free proteinases K, respectively. Subsequently, a second energy minimization was performed until no significant energy change could be detected. The systems were then simulated by five successive 200 ps position-restrained dynamics runs with decreasing harmonic positional restraint force constants on the solutes (K_{posres}=1000, 1000, 100, 10 and 0 kJ mol^{-1} nm^{-2}). Finally, the systems were visually examined to see if the integrity was preserved, revealing that some of the buried crystallographic-water positions, for example, those occupied by crystallographic waters 326 and 328, were occupied by the explicit waters, *e.g.*, with index 5205 and 9139 in the native proteinase K and index 9132 and 9163 in the Ca^{2+}-free proteinase K, respectively.

Production MD simulations were performed for 50 ns with coordinates saved every 10 ps. Solute, solvent and counter-ions were independently coupled to a reference temperature bath at 300 K with a coupling constant τ_t of 0.1 ps [17]. The pressure was maintained by weakly coupling the system to an external pressure bath at 1 atmosphere with a coupling constant τ_p of 0.5 ps. The non-bonded pair was updated every 10 steps and the non-bonded interactions were calculated using twin range cutoffs of 8 Å and 14 Å. Long range electrostatic interactions beyond the cutoff were treated with the generalized reaction field (GRF) model using a dielectric constant of 54 [18]. The LINCS algorithm [19] was used to constrain the bond lengths to their equilibrium positions. A two fs time step was used for the integration of the equation of motion.

Analysis techniques

The conventional geometrical analyses such as number of hydrogen bonds (NHB), number of native contacts (NNC), number of residues in the secondary structure elements (SSE), radius of gyration (Rg), solvent accessible surface area (SASA) and root mean square deviation (RMSD) were performed using the programs g_hbond, g_mindist, do_dssp, g_gyrate, g_sas, and g-rms within GROMACS, respectively. The hydrogen bond network formed between residues was represented by residue1⋯[residue2, residue3, residue4...], where residue1 forms hydrogen bonds with residue2, residue3 and residue4, *etc*. The salt bridge network was represented by residue1:residue2:residue3, where residue2 forms salt bridges with residue1 and residue3. The potential energy (ENE) was calculated by the program g_energy using the GROMOS96 43a1 force field.

Essential dynamics

The ED method [20, 21] is a powerful tool for filtering large concerted motions from an ensemble of structures, *i.e.*, a set of crystal structures [22] or a trajectory of snapshots derived from computer simulation [23–27]. This method is based on the diagonalization of the covariance matrix of atomic fluctuations. The resulting eigenvectors indicate directions in a $3N$-dimensional (the N is the number of atoms used for constructing the covariance matrix) configurational space and describe concerted fluctuations of the atoms. The eigenvalues are a measure of the mean square fluctuation of the system along corresponding eigenvectors. The central hypothesis of ED is that only the eigenvectors with large corresponding eigenvalues are important for describing the overall motion of the protein. Projection of the trajectory onto an eigenvector allows for a study of the time dependence of the conformational changes along this eigenvector. ED analyses were performed using the programs g_covar and g_anaeig within GROMACS with inclusion of only C_α atoms.

A useful method for comparing the ED properties of two simulations on similar systems is the so-called combined analysis [28]. In this method, ED analysis can be performed on a combined trajectory constructed through concatenating the individual simulations. Projections of trajectory onto the resulting eigenvectors and comparison of these projection properties provide a powerful tool for evaluating similarities and differences in essential motions between different simulations. There are two main effects to be investigated: i)

the differences in the distribution of these projections, which indicate that simulations have different average displacements along eigenvectors; ii) the differences in the mean square displacement (MSD) of these projections, which show differences in dynamics along different eigenvectors. Here in this study the combined analysis was applied to the concatenated trajectory of the native and Ca^{2+}-free proteinase K simulations.

Results

Geometrical property analyses

Because of the large system sizes, MD simulations required ~ 10 ns to reach equilibrium (see Online Resource Fig. S1). Table 1 shows a comparison of geometrical properties of the native and Ca^{2+}-free proteinases K calculated over 10–50 ns MD simulations. These properties are very similar, as only minor differences can be observed in their average values. However, subtle changes in some geometrical properties can still reflect the effect of Ca^{2+} on protein dynamics. For example, slight reductions in the NHB, NNC and SSE upon Ca^{2+} removal indicate reduced *inter*-atomic interactions in protein, which in turn lead to slight increases in SASA and Rg. In addition, the Ca^{2+}-free form has higher RMSD values and ENE than the native form, suggesting a relative destabilization of the enzyme structure when Ca^{2+} is removed. This is also reflected by increased standard deviations (SD) of certain geometrical properties of the Ca^{2+}-free proteinase K in comparison to those of the native proteinase K. Taken together, the protein is on average in a more open and relaxed conformation in the absence of Ca^{2+}.

Also worth noting is the average RMSD values calculated with respect to the starting models after superposition on the SSE backbones ("SS bb"). For both forms of proteinases K, the "All atom", "All bb" (backbone atoms) and "SS bb" have the largest, moderate and lowest RMSD values, respectively, indicating that the relatively large conformational fluctuations originate mainly from the loops/links located between SSEs (Table 1). Furthermore, The removal of Ca^{2+} leads to larger fluctuations in the loops/links as the increases in RMSD are larger for the Ca^{2+}-free form (0.77 Å from "SS bb" to "All bb" and 0.85 Å from "All bb" to "all atom") than for the native form (0.59 Å from "SS bb" to "All bb" and 0.77 Å from "All bb" to "All atom").

