Available online at www.sciencedirect.com
ScienceDirect
Biochimica et Biophysica Acta 1764 (2006) 1397–1406
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
www.elsevier.com/locate/bbapap

Flexibility and enzymatic cold-adaptation: A comparative molecular dynamics investigation of the elastase family

Elena Papaleo, Laura Riccardi, Chiara Villa, Piercarlo Fantucci, Luca De Gioia *

Department of Biotechnology and Bioscience, University of Milano-Bicocca, P.za della Scienza 2, 20126, Milan, Italy

Received 31 March 2006; received in revised form 14 June 2006; accepted 26 June 2006
Available online 1 July 2006

Abstract

Molecular dynamics simulations of representative mesophilic and psycrophilic elastases have been carried out at different temperatures to explore the molecular basis of cold adaptation inside a specific enzymatic family. The molecular dynamics trajectories have been compared and analyzed in terms of secondary structure, molecular flexibility, intramolecular and protein–solvent interactions, unravelling molecular features relevant to rationalize the efficient catalytic activity of psychrophilic elastases at low temperature. The comparative molecular dynamics investigation reveals that modulation of the number of protein–solvent interactions is not the evolutionary strategy followed by the psycrophilic elastase to enhance catalytic activity at low temperature. In addition, flexibility and solvent accessibility of the residues forming the catalytic triad and the specificity pocket are comparable in the cold- and warm-adapted enzymes. Instead, loop regions with different amino acid composition in the two enzymes, and clustered around the active site or the specificity pocket, are characterized by enhanced flexibility in the cold-adapted enzyme. Remarkably, the psycrophilic elastase is characterized by reduced flexibility, when compared to the mesophilic counterpart, in some scattered regions distant from the functional sites, in agreement with hypothesis suggesting that local rigidity in regions far from functional sites can be beneficial for the catalytic activity of psychrophilic enzymes.

Keywords: Molecular dynamics simulations; Serine protease; Elastase; Cold adaptation; Psychrophilic enzyme; Flexibility

1. Introduction

Psychrophilic organisms live at low temperature, where most other species cannot grow. Enzymes from cold-adapted organisms are able to catalyze reactions at temperatures at which enzymes from mesophilic or thermophilic organisms are generally unable to sustain a viable metabolism [1–3]. In recent years, increasing interest has been directed to the origin of enzyme cold adaptation both for understanding key factors affecting protein folding and for designing biocatalysts with enhanced activity at low temperature or increased thermostability [2–5]. Moreover, enzymes synthesized by cold-adapted organisms have considerable biotechnological potential, offering environmentally acceptable and energy saving applications [4,5].

Cold active enzymes are commonly characterized by increased catalytic efficiency at low temperatures, measured as k_{cat}/K_m ratio, and by significantly increased thermolability, which is believed to be a consequence of enhanced peptide chain flexibility [2,6,7]. It is generally assumed that molecular flexibility at low temperature has been obtained during evolution by decreasing or weakening intramolecular interactions that stabilize the mesophilic counterparts [2]. The question whether low thermal stability reflects increased molecular flexibility has been widely discussed [2,8] and recently the relationship between flexibility and catalytic activity has been investigated in uracil DNA glycosidase [8], corroborating the hypothesis that cold adaptation is not necessarily ascribed to overall increased flexibility but it is most likely the result of increased flexibility of only some of the protein regions [2,6].

The recent elucidation of X-ray structures of cold-adapted enzymes [2,10–15], as well as homology modeling studies [16–18], has shed light on some key and peculiar structural characteristics of psychrophilic enzymes. Cold-adapted enzymes generally differ from their mesophilic or thermophilic homologues only for limited structural characteristics

* Corresponding author. Tel.: +39 0264483463.
E-mail address: luca.degioia@unimib.it (L. De Gioia).

1570-9639/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2006.06.005

and/or few amino acidic substitutions. It has been shown that amino acid substitutions in functional regions, as well as more extended solvent exposed hydrophilic surfaces, can lead to enhanced structural flexibility [2]. Moreover, the evaluation of electrostatic free energy contribution to protein stability has showed that electrostatics can play an important role in both cold- and heat-adaptation of enzymes [19,20].

Comparative studies aimed at clarifying key factors for adaptation of psychrophilic enzymes have also been recently reported [21,22], giving valuable indications on evolutionary strategies leading to cold-adaptation. However, a unified theory cannot be formulated because cold-adaptation in different enzyme families has been most likely obtained according to different evolutionary strategies [2].

Molecular dynamics (MD) is a suitable tool to evaluate flexibility and correlate it to protein function [8,23–25]. In particular, MD simulations can be useful in the rationalization and interpretation of experimental results, giving the possibility to interpolate or extrapolate experimental data into regions hardly accessible experimentally [26–28]. MD studies aimed at clarifying structure–function relationships in cold-adapted enzymes have been recently reported for pancreatic trypsins [29], which are homologous to elastase, and for uracil DNA glycosidases [8].

In light of the above observations, the molecular basis of cold adaptation inside the specific enzymatic class of pancreatic elastases, which are endopeptidases of the serine-protease family, have been explored by MD simulations. The choice of the elastase family was driven by the observation that serine proteases are thoroughly characterized and a sufficiently large number of sequences and structural data are available. To sample efficiently the potential energy surface [26–28], multiple MD simulations for a psychrophilic (*Salmo salar*, Atlantic salmon) and a mesophilic (*Sus scrofa*, native porcine) elastase were carried out in explicit solvent at two different temperatures (283 and 310 K). The temperatures of 283 and 310 K have been chosen *ad hoc* as they reflect the growth temperature of the psychrophilic and mesophilic organisms, respectively [30]. In sum, trajectories of 48 ns have been collected for each protein investigated.

The MD trajectories of native porcine (PE) and Atlantic salmon (SE) elastases have been analyzed in terms of secondary structure content, molecular flexibility, solvent accessibility degree, intramolecular interactions and protein–solvent interactions. Results lead to the conclusion that the Atlantic salmon elastase is characterized by increased flexibility in loop regions localized around the active site, whereas the most flexible protein portions in the mesophilic porcine elastase are localized in regions far from the active site.

2. Materials and methods

MD simulations were performed using the GROMACS software and GROMACS force field [31], implemented on a parallel architecture. The X-ray structures of native porcine (PE) and Atlantic salmon (SE) elastases (PDB entries 1LYV [32] and 1ELT [33], respectively) were used as starting point for MD simulations. Protein structures, including crystallographic water molecules and the calcium ion, were soaked in a dodecahedral box of 4682 (PE simulations) or 5181 (SE simulations) SPC water molecules [34]. All the protein atoms are at a distance equal or greater than 0.5 nm from the box edges.

The ionization state of charged residues was set to be consistent with neutral pH: Lys and Arg residues were positively charged, whereas Asp and Glu were negatively charged. The tautomeric form of histidine residues was derived using GROMACS tools and confirmed by visual inspection of the molecular environment of each histidine. With the above assumptions, both SE and PE are characterized by positive overall charge. In order to neutralize the overall charge of the system, a number of water molecules equal to the protein net charge was replaced by Cl^- ions.

In summary, the PE simulations were carried out on a system containing 16402 atoms, including 2350 protein atoms, 14046 water atoms and 1 Ca^{2+} and 6 Cl^- ions. The SE simulations were carried out on a system of 17808 atoms, including 2260 protein atoms, 15543 water atoms and 1 Ca^{2+} and 5 Cl^- ions.

Initially, solvent molecules were relaxed by molecular mechanics (steepest descent method, 1000 steps). The optimization step was followed by 30 ps MD at 300 K (time step 1 fs) while restraining protein atomic positions using a harmonic potential. During equilibration, the coupling constant to the external bath (τ) [35] was set to 0.001 ps for both protein and non-protein elements.

MD simulations were performed in the NPT ensemble at 283 and 310 K, applying periodic boundary conditions and using an external bath with a coupling constant τ of 0.1 ps for the protein and for non-protein groups. Pressure was kept constant (1 atm) by modifying the box dimensions and the time-constant for the pressure coupling was set to 1 ps [35]. The LINCS algorithm [36] was used to constrain bond lengths, allowing to use a 2 fs time step. Electrostatic interactions were calculated using the Particle-mesh Ewald (PME) summation scheme [37]. Van der Waals and Coulomb interactions were truncated at 1.0 nm. The nonbonded pair list was updated every 10 steps and conformations were stored every 2 ps.

To improve the conformational sampling, two 12 ns simulations obtained initializing the MD runs with different Maxwellian distributions of initial velocities, were carried out for each protein system, both at 283 and 310 K.

The rmsd (root mean square deviation), which is a crucial parameter to evaluate the stability of MD trajectories, was computed for mainchain atoms using as reference the starting structure of the MD simulations. In addition, MD trajectories differing only for the distribution of initial velocities have been merged to compute rmsd matrices, in which the rmsd values computed comparing all pairs of frames are organized in a two-dimensional map which allows an evaluation of the re-sampling of similar substructures. The rmsd matrices representing merged trajectories have been processed using the Jarvis–Patrick method [38], to extract information on possible clusters of conformations. The average structure of the clusters, defined as the protein structure with the smallest average distance to the other structures belonging to the same cluster, have also been calculated [31].

The secondary structure content was calculated for all stored conformations using the DSSP program [39]. The root mean square fluctuation per residue (rmsf) was calculated on mainchain atoms. The protein–protein and protein–solvent hydrogen bonds (H-bonds) were evaluated by GROMACS subroutines. The solvent accessibility degree was evaluated with the NACCESS program [40] on structures collected from the trajectories every 10 ps, whereas the solvent accessible area was calculated on the three-dimensional average structures. The number of protein–solvent H-bonds has been normalized calculating the ratio between protein–solvent H-bonds and the solvent accessible area of polar atoms. Protein–protein H-bonds have been normalized calculating the ratio between the number of protein–protein H-bonds and the number of protein polar atoms.

Generalized order parameters (S^2) which are a measure of the degree of spatial restriction of motion have been calculated for the mainchain N–H bonds from the equilibrium trajectories as an approximation of the asymptotic value of the bond rotational autocorrelation function [41,42], according to:

$$S^2 = \lim_{t\to\infty}(C_I(t))$$

where $C_I(t)$ is defined as:

$$C_I(t) = \langle P_2(\mu(T)^*\mu(T+t))\rangle$$

$\mu(T)$ is the unit vector that describes the orientation of the N–H bond at time T, measured in the molecule-fixed frame, and $P_2(x)$ is the second Legendre polynomial. S^2 can take values ranging from 0, corresponding to completely isotropic motions, to 1, if the motions are completely restricted.

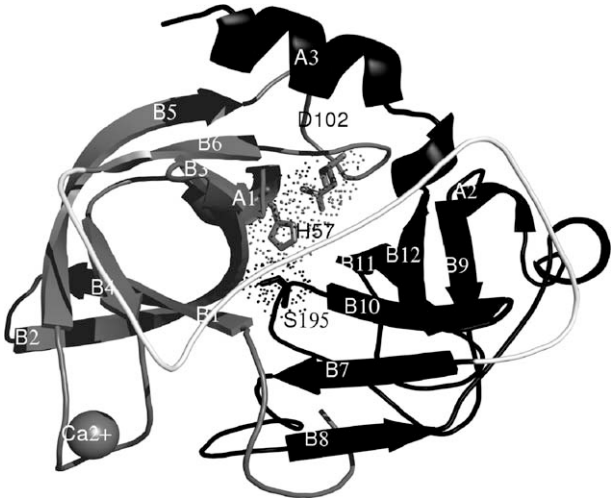

Fig. 1. 3D structure of elastases. The N-terminal domain, the interdomain loop, the C-terminal domain are shown in grey, light grey and black, respectively. The secondary structure elements are shown as ribbons, whereas the calcium ion and the residues forming the catalytic triad (H57, D102, S195) are shown as sphere/sticks and dots, respectively.