Comparison of B-factors

B-factors were computed based on the root mean square fluctuation (RMSF) calculated from 10–50 ns MD simu-

Table 1 Average geometrical property (Standard deviations are shown in parentheses) statistics (10–50 ns)

	NHB[a]	SASA[b](Å^2)	NNC[c]	Rg[d] (Å)	ENE[e] (kJmol^{-1})	RMSD[f] (Å)			SSE[j]		
						All atom[g]	All bb[h]	SS bb[i]	α helix	β sheet	Turn
Native	219 (7)	18043.3 (1215.6)	136087 (785)	16.6 (0.5)	−715966 (877)	2.58 (0.06)	1.81 (0.07)	1.22 (0.05)	72 (2)	60 (3)	32 (5)
Ca^{2+}-free	212 (8)	18380.8 (2626.8)	135789 (794)	16.7 (0.7)	−711359 (886)	2.94 (0.10)	2.09 (0.09)	1.32 (0.09)	66 (3)	59 (4)	29 (5)

[a] Number of hydrogen bonds. A hydrogen bond is considered to exist when the donor-hydrogen-acceptor angle is larger than 120° and the donor-acceptor distance is smaller than 3.5 Å

[b] Total solvent accessible surface area

[c] Number of native contacts. A native contact is considered to exist if the distance between two atoms is less than 6 Å

[d] Radius of gyration

[e] Potential energy calculated with the GROMOS96 43a1 force field using a twin range cutoffs of 8 Å and 14 Å with a reaction filed correction

[f] RMSD with respect to the starting structure. The RMSD values were calculated with respect to starting structure after superposition on the secondary structure element backbones as defined by DSSP in the starting structure

[g] All atom (including hydrogen atoms) RMSDs

[h] Backbone RMSD values of all residues

[i] Backbone RMSDs of the secondary structure elements

[j] Number of residues in the corresponding secondary structure elements

lations. Figure 2 shows a comparison of B-factors between the native and Ca^{2+}-free proteinases K. Although these two curves are similar for many structural regions, changes in the magnitude of B-factors can reflect the effect of Ca^{2+} on the stability of the protein. The B-factors averaged over the 279C_α atoms of the native and Ca^{2+}-free proteinases K are 16.7 and 28.2 $Å^2$, respectively, reflecting an overall increase in flexibility of the protein upon Ca^{2+} removal.

Close examination of Fig. 2 reveals that a majority of regions hve higher B-factors in the Ca^{2+}-free proteinase K than in the native form except for a limited number of regions showing lower B-factors in the Ca^{2+}-free form. For regions with increased B-factor, some are situated at the beginning or end of SSEs, whereas most of them are located in the surface-exposed loops. The most significant increase is seen in the N- and C-termini and the loops surrounding residues 61, 81, 100, 214 and 241 (Fig. 2), although many of them are located far away from the two Ca^{2+}-binding sites.

The regions exhibiting reduced B-factors upon Ca^{2+} removal include residues 132–136, 160–169 and 220–223. Of these, segment 132–136 forms part of both S1 and S4 substrate-binding pockets; and regions 160–169 and 220–223 are parts of the substrate-binding pocket S1 [10]. In order to ascertain whether MD simulation can reproducibly predict such B-factor differences, two additional 50 ns MD simulations on the Ca^{2+}-bound and Ca^{2+}-free proteinases K were performed using particle mesh Ewald (PME) electrostatics [29] for computing electrostatic interactions instead of GRF [18] as described in the "Materials and methods" section. This simulation protocol produced the consistent results (see Online Resource Fig. S2), suggesting that the higher flexibility of these substrate-binding regions in the native proteinase K may be functionally relevant.

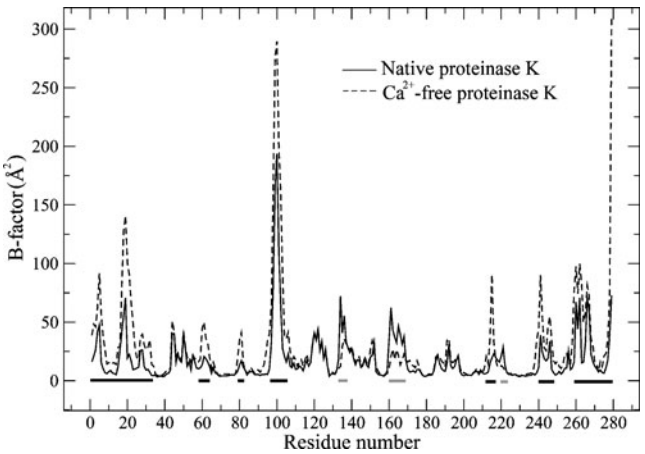

Fig. 2 Comparison of the C_α B-factors calculated over the 10–50 ns simulations of the native (Solid lines) and Ca^{2+}-free (dashed lines) proteinases K. The regions exhibiting significant increase and reduction in B-factors upon Ca^{2+} removal are indicated with black and gray bars along the horizontal axis, respectively

It is worth noting that the differences in B-factors of the catalytic triads of these two forms of proteinase K are only minute. The values for D39, H69 and S224 are 3.5, 4.0 and 6.2 $Å^2$ in the native proteinase K and 4.9, 4.2 and 5.4 $Å^2$ in the Ca^{2+}-free proteinase K, respectively, indicating that Ca^{2+} removal has minor effect on the thermal motions of the catalytic triad. For the native and Ca^{2+}-free proteinases K, the B-factors averaged over the Ca1 site residues are 6.86 and 23.6 $Å^2$, respectively. Such a large increase in flexibility of this Ca^{2+}-binding site is not surprising, as the Ca1 does not tether its binding residues anymore. The difference in B-factors of the Ca2 site residues are only minor (40.7 and 43.1 $Å^2$ for the native and Ca^{2+}-free proteinases K, respectively). It is possible that the diffusion of Ca2 from its binding site results in this minor difference (see below).

Ca^{2+} motions

The stability of thje Ca1 and Ca2 during simulation was examined by monitoring the distances between the calciums and the centers of mass of their binding residues. Figure 3a shows that the Ca1-Ca1 site distance is rather stable, whereas the Ca2-Ca2 site distance fluctuates dramatically, indicating that the Ca2 has the ability to diffuse away from its binding site. To further determine whether Ca2's diffusion is the consequence of an inaccurate description of calcium–protein interactions by the chosen simulation protocol, we have conducted two additional 50 ns MD simulations on the native proteinase K: i) the simulation protocols were similar to those described in the "Materials and methods" section except for a different "seed" value used to initialize random generator for random velocities to atoms; ii) the PME electrostatics was used for computing electrostatic interactions as mentioned above. Examination of the Ca1-Ca1 site and Ca2-Ca2 site distances gives similar results (see Online Resources Fig. S3 and S4), indicating that the Ca2 diffusion may not be the consequence of the simulation artifact.