The N–H bond autocorrelation function $C_I(t)$ was computed for each residue (with the exception of the N-terminal valine and all proline residues of PE and SE). In particular, S^2 order parameters were calculated as the average value of $C_I(t)$ starting from $t=1$ ns.

To evaluate the significance of the obtained order parameters we estimated the drift of the $C_I(t)$ portion used to calculated S^2 by determining the skewness of its distribution.

The generalized order parameters were calculated for all the equilibrium portions of the trajectories and the average values for each protein residue.

To evaluate the dynamic behavior in the surrounding of an amino acid x, residues in the range of 0.5 nm from the center of mass of x were monitored during MD simulations. In particular, the persistency degree of each residue was defined as the ratio between the number of times that a residue appears in the surrounding of x and the total number of conformations analyzed.

The visual analysis of protein structures was carried out using VMD [43] and PyMOL (http:www.pymol.org).

The data extracted from the simulations have been compared to the data derived from the analysis of the experimental crystallographic structures, when available.

In the following, the numbering of residues is referred to PE unless otherwise specified. The N- (16–108) and C-terminal (136–245) domains of elastases are labeled as N and C, respectively. The β-strands are numbered from β1 to β12 while the α-helices are numbered from α1 to α3. PE and SE are formed by 240 and 236 amino acids, respectively.

3. Results and discussion

3.1. Sequence and structural features of salmon and porcine elastases

To better discuss results from MD simulations, a brief introduction about SE and PE structures, as well as their aminoacid sequences, is in order. The 3D-structure of elastases is folded into two domains characterized by an antiparallel β type fold: 12 β-strands and 3 α-helices are present both in PE and SE (Fig. 1, see also supplementary materials, Table 1S).

The primary sequences of PE and SE are extremely similar (161/236 identical amino acids, 68.2% sequence identity). Among the 76 amino acids substitutions, 45 correspond to completely unrelated amino acids (Fig. 2). Some differences in the amino acid composition of PE and SE are worth mentioning (Table 1). A lower number of positively charged (Arg and Lys) and long-chain hydrophobic (mainly Val) residues are present in SE, whereas the number of Gly and Met residues, which are often abundant in psychrophilic enzymes [2,3,18,21,22], is larger in SE. PE presents an Arg/(Arg+Lys) ratio two fold greater than SE (0.80 vs. 0.46, Table 1), in agreement with the observation that cold adapted enzymes are often characterized by a low relative arginine content [Arg/(Arg+Lys)] compared to mesophilic and thermophilic enzymes [2,3,18,21,22].

3.2. Stability of the MD trajectories and overall structural properties

MD simulations of SE and PE were carried out at 283 and 310 K (see Materials and methods). Several criteria have been used to evaluate convergence and stability of the MD trajectories. Indeed, it is well known that the analysis of multiple trajectories helps in the identification of recurring features and can allow to avoid artifacts arising from the simulation procedure [44,45]. Therefore, in order to sample efficiently the conformational space, two independent 12 ns MD simulations were carried out for every system both at 283 and 310 K, starting from the same atomic coordinates but using different sets of initial velocities, taken from a Maxwellian distribution. In the following, MD trajectories collected for the

```
SE          VVGGRVAQPNSWPWQISLQYKSGSSYYHTCGGSLIRQGWVMTAAHCVDSARTWRVVLGEH
PE          VVGGTEAQRNSWPSQISLQYRSGSSWAHTCGGTLIRQNWVMTAAHCVDRELTFRVVVGEH
            ****  ** **** ******:****: *****:**** **********   *:***:***

SE          NLNTNEGKEQIMTVNSVFIHSGWNSDDVAGGYDIALLRLNTQASLNSAVQLAALPPSNQI
PE          NLNQNNGTEQYVGVQKIVVHPYWNTDDVAAGYDIALLRLAQSVTLNSYVQLGVLPRAGTI
            *** * * ** : *: :::*  **:****:********* :: :*** ***: **   :*

SE          LPNNNPCYITGWGKTSTGGPLSDSLKQAWLPSVDHATCSSSGWWGSTVKTTMVCAGG-GA
PE          LANNSPCYITGWGLTRTNGQLAQTLQQAYLPTVDYAICSSSSYWGSTVKNSMVCAGGDGV
             * **:******** * * * *  :* **:**:** * **** :******::****** *

SE          NSGCNGDSGGPLNCQVNGSYYVHGVTSFVSSSGCNASKKPTVFTRVSAYISWMNGIM---
PE          RSGCQGDSGGPLHCLVNGQYAVHGVTSFVSRLGCNVTRKPTVFTRVSAYISWINNVIASN
            ***:******* * *** * *********  *** ::***************:* ::
```

Fig. 2. Pairwise sequence alignment of SE and PE. The identical (asterisk) and similar residues (double points) are highlighted.