The electrostatic interaction energies between the Ca^{2+} and the protein and between the Ca^{2+} and the solvent were further calculated to ascertain the factors responsible for the distinct behavior of Ca1 and Ca2. As shown in Fig. 3b, the Ca1-protein interaction energy is relatively stable, fluctuating around an average value of −1144.9 (SD = 88.8) kJ mol^{-1}. In contrast, the Ca2-protein interaction energy is rather unstable and increases from −880.0 to 84.3 kJ mol^{-1} within the first 1 ns, after which it fluctuates around 0 kJ mol^{-1} throughout the remainder of the simulation, with occasionally durative zero interaction energy observed when Ca2 moves far away from the protein. The average value of the Ca2-protein interaction energy is −69.3 (SD= 113.2) kJ mol^{-1}. These results indicate that the interaction

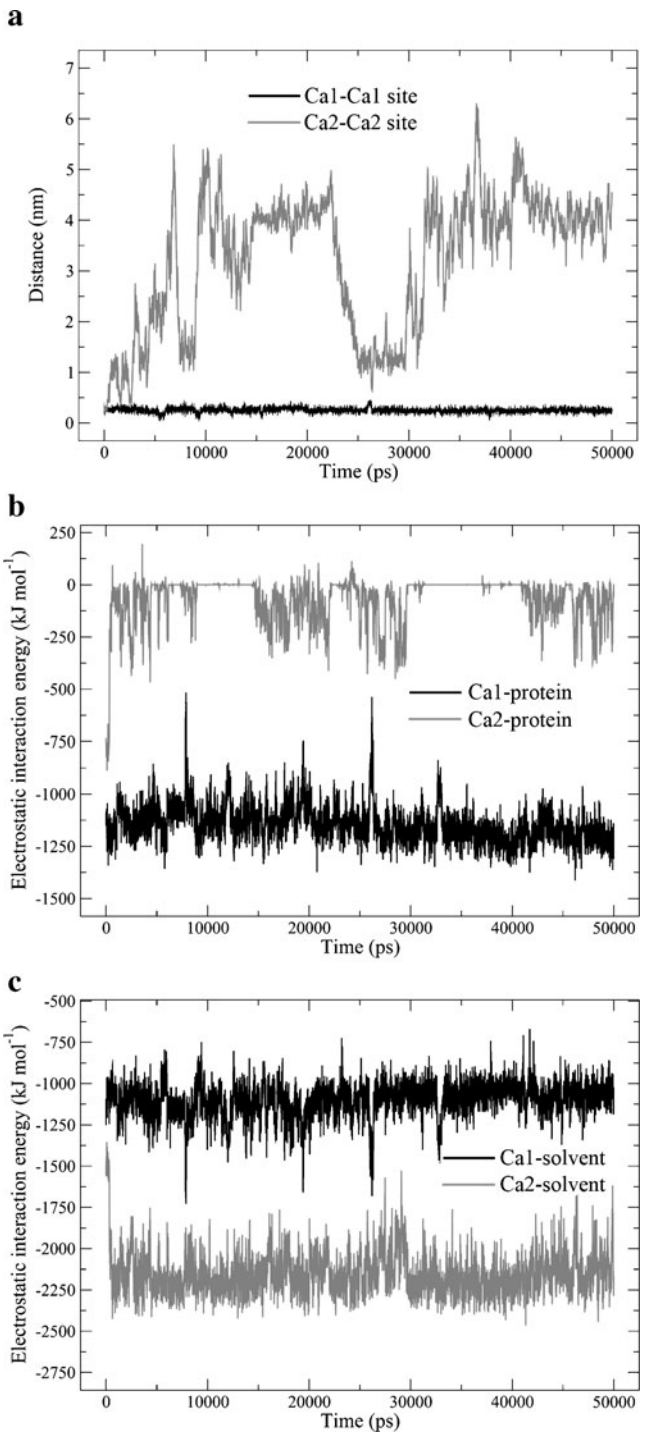

Fig. 3 Distances and interaction energies between the Ca^{2+} and the protein and between the Ca^{2+} and the solvent during simulation. (**a**) The distances between the two calciums and the centers of mass of their binding sites as a function of time. The distances of the Ca1-Ca1 site and Ca2-Ca2 site are shown in black and gray lines, respectively. (**b**) The electrostatic interaction energies between the two calciums and the protein. The interaction energies of the Ca1-protein and the Ca2-protein are shown in black and gray lines, respectively. (**c**) The electrostatic interaction energies between the two calciums and the solvent. The interaction energies of the Ca1-solvent and the Ca2-solvent are shown in black and gray lines, respectively. The energies are the sum of the short (SR) and long range (LR) terms; the 1–4 terms are not included

of the Ca1-protein is stronger and more stable than that of the Ca2-protein. The average values (SD) for interaction energies of the Ca1-solvent and Ca2-solvent are −1099.8 (107.4) and −2152.8 (129.8) kJ mol^{-1}, respectively, indicating that Ca2-solvent interaction is stronger than Ca1-solvent interaction (Fig. 3c). Therefore, the stronger electrostatic interaction of Ca2 with water molecules provide the potential for Ca2 to diffuse away from the protein, whereas the interaction of the Ca1-water can not overcome the interaction of the Ca1-protein, thus keeping Ca1 at its binding site.

It should be pointed out that van der Waals interactions of the Ca^{2+}-protein and Ca^{2+}-solvent are only minute (~27 and ~98 kJ mol^{-1} for Ca1; ~−0.3 and ~116 kJ mol^{-1} for Ca2) in comparison to electrostatic interactions. Therefore, we consider that the van der Waals interactions contribute negligibly to Ca^{2+} binding affinity.

Inter-residue hydrogen bonds and salt bridges around the Ca^{2+}-binding sites

Analyses of RMSD and B-factor reveal that Ca^{2+} removal increases the overall flexibility of proteinase K to a certain extent. However, the Ca^{2+}-free proteinase K appears to be stable during simulation because of its very similar geometrical properties to those of the native form. It is well known that hydrogen bonds play an important role in stabilizing protein structure due to their large number and wide distribution, and that salt bridges mainly make a contribution to the stability of local structure due to their localized distribution [6, 30]. Here in this study, we mainly focus on the hydrogen bonds and salt bridges contributing to the stability of the local structures around the Ca1 and Ca2 cations.

The hydrogen bonds involving residues that are within 8 Å of Ca1 and Ca2 with occurrence greater than 10% during the 10–50 ns simulations of both forms of proteinases K are shown in Table 2. The long-lived and low-frequency hydrogen bonds are arbitrarily defined as those having an occurrence greater than 70% and less than 50%, respectively. In the native proteinase K, the long-lived hydrogen bonds around Ca1 include Q149···V155, V157···C178, S176···E174, T179···D200, T179···I201 and V180···V157, which are also long-lived hydrogen bonds in the Ca^{2+}-free form—an indication that Ca1 removal does not disrupt these *inter*-residue interactions crucial for the stability of this local structure. Only in one residue pair, C178···V155, can hydrogen bond occurrence be found with significant increase upon Ca1 removal. This enhanced hydrogen bond seems to make more contribution to the stability of the Ca1 site when compared to the native proteinase K. Analogously, five common long-lived hydrogen-bonding residue pairs around the Ca2 site

Table 2 *Inter*-residue hydrogen bonds around the Ca1 and Ca2 site (10–50 ns)

Hydrogen bonds	Occurrence (%)[a]		
	Native	Ca^{2+}-free	Change[b]
Around Ca1 site			
Q149···V155	76.2	82.9	+
Q149···S176	13.4	3.9	−
Q149···C178	17.5	7.4	−
V157···C178	97.2	96.5	−
S176···E174	87.4	80.8	−
V177···E174	31.6	38.9	+
C178···V155	66.8	88.2	+ +
T179···D200	94.1	99.8	+
T179···I201	73.2	78.9	+
Around Ca2 site			
R12···W8	86.5	93.3	+
R12···S15	59.9	68.5	+
R12···T16	38.3	30.2	−
R12···D187	77.0	72.5	−
R12···N257	82.4	78.9	−
S15···S17	20.4	28.2	+
S15···N257	22.3	20.9	−
T16···Y274	14.2	17.1	+
D187···T262	79.4	81.1	+
D257···K258	13.1	10.6	−
N257···L272	94.7	97.1	+
L261···G259	9.2	14.9	+

[a] Only hydrogen bonds occurring for more than 10% during any of the 10–50 ns simulations are listed

[b] Changes in the hydrogen bond occurrence when Ca^{2+} is removed. The slight increase/decrease (less than ~10%) is indicated by "+"/"–", respectively; the significant increase/decrease (more than ~20%) is indicated by "+ +"/"– –", respectively

(Table 2), which show only minimal differences in occurrence between these two simulations, may act as an important factor in stabilizing the Ca2 site local structure. The minor differences in occurrences of both long-lived and low-frequency hydrogen bonds around the Ca2 site may be explained by the diffusion of Ca2 from its binding site during simulation of the native proteinase K.

A salt bridge is considered to be formed if the distance between the center of mass of the carboxyl group in acidic residue and that of the amino group in basic residue is within 10 Å. Among those residues that are within 8 Å of the Ca1, only one salt bridge, D200:R250 is found, with average distance (SD) between their charged groups being 9.4 (1.5) and 7.9 (2.1) Å for the native and Ca^{2+}-free proteinases K, respectively. D200 is one of the Ca1-binding residues, and therefore Ca1 removal provides a certain flexibility for D200 to shift toward R250, thus explaining their reduced distance in the Ca^{2+}-free form. However, such an enforced electrostatic interaction could make more contribution to the stability of the Ca1 site upon Ca^{2+} removal.

There are three salt bridges, R12:D187, R12:D260 and D254:K258 that are found near the Ca2 site in both forms of proteinases K. The average distances (SDs) between charged groups of these three residue pairs are 8.1 (0.8), 5.8 (1.4) and 7.2 (1.5) Å in the native proteinase K and 8.6 (1.1), 5.3 (1.3) and 7.7 (1.9) Å in the Ca^{2+}-free form, respectively, indicating that Ca2 diffusion (in the native form) or removal (in the Ca^{2+}-free form) does not break these electrostatic interactions contributing to the stability of the Ca2 site local structure.

Essential dynamics

ED properties of the native and Ca^{2+}-free proteinases K were analyzed and compared. Plots of the eigenvalues and the cumulative contribution of eigenvectors to the total mean square fluctuation (TMSF) as a function of eigenvector index were shown in Fig. S5. Both eigenvalue curves are steep, with the first 30 eigenvectors contributing 74.6% and 76.2% to TMSF of the native and Ca^{2+}-free proteinases K, respectively, indicating that most of the internal motions of proteinase K are confined within a subspace of very small dimensions.

In order to estimate the influence of Ca^{2+} removal on the flexibility of proteinase K, TMSF and the fluctuations within a subspace defined by the first 30 eigenvectors were calculated. The results reveal that these fluctuations increase by 26.5% and 30.1%, respectively. To further evaluate the degree of motion similarity between these two states, the first 30 eigenvectors from different simulations were projected onto each other, yielding an overlap of ~75% between these two essential subspaces. Accordingly, we may conclude that, although Ca^{2+} removal increases the overall conformational freedom of the enzyme structure, these two forms of proteinase K still share many common motional modes.

The differences in molecular motions between these two forms of proteinase K were investigated by performing a combined analysis. Figure 4 shows the properties of the projected eigenvectors. Only in the case of the first eigenvector can the projection be found with distinct distributions (Fig. 4a and b)—an indication that significant differences in large concerted motions occur only in the subspace described by eigenvector 1. The second combined eigenvector demonstrates a large degree of overlap in positional displacement distribution, suggesting many similar motions occurred in these two forms of proteinases K. The narrow Gaussian displacement distributions observed after the first few eigenvectors suggest the restrained harmonic motions common to both forms of proteinase K.