Table 1
Amino acid compositions of SE and PE in comparison to the average composition of globular proteins

Amino acid	SE	PE	Frequency in globular proteins (%)
A	16 (6.8%)	17 (7.1%)	8.3
C	8 (3.4%)	8 (3.3%)	1.7
D	7 (3.0%)	7 (2.9%)	5.3
E	3 (1.3%)	4 (1.7%)	6.2
F	3 (1.3%)	3 (1.2%)	3.9
G	28 (11.9%)	25 (10.4%)	7.2
H	6 (2.5%)	6 (2.5%)	2.2
I	9 (3.8%)	10 (4.2%)	5.2
K	7 (3.0%)	3 (1.2%)	5.7
L	15 (6.4%)	18 (7.5%)	9.0
M	5 (2.1%)	2 (0.8%)	2.4
N	18 (7.6%)	17 (7.1%)	4.4
P	10 (4.2%)	7 (2.9%)	5.1
Q	10 (4.2%)	15 (6.2%)	4.0
R	6 (2.5%)	12 (5.0%)	5.7
S	31 (13.1%)	22 (9.2%)	6.9
T	16 (6.8%)	19 (7.9%)	5.8
V	20 (8.7%)	27 (11.2%)	6.6
W	10 (4.2%)	7 (2.9%)	1.3
Y	8 (3.4%)	11 (4.6%)	3.2
KR	13 (5.5%)	15 (6.2%)	
ED	10 (4.2%)	11 (4.6%)	
LVIFM	52 (22.0%)	60 (25.0%)	
AG	44 (18.7%)	42 (17.5%)	
STYNQ	83 (35.1%)	84 (35.0%)	
FWY	21 (8.9%)	21 (8.7%)	
R/(R+K)	0.46	0.80	

The amino acid composition of SE and PE was calculated according to SAPS [50] and the average composition for globular proteins has been taken from McCaldon and Argos [51]. The values for positively charged residues (KR), negatively charged residues (ED), short chain hydrophobic residues (AG), long chain hydrophobic residues (LVIFM), aromatic residues (FWY), polar residue (STYNQ) and the ratio R/[R+K] are also reported.

same system but characterized by different initial velocities are referred to as $_{sim1}$ and $_{sim2}$.

The rmsd values calculated for the mainchain atoms of SE and PE generally needed 4.5–5 ns to reach the stable states (Fig. 3, see supplementary materials, Fig. 1S).

As a consequence, to ensure that calculated parameters reflect the intrinsic properties of each system, the analyses of MD trajectories have been carried out by discarding the first 4.5 ns for all simulations.

Comparison of the initial structures with the protein conformations collected after 4.5–5 ns of simulation shows that no drastic conformational changes have taken place and only some secondary structure motifs present different length and a few surface loops change orientation.

The interatomic distances between Ca^{2+} and the atoms of the coordinating residues (70, 72, 75, 77 and 80), the potential and total energy of the system, as well as the protein gyration radius, were characterized by negligible fluctuations throughout the MD simulations, indicating stable trajectories (Table 2). In particular, the coordination sphere and geometry of Ca^{2+} observed during the trajectories is in agreement with the observed coordination in the X-ray structures of PE [32] and SE [33]. The gyration radii calculated on the X-ray structures are 1.67 and 1.63 nm for PE and SE, respectively, in good agreement with the protein gyration radii calculated during the trajectories (Table 2). Interestingly, experimental gyration radii are available also for bovine β-trypsin, which is a serine-protease evolutionary related to elastases, and the experimental values calculated in presence of Ca^{2+} fully agree with our estimated values [46].

Further insight into the configurations visited by the system have been obtained carrying out a cluster analysis on rmsd matrices (see Materials and methods). Two different clusters, corresponding to simulations initialized with different Maxwellian distributions, have been identified. Therefore, simulations initialized with different initial velocities apparently converge to different conformational basins. However, when the protein average structures representing the two clusters of each protein system were analyzed, it turned out that the only region differing significantly is the short α-helix α1 (see supplementary materials, Table 1S). In fact, the pairwise rmsd computed on the mainchain atoms of the other protein portions is as low as 0.10 nm (see supplementary materials, Fig. 2S).

When the average secondary structure content over time is considered (Table 3), differences between SE and PE are evident, mainly at 283 K: the cold adapted enzyme presents a lower number of amino acids characterized by β-sheet and α-helix structures, balanced by an increased number of residues in β-bridge and turn conformations. In particular, the α3-helix, at the C-terminal end, is shorter in the SE average structures (residues 231–243 and 231–240 in PE and SE simulations, respectively; see supplementary materials). The surface constitutes the interface through which a protein senses the surrounding environment and its characteristics can affect the dynamic properties of the native state. In particular, a more hydrophobic surface implies a lower number of favorable protein–solvent interactions and consequently reduced flexibility of the protein structure [2,3,29]. The intramolecular interactions, evaluated computing the total number of H-bonds during the MD trajectories, as well as the ratio between protein–protein H-bonds and polar protein atoms, are comparable in PE

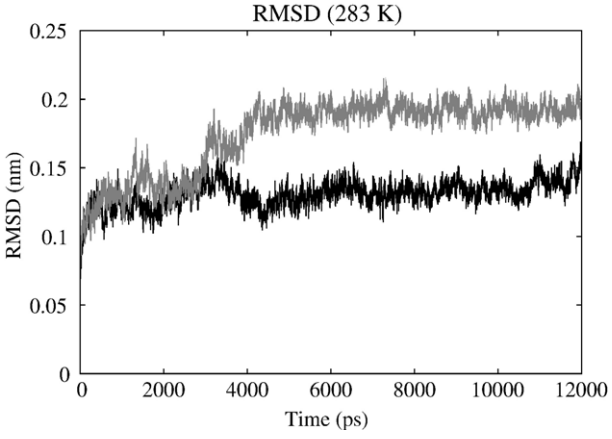

Fig. 3. Rmsd of mainchain atoms as a function of time. Rmsd as a function of time for SE (grey line) and PE (black line) at 283 K. For the sake of clarity only values relative to the simulation $_{sim1}$ are shown. For further information about rmsd profiles as a function of time see supplementary materials (Figure 1S).

Table 2
Average values and standard deviations (in parenthesis) of total and potential energy, gyration radius and distances between Ca^{2+} ion and the coordinating residues

	PE283	PE283II	PE310	PE310II	SE283	SE283II	SE310	SE310II
En. tot. (kJ/mol)	−188037.24 (423.07)	−188132.85 (423.50)	−176081.57 (432.00)	−176132.89 (441.15)	−203935.11 (428.99)	−203952.84 (438.36)	−190898.76 (450.71)	−190979.84 (446.89)
En. pot. (kJ/mol)	−226533.80 (421.20)	−226635.77 (425.11)	−218252.67 (434.25)	−218310.95 (447.13)	−245741.98 (431.49)	−245762.37 (439.42)	−236693.03 (455.02)	−236788.35 (448.32)
Rg (nm)	1.66 (0.005)	1.65 (0.005)	1.66 (0.006)	1.66(0.007)	1.62 (0.007)	1.63 (0.006)	1.63 (0.005)	1.63 (0.004)
Ca^{2+}-E70_OE1 (nm)	0.20 (0.05)	0.19 (0.06)	0.19 (0.06)	0.19 (0.07)	0.20 (0.04)	0.20 (0.03)	0.20 (0.05)	0.20 (0.04)
Ca^{2+}-O (72) (nm)	0.25 (0.05)	0.25 (0.08)	0.24 (0.07)	0.22 (0.06)	0.23 (0.05)	0.23 (0.05)	0.25 (0.05)	0.24 (0.06)
Ca^{2+}-O (75) (nm)	0.24 (0.06)	0.23 (0.08)	0.24 (0.07)	0.21 (0.09)	0.20 (0.06)	0.20 (0.05)	0.22 (0.06)	0.23 (0.09)
Ca^{2+}-O1 (77) (nm)	0.20 (0.08)	0.19 (0.08)	0.20 (0.07)	0.18 (0.09)	0.21 (0.08)	0.21 (0.09)	0.22 (0.09)	0.21 (0.09)
Ca^{2+}-E80_OE2 (nm)	0.19 (0.06)	0.18 (0.08)	0.18 (0.09)	0.18 (0.09)	0.16 (0.05)	0.16 (0.03)	0.16 (0.04)	0.16 (0.06)

and SE (Table 4). The average number of the protein–protein H-bonds calculated for PE and SE during the MD trajectories is slightly greater than the corresponding values computed on the X-ray structures (172 in PE and 167 in SE). Significant differences in protein–solvent interactions are evident comparing PE and SE: at both temperatures the average number of protein–solvent interactions is greater for PE (Table 4). However, it should be noted that the solvent accessible surface of PE and SE polar atoms is different in the two proteins. Therefore, the total number of protein–solvent H-bonds was normalized calculating the ratio between the number of protein–solvent H-bonds and the solvent accessible surface of polar atoms. Results indicate that the difference in the normalized number of protein–solvent H-bonds between PE and SE is negligible (Table 4).

3.3. Analysis of structural flexibility

As mentioned in Introduction, enhanced protein flexibility has been often related to cold-adaptation [2,3,8,28]. In order to evaluate and compare flexibility in PE and SE, the root mean square deviation (per residue) from the average structure obtained from merged MD trajectories (root mean square fluctuation, rmsf) was calculated and adopted as flexibility index (Fig. 4). The reliability of the rmsf analysis was evaluated splitting the MD trajectories (after equilibration) and computing rmsf values for every trajectory portion (not shown). Protein regions characterized by large rmsf values are generally located at corresponding positions in PE and SE, even though the intensity of the fluctuations is often different. As expected, regular secondary structure regions show small fluctuations during the simulations, whereas pronounced fluctuations are observed for some loop regions.

The rmsf calculated from the simulations exhibits a qualitative correspondence to the crystallographic B-factors in both proteins (Fig. 4). Indeed, it should be noted that the temperature factors may be affected by additional effects besides the internal fluctuations, such as crystal-packing forces and intermolecular interactions with nearest neighbors [47,48].

An interesting descriptor of backbone motion is the generalized order parameter (S^2), which characterizes the angular correlation for the dynamics of the N–H bond vector (Fig. 5). Results from the computation of the order parameters are in reasonable agreement with the rmsf profiles, with the greatest flexibility occurring in loop regions, while other secondary structural elements are more constrained.

To better highlight regions characterized by different flexibility in SE and PE, the rmsf profile of SE was subtracted from the corresponding profile obtained for PE (rmsf-diff, Fig. 6). The correspondence between residues of SE and PE was determined from a sequence alignment of the two proteins (Fig. 2). The regions characterized by an rmsf-diff value lower than −0.10 nm (indicating greater flexibility in SE) or higher than 0.10 nm (greater flexibility in PE) have been further analyzed (Table 5 and Fig. 7).

All residues characterized by significantly different flexibility in PE and SE are well exposed to the solvent (average

Table 3
Average number of residues with a given secondary structure in SE and PE at 310 and 283 K, as well as in the corresponding X-ray structures, according to the DSSP software

	PE			SE		
	X-ray	283 K	310 K	X-ray	283 K	310 K
COIL	62 [25.8%]	62.8 (4.7) [26.2%]	66.9 (4.1) [27.8%]	63 [26.7%]	63.2 (4.6) [26.8%]	63.4 (4.5) [26.9%]
β-sheet	82 [34.2%]	86.5 (5.0) [36.0%]	80.4 (4.5) [33.5%]	81 [34.3%]	79.7 (4.7) [33.8%]	79.1 (5.1) [33.5%]
Â-bridge	10 [4.1%]	5.7 (1.9) [2.4%]	8.4 (2.2) [3.5%]	6 [2.5%]	10.9 (2.8) [4.6%]	9.9 (2.3) [4.2%]
Bend	20 [8.3%]	35.8 (4.2) [14.9%]	36.0 (4.9) [15.0%]	22 [9.3%]	35.0 (4.9) [14.8%]	36.8 (5.2) [15.6%]
Turn	41 [17%]	28.5 (4.6) [11.9%]	30.8 (5.0) [12.8%]	39 [16.5%]	32.1 (4.9) [17.7%]	30.7 (4.5) [13.0%]
Helix	13 [5.4%]	19.6 (2.5) [8.2%]	14.8 (3.1) [6.2%]	15 [6.4%]	12.3 (3.0) [5.2%]	14.4 (2.9) [6.1%]
3_{10}-helix	12 [5.0%]	1.2 (1.8) [0.5%]	2.6 (2.4) [1.1%]	10 [4.2%]	2.8 (2.2) [1.2%]	1.8 (2.0) [0.8%]

Standard deviations and percentage values are given in parenthesis and square brackets, respectively.