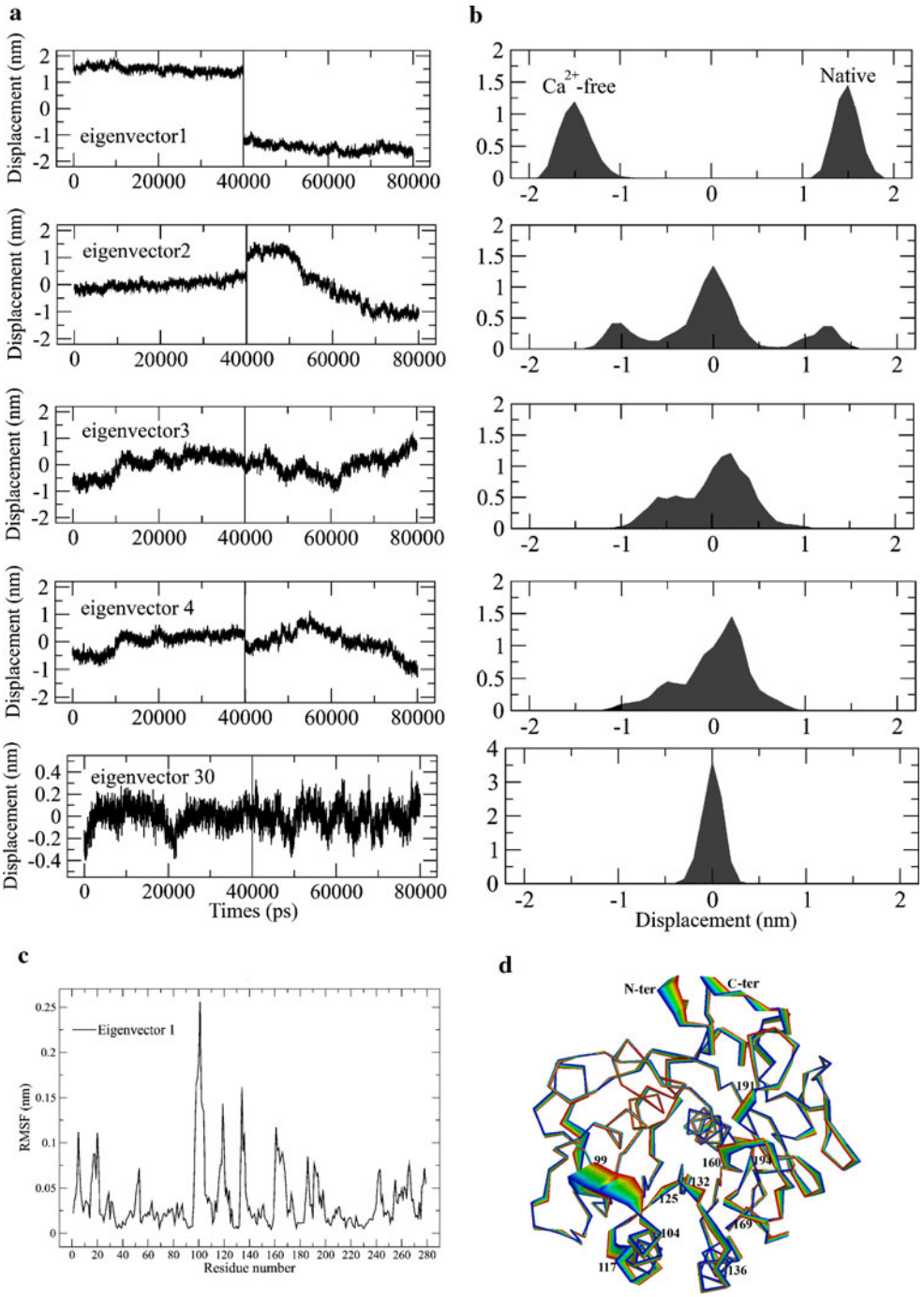

Fig. 4 Properties of the projections of the merged trajectory onto the "combined" eigenvectors. (**a**) Projections of the merged trajectory (native: 0–40 ns; Ca^{2+}-free: 40–80 ns) onto eigenvectors 1–4 and 30. (**b**) Plots of the distributions of these projections. Distinct distribution can only be found in the projection along eigenvector 1. (**c**) C_α RMSF of the first eigenvector as a function of residue number. (**d**) Projection extremes along the first eigenvector. The linear interpolations between the two extremes are colored from blue (the native proteinase K) to red (the Ca^{2+}-free proteinase K) to highlight the primary structural differences between these two states

The C_α RMSF and the projection extremes of eigenvector 1 are shown in Fig. 4c and d, respectively. Note that the linear interpolations between the two extremes are not a conformational transition pathway between the Ca^{2+}-bound and Ca^{2+}-free states. We emphasize that they offer only an impression of the character of a probably preferred mode of protein motions. The significant shifts (RMSF > 0.075 nm) are observed in regions of the N-, C-termini and residues 99–104, 117–125, 132–136, 160–169, 191–194, and 241–243. Of these, residues 160–169 and 191–194 are located relatively close to Ca1 site. Visualization of the motions along eigenvector 1 (Fig. 4d) reveals that the displacement of segment 160–169 originates from the concerted displacement of segment 191–194, which moves away from the Ca1 site when Ca^{2+} is removed. This motion is then transmitted from the Ca1 site *via* segment 160–169 to segment 132–136, which is part of the substrate-binding pockets (S1 and S4) and moves away from the S1 pocket, thus resulting in the opening of this pocket. The largest displacement is seen in the segment 100–104. This

segment, together with segment 132–136, moves concertedly toward the S4 pocket leading to the closure of this pocket. Another large displacement is observed in the loop region comprising residues 117–125, which is located between the α3 following segment 100–104 and the β4 preceding segment 132–136. Both the α helix and β strand show only small fluctuations, indicating that the large displacements of segments 100–104 and 132–136 are mediated by structural changes in the loop region 117–125. This is an interesting phenomenon, as several studies [31–33] have demonstrated that the dynamic behavior in regions opposite the substrate-binding site could play a role in modulating the dynamics of the substrate-binding pockets.

Despite these concerted motions caused by Ca^{2+} removal, no apparently large displacement is observed for the catalytic triad. The RMSF values of D39, H69 and S224 are 0.005, 0.006 and 0.018 nm, respectively. These results are in agreement with the results of B-factor analyses, both indicating that Ca^{2+} removal does not affect the fluctuations in the catalytic triad.

Figure 5 shows a comparison of the MSD of the projections of individual trajectories onto combined eigenvectors, which reveals that the third eigenvector of the native proteinase K and the second eigenvector of the Ca^{2+}-free proteinase K have the largest MSD values, indicating that the most significant conformational travel/motion occurs in these two subspaces. A comparison of RMSF values between these two eigenvectors (see Online Resource Fig. S6) exhibits that upon Ca^{2+} removal, the motions in many regions such as the N- and C-termini and the loops surrounding residues 60, 100, 119, 193, 215 and 241 increased, whereas the motions in segments 132–135, 160–169, and 218–222 decreased. These results agree,

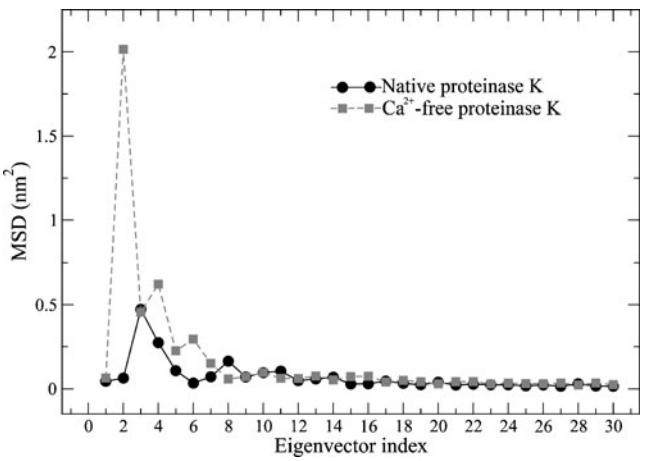

Fig. 5 MSD of the projections of individual trajectories onto the "combined" eigenvectors as a function of eigenvector index. The MSD values of the native and Ca^{2+}-free proteinases K are indicated by black circles and gray squares, respectively

to a large extent, with the results of comparative B-factor analyses, both suggesting that Ca^{2+} removal enhances the overall flexibility of the structure while reducing mobility in a limited number of regions around the substrate-binding site. It should be noted that the large conformational displacements have already been seen along eigenvector 1 derived from the combined analysis (Fig. 4c and d), which describes the static shift between the equilibrium structures of these two simulations.

Discussions

Although both forms of proteinase K displayed very similar geometrical properties during MD simulations, subtle differences in certain geometrical properties can still reflect the fact that the Ca^{2+}-free proteinase K assumes a more open and relaxed conformation than the native form. The removal of Ca^{2+} mainly increases flexibility in the loops/links as well as in the N- and C-termini, whereas the structural rigidity of the helices and central β sheet appears not to be affected. Interestingly, both B-factor and ED analyses reveal that the three segments comprising residues 132–136, 160–169 and 220–223 show larger fluctuations in the native proteinase K than in the Ca^{2+}-free form (Fig. 2 and Fig. S6). These segments participate in both formations of the S1 or S4 pockets [6, 10], thus resulting in higher flexibility of these pockets in the native proteinase K.

Many studies [13, 31–36] have suggested that the presence of flexibility within the substrate-binding site is necessary for efficient substrate recognition and binding due to the mechanism of induced fit [37, 38] or conformational selection [39]. Particularly, the crystallographic and biochemical studies [40] on a proteinase K-like enzyme, SPRK from the psychrotroph *Serratia* species (PDB code 2B6N), revealed that its rigid substrate-binding site not only reduced the binding affinity toward the synthetic substrate suc-Ala-Ala-Pro-Phe-nitroanilide, but also influenced its substrate specificity profile. Mutagenesis studies on α-Lytic protease (a serine protease belonging to the trypsin family) [41] and cold-adaptive alkaline phosphatase [35] revealed that higher substrate binding affinity was directly correlated to higher structural flexibility of the substrate-binding site. In the case of proteinase K, experimental study revealed that the native proteinase K showed slightly higher affinity toward the synthetic substrate succinyl-A-A-A-p-nitroanilide than the Ca^{2+}-free form [12]. Taken together, we consider that the observed difference in the flexibility of the substrate-binding regions between these two forms of proteinase K, although not very apparent, could be proposed to explain their differential substrate affinity. The relatively higher flexibility of the substrate-binding regions observed in the native proteinase K may make it

easy to overcome the entropic barriers, thereby contributing to efficient binding of the substrate to this enzyme form.

The two segments of residues 100–104 and 132–136 are the most important substrate-binding regions as they can form a three-stranded antiparallel β sheet with the substrate [10]. Our MD simulations indicate the segment 100–104 undergoes larger conformational fluctuations than the segment 132–136. The possible reasons for this are: i) the segment 100–104 is not involved in the secondary structure hydrogen bonding but exposes well to the solvent, thus resulting in a high conformational freedom; ii) for the segment 132–136, its N-terminus is linked to the central β sheet through β6 and its C-terminal residues 135–136 form an antiparallel β-sheet with residues 169–170, resulting in a tight anchor with the bulk of the enzyme structure. However, the flexibility of the segment 132–136 may be more important for substrate binding than that of the segment 100–104, as the former is part of both the S1 and S4 pockets. Sequence comparison of the proteinases K from various species reveals a high degree of conservation for three consecutive glycines at positions 134–136 [42], which may aid in enhancing the flexibility of this segment, facilitating in turn the binding of the P1 and P4 substrate residues with large side chains.