Table 4
Average numbers of intramolecular (protein–protein) and protein–solvent H-bonds in PE and SE simulations at 283 and 310 K

	PE 283K	PE 310 K	SE 283 K	SE 310 K
H-bond protein–protein	181.7 (6.6)	185.0 (6.9)	181.7 (6.8)	181.4 (6.7)
H-bond protein–solvent	425.2 (12.0)	391.0 (13.2)	379.7 (11.5)	354.5 (12.1)
H-bond protein (mainchain)–solvent	171.3 (7.7)	155.8 (8.1)	161.7 (9.1)	147.9 (8.7)
H-bond protein (sidechain)–solvent	254.5 (8.8)	235.2 (7.9)	218.4 (7.9)	206.8 (8.1)
Solvent accessible area of polar atoms (nm^2)	45.9	43.2	40.1	40.1
Solvent accessible area of all atoms (nm^2)	105.3	103.3	95.7	95.4
Solvent accessible area of polar atoms/Solvent accessible area of all atoms	0.436	0.418	0.419	0.420
Number of protein polar atoms	677	677	651	651
H-bond protein–protein/number of protein polar atoms	0.268	0.273	0.279	0.279
H-bond protein–solvent/solvent accessible area of all atoms ($1/nm^2$)	4.04	3.79	3.97	3.71
H–bond protein–solvent/solvent accessible area of polar atoms ($1/nm^2$)	9.26	9.05	9.47	8.84

Standard deviations are given in parenthesis. The normalized values are calculated as explained in the text.

solvent accessibility degree of sidechain atoms >45%) with the exception of the residue 77 (Glu in SE and Asn in PE; average solvent accessibility degree of sidechain atoms in the range 10–20%). Eighteen of the amino acids differing in SE and PE, as inferred by sequence alignment (Fig. 2), are localized in regions with rmsf-diff lower than −0.05 or greater than 0.05.

In the 91–99a region, which is characterized by greater flexibility in SE at both temperatures, the aminoacids P92 and Y93 in PE are substituted by S and G in SE, respectively. The molecular environment of the amino acids 92 and 93 is similar in the two proteins and mainly formed by hydrophobic amino-acids (see supplementary materials, Table 2S). The enhanced flexibility of this region might be related to the lack of hydrophobic interactions in the cold-adapted enzyme, in which short sidechain residues (serine and glycine) are present at 92 and 93 positions. Notably, the 91–99a residues are located in a loop region in proximity of the catalytic site and are expected to modulate the access of the substrate to the catalytic sites.

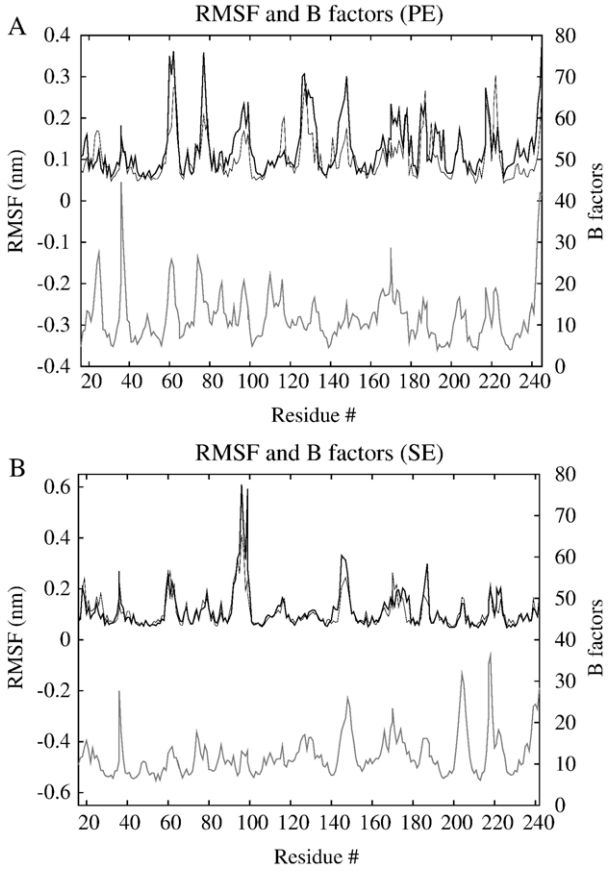

Fig. 4. Rmsf and crystallographic temperature factors as a function of residue numbers. The rmsf values of each residue relative to the average structure of PE (panel A) and SE (panel B) at 283 (grey line) and 310K (black line) are shown, whereas the crystallographic B-factors profiles are shown as grey heavy line.