Our MD simulations indicate that the segment 100–104 has a larger conformational freedom in the Ca^{2+}-free proteinase K than in the native form. This may be caused by the overall relaxation of the molecule in the absence of Ca^{2+}. However, why does the presence of Ca^{2+} cause a higher flexibility in the segment 132–136? ED analyses indicate that, when Ca^{2+} is present, the large displacements of the segment 132–136 originate from large fluctuations of the segment 160–169, which are transmitted to segment 132–136 through the antiparallel β-sheet formed between the C-terminal residues of these two segments. We conclude that, therefore, the presence of the Ca1 could not only stabilize the Ca1 site, but also enhance the conformational freedom of the substrate-binding site. As explained in the previous section, this higher flexibility could enhance substrate affinity to some certain extent, although the experimentally determined K_m values for both forms of proteinase K are in the same order of magnitude [12].

As mentioned in the "Introduction" section, the catalytic activities of the proteinase K toward the same substrate upon Ca^{2+} removal have been reported to be almost unchanged or reduced. Bajorath and co-workers attributed the observed reduction in enzymatic activity of the Ca^{2+}-free proteinase K to the changes in geometries of the substrate-binding region and catalytic triad: the substrate-binding segment 100–104 shifted by 0.8 Å and the hydrogen bond distance between S224 O_γ and H69 N_ε increased by 0.5 Å upon Ca^{2+} removal [11]. Our MD simulations reveal that in both the native and Ca^{2+}-free proteinases K, the segment 100–104 exhibits very large fluctuations with amplitudes greater than 1.5 Å; and the S224 $O_\gamma \cdots$ H69 N_ε hydrogen bond length distribution peaks at 2.8 Å with a half width of 0.35 Å. These results imply that there are no remarkable changes in dynamics of segment 100–104 (although larger conformational freedom was observed in this segment in the Ca^{2+}-free proteinase K as discussed above) and in hydrogen bond distance of S224 $O_\gamma \cdots$ H69 N_ε upon Ca^{2+} removal. Therefore, the subtle geometrical changes observed between the static crystal structures may not be enough to interpret the reduced activity of the Ca^{2+}-free proteinase K. Interestingly, in a latter study Müller and co-workers suggested that the Ca^{2+}-free proteinase K tended to precipitate irreversibly when air bubbles existed in buffer, thus leading to a much reduced effective concentration in assay that was responsible for the experimentally observed lowered activity of the Ca^{2+}-free proteinase K [12]. Furthermore, our comparative B-factor analyses (Fig. 2) reveal that the removal of Ca^{2+} has only minor influence on the thermal motion of the catalytic triad; and the ED analyses (Fig. 4c, d and Fig. S6) reveal that the catalytic triad residues are not involved in the large concerted fluctuations originating from Ca1 site upon Ca^{2+} removal. These results support, and further rationalize Müller's experimental results, namely that Ca^{2+} removal does not affect dramatically the catalytic activity of proteinase K.

Diffusion of the Ca2 from its binding site suggests that Ca2 is indeed a weakly bound Ca^{2+}. This is consistent with an NMR study on the serine protease PB92 [13], which revealed that the weak Ca^{2+}-binding site was only partially occupied by Ca^{2+} in solution. In addition, crystallographic studies on the proteinase K-like proteases, SPRK (PDB code 2B6N) [40] and VPRK (PDB code 1SH7) [30], show the lack of Ca^{2+} occupancy at the structurally equivalent Ca2 site. We consider that the presence of Ca2 in the crystal structure of proteinase K is the consequence of relatively high calcium concentration in crystallization buffer [43, 44]; and the bound Ca2 could be further stabilized by the cryocooling of the crystallization condition. In contrast, the relatively higher simulation temperature and the lack of extra calcium in solvation box lead to the limited occupation of Ca2 at its binding site. Despite its stabilizing effect on the N- and C-terminal regions of proteinase K, it seems likely that Ca2 is not crucial for structural stability as its removal does not impair the overall stability of the enzyme. We suggest that other factors such as hydrogen bonds and salt bridges are important for the stability of this local structure. For example, four long-lived *inter*-residue hydrogen bonds (R12···W8, R12···S15, R12···D187 and D187···T262), which are common to both forms of proteinases K, form two *inter*-crossed hydrogen bond

networks, R12···[W8, S15, D187] and D187···[R12, T262]. In conjunction with a common salt bridge network, D187:R12:D260 (see Online Resource Fig. S7), these *inter*-residue interaction networks make a substantial contribution to the stability of the N- and C-terminal regions whether Ca2 is present or not.

Our simulations showed that Ca1 resided stably at its binding site. However, in the crystal structure of SPRK there is no Ca^{2+} that is found to occupy the structurally equivalent Ca1 site, implying that Ca1 may not be indispensable for the stability of the protein structure. Similar to the situation of the Ca2 site, five long-lived hydrogen bonds and one salt bridge around the Ca1 site, which were found commonly in both simulations, could contribute to the stability of this local structure.

ED analyses reveal that TMSF of the enzyme molecule increases by ~30% upon Ca^{2+} removal, but at the same time these two forms of proteinase K still share many common motional modes. Although the conformational displacements originating from the Ca1 site upon Ca^{2+} removal can be transmitted to substrate-binding regions and lead to the closings/openings of the substrate-binding pockets, such breathing motions cannot be related directly to a specific functional task (such as substrate binding and product release as described for *streptococcus agalactiae* hyaluronate lyase [45]) as our ED analysis is performed on the combined trajectory. Here we merely emphasize large conformational differences resulting from Ca^{2+} removal and focus on the changes in flexibility of substrate-binding regions when Ca^{2+} is present or absent. Individual ED analyses also indicate that the segments of residues 100–104, 132–136 and 160–169 undergo large concerted conformational displacements, which lead to the closings/openings of the substrate-binding pockets and facilitate binding/release of the peptide substrate/product [42].

Conclusions

In summary, we have investigated the effect of calciums on the stability and molecular motions of proteinase K by using MD simulations. Comparative analyses of the geometrical properties, B-factors and ED properties point to a common conclusion, namely that Ca^{2+} removal increases the overall flexibility of proteinase K. This provides an explanation for the decreased thermal stability of the Ca^{2+}-free proteinase K when compared to the native form. Combined ED analyses reveal that, although large static differences can be observed in equilibrium structures between the native and Ca^{2+}-free proteinases K, Ca^{2+} removal does not affect the fluctuations of catalytic triad residues, thus supporting and explaining the almost unchanged catalytic activity of proteinase K upon Ca^{2+} removal. The enhanced flexibility of some substrate-binding segments caused by the presence of Ca1 could be used to explain the increased substrate binding affinity of the native proteinase K when compared to the Ca^{2+}-free form. This work not only elucidates the mechanism responsible for the stabilizing effect of Ca^{2+} on the enzyme structure, but also facilitates the interpretation of the previously published experimental data.

Acknowledgments The authors thank High Performance Computer Center of Yunnan University for computational support. This work was funded by the National Natural Science Foundation of China (approved number 30860011) and partially supported by grants from Yunnan Province (2006C008M, 2007C163M, 07Z10756 and 2007PY-22) and Innovation Group Project from Yunnan University (KL070002).

References

1. Ebeling W, Hennrich N, Klockow M, Metz H, Orth HD, Lang H (1974) Eur J Biochem 47:91–97
2. Pahler A, Banerjee A, Dattagupta JK, Fujiwara T, Lindner K, Pal GP, Suck D, Saenger W (1984) EMBO J 3:1311–1314
3. Siezen RJ, Leunissen JAM (1997) Protein Sci 6:501–523
4. Lizardi PM, Engelberg A (1979) Anal Biochem 98:116–122
5. Shaw WV (1987) Biochem J 246:1–17
6. Liu SQ, Meng ZH, Yang JK, Fu YX, Zhang KQ (2007) BMC Struct Biol 7:33
7. Betzel C, Pal GP, Struck M, Jany KD, Saenger W (1986) FEBS Lett 197:105–110
8. Betzel C, Pal GP, Saenger W (1988) Eur J Biochem 178:155–171
9. Betzel C, Gourinath S, Kumar P, Kaur P, Perbandt M, Eschenburg S, Singh TP (2001) Biochemistry 40:3080–3088
10. Wolf WM, Bajorath J, Müller A, Raghunathan S, Singh TP, Hinrichs W, Saenger W (1991) J Biol Chem 266:17695–17699
11. Bajorath J, Raghunathan S, Hinricks W, Saenger W (1989) Nature 337:481–484
12. Müller A, Hinrichs W, Wolf WM, Saenger W (1994) J Biol Chem 269:23108–23111
13. Martin JR, Mulder FAA, Karimi-Nejad Y, van der Zwan J, Mariani M, Schipper D, Boelens R (1997) Structure 5:521–532
14. Lindahl E, Hess B, van der Spoel D (2001) J Mol Model 7:306–317
15. Kutzner C, van der Spoel D, Fechner M, Lindahl E, Schmitt UW, de Groot BL, Grubmuller H (2007) J Comput Chem 28:2075–2084
16. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J (1981) Interaction models for water in relation to protein hydration. In: Pullman B (ed) Intermolecular forces. Reidel, Dordrecht, the Netherlands, pp 331–342
17. Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A, Haak JR (1984) J Chem Phys 81:3684–3690
18. Tironi IG, Sperb R, Smith PE, van Gunsteren WFA (1995) J Chem Phys 102:5451–5459
19. Hess B, Bekker H, Berendsen HJC, Fraaije J (1997) J Comput Chem 18:1463–1472
20. Amadei A, Linssen ABM, Berendsen HJC (1993) Proteins 17:412–425
21. Balsera MA, Wriggers W, Oono Y, Schulten K (1996) J Phys Chem 100:2567–2572

22. de Groot BL, Hayward S, van Aalten DMF, Amadei A, Berendsen HJC (1998) Proteins 31:116–127
23. van Aalten DMF, Jones PC, de Sousa M, Findlay JBC (1997) Protein Eng 10:31–37
24. Vreede J, van der Horst MA, Hellingwerf KJ, Crielaard W, van Aalten DMF (2003) J Biol Chem 278:18434–18439
25. Liu SQ, Fu YX, Liu CQ (2007) Chin Sci Bull 52:3074–3088
26. Liu SQ, Liu CQ, Fu YX (2007) J Mol Graphics Modell 26:306–318
27. Liu SQ, Liu SX, Fu YX (2008) J Mol Model 14:857–870
28. van Aalten DMF, Amadei A, Linssen ABM, Eijsink VGH, Vriend G (1995) Proteins Struct Funct Genet 22:45–54
29. Darden TA, York DM, Pedersen LG (1993) J Chem Phys 98:10089–10092
30. Arnorsdottir J, Kristjansson MM, Ficner R (2005) FEBS J 272:832–845
31. Peters GH, Bywater RR (1999) Protein Eng 12:747–754
32. Peters GH, Frimurer TM, Andersen JN, Olsen OH (1999) Biophys J 77:505–515
33. Peters GH, Frimurer TM, Andersen JN, Olsen OH (2000) Biophys J 78:2191–2200
34. Wade BC, Gabdoulline RR, Ludemann SK, Lounnas V (1998) Proc Natl Acad Sci USA 95:5942–5949
35. Gudjónsdóttir K, Ásgeirsson B (2008) FEBS J 275:117–127
36. Miller DW, Agard DA (1999) J Mol Biol 286:267–278
37. Koshland DE (1958) Proc Natl Acad Sci USA 44:98–104
38. Henzler-Wildman K, Kern D (2007) Nature 450:964–972
39. Lange OF, Lakomek NA, Farès C, Schröder GF, Walter KF, Becker S, Meiler J, Grubmüller H, Griesinger C, de Groot BL (2008) Science 320:1471–1475
40. Helland R, Larsen AN, Smalas AO, Willassen NP (2006) FEBS J 273:61–71
41. Bone R, Silen JL, Agard DA (1989) Nature 339:191–195
42. Liu SQ, Meng ZH, Fu YX, Zhang KQ (2010) J Mol Model 16:17–28
43. Bajorath J, Hinrichs W, Saenger W (1988) Eur J Biochem 176:441–447
44. Siezen RJ, de Vos WM, Leunissen JAM, Dijkstra BW (1991) Protein Eng 4:719–737
45. Mello LV, Groot BL, Li S, Jedrzejas MJ (2002) J Biol Chem 277:36678–36688