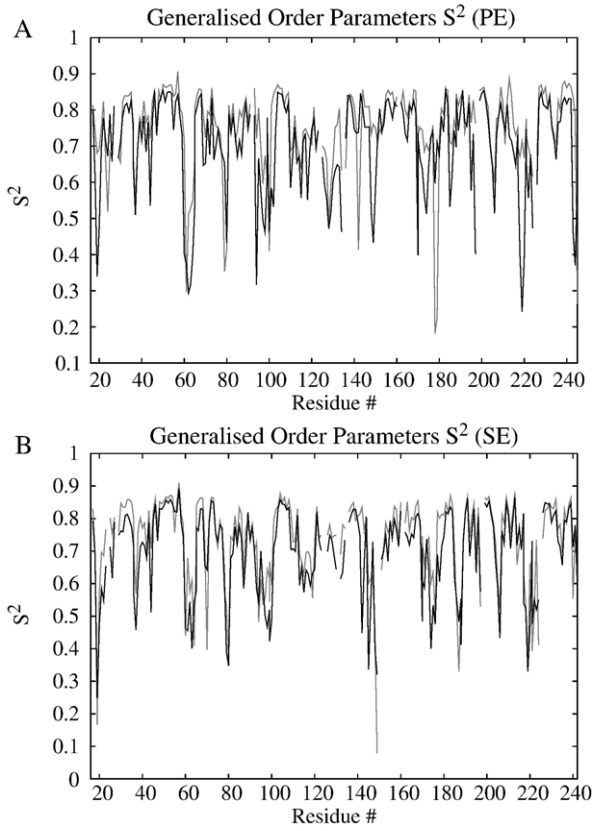

Fig. 5. N-H generalized order parameters S^2. The order parameters as a function of residue numbers for PE (panel A) and SE (panel B) at 283 (grey line) and 310 K (black line) are shown.

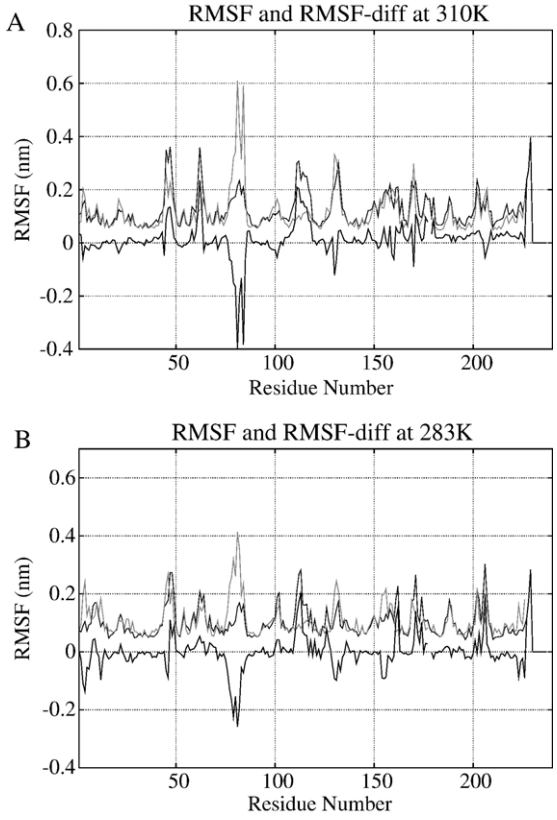

Fig. 6. Rmsf and rmsf-diff as a function of residue numbers. The rmsf values of each residue relative to the average structure of SE (grey line) and PE (black line) at 310 (A) and 283 K (B) are shown as thin lines, whereas the rmsf-diff is shown as a black heavy line.

The protein portion including the residue 145, which differs in PE and SE (R145 in PE is replaced by S in SE), is characterized by greater flexibility in SE at both temperatures. Notably, the analysis of the 3D average structures revealed that the region 143–151, which is localized in proximity of the bottom of the specificity pocket, assumes a β-strand structure in PE simulations, whereas it remains unstructured in SE (see supplementary materials, Table 1S). The amino acid 145, located in proximity of the substrate pocket, presents totally different molecular environments, with S145 surrounded by positively charged residues (K156 and K143) in SE and R145 by polar or hydrophobic residues (Q150, Q156 and L143) in PE (see supplementary materials, Table 2S).

The protein portion 18–21, in which T20 and E21 in PE are replaced by R and V in SE, respectively, is characterized by enhanced fluctuations in SE.

The protein regions of PE that are characterized by enhanced fluctuations when compared to SE include residues: 60–64 (where the SAR residues 61–63 in SE are replaced by REL in PE); 76–77 (where E77 in SE is replaced by N77 in PE); 125–133 (where the PSN residues 125–127 and P131 in SE are substituted by RAG and A in PE, respectively) and 217–222 (where S217a and A221a in SE are replaced by R and V in PE, respectively). These regions are distant both from the catalytic site and the specificity pocket for the substrate: the 60–64 residues are located in the loop between α1 and β4, the residues 76–78 in the loop between β4 and β5; the residues 125–133 in the interdomain loop and the 217a–222 portion in the loop between β10–β11. For some of these regions the different flexibility in PE and SE might be rationalized on the ground of local properties. In fact, the presence of Pro residues at positions 125 and 131 in SE can explain the local increased rigidity in the psychrophilic enzyme (Table 5). In addition, the analysis of the environment of residues 126, 127 and 217a reveals the presence of polar residues (Thr, Ser, Asn and Gln) which form a more extended hydrogen-bond network with S126, N127, S217a in SE than with the mesophilic counterparts A126, G127 and R217a (see supplementary materials, Table 2S).

3.4. Analysis of the catalytic site and the substrate specificity pocket

The molecular environment, as well as the solvent accessibility degree, of the residues composing the catalytic triad (H57, D102, S195) and the specificity pocket for the substrate (S189, T216, V226) has been analyzed (see supplementary materials, Table 3S). The flexibility of the catalytic triad is comparable in the cold- and warm-adapted elastases (Supplementary materials, Table 3S), as well as the solvent accessibility degree, with the exception of H57, which is more solvent exposed in SE. Also the molecular environment of the catalytic residues is quite similar in SE and PE. The differences in the surrounding of H57 and D102 are ascribable to residues characterized by different flexibility in SE and PE. In particular, the flexible W94 residue is less persistent in the surrounding of H57 in SE. N95, S96, G99b and F215 interact with D102 only in SE. In both enzymes, the molecular environment of S195 is composed by the same

Table 5
Protein regions with significantly different rmsf values in SE and PE (rmsf-diff lower and larger than 0 indicate higher SE and PE flexibility, respectively)

Sequence number	Rmsf-diff range (nm)	Amino acid composition in SE	Amino acid composition in PE
18–19–20–21	(−0.05/−0.15)	G–G–R*–V*	G–G–T*–E*
91–92–93–94–95–96–97–98–99–99A	(−0.05/−0.40)	H–S*–G*–W–N–S–D–D–V–A	H–P*–Y*–W–N–T–D–D–V–A
145–147	(−0.05/−0.15)	S*–T	R*–T
60–61–62–63–64	(0.05–0.15)	D–S*–A*–R*–T	D–R*–E*–L*–T
76–77–78	(0.05–0.25)	N–E*–G	N–N*–G
125–126–127–128–129–130–131–132–133	(0.05–0.25)	P*–S*–N*–Q–I–L–P*–N–N	R*–A*–G*–T–I–L–A*–N–N
217A	(0.05–0.15)	S*	R*
219–221–221A–222	(0.05–0.20)	G–C–N–A*–S	G–C–N–V*–T

The amino acid composition is shown using the single letter code and the relevant differences between SE and PE sequences are labeled with an asterisk.

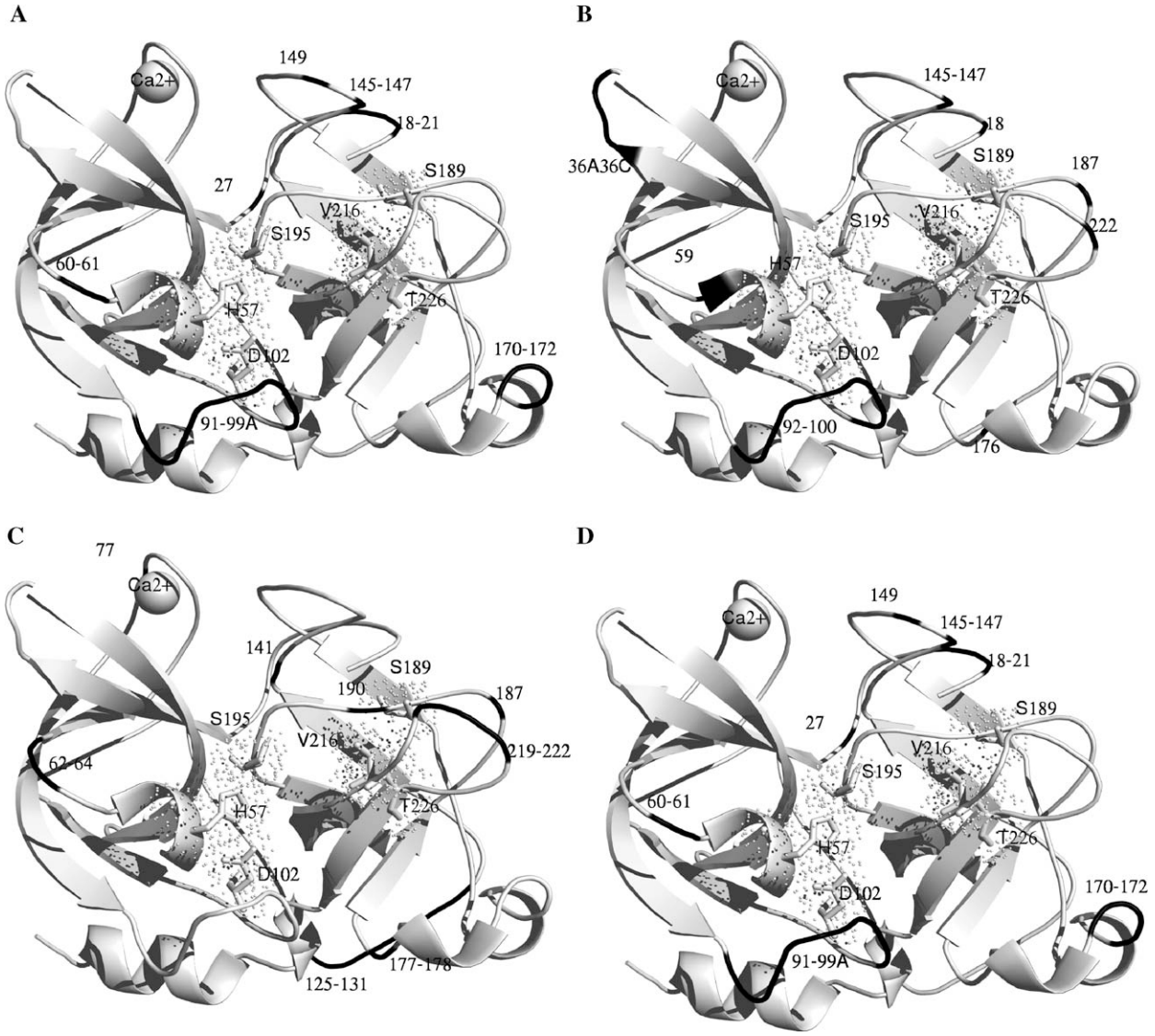

Fig. 7. Regions characterized by different flexibility in SE and PE. The calcium ion and the regions characterized by greater rmsf in SE (A,B) or PE (C,D) at 283 (A,C) and 310 K (B,D) are indicated in grey and black, respectively. The secondary structure elements are shown as ribbons. The catalytic triad (S195, H57 and D102) and the specificity pocket amino acids (V216, T226 and S189) are shown as sticks and dots, whereas the calcium ion is shown as a grey sphere.

residues, which present low fluctuations (rmsf <0.16 nm) and a persistency degree >25%, indicating stable interactions.

The analysis of the molecular environment of the substrate specificity pocket (S189, V216 and T226) reveals that the environment of S189, which is located at the bottom of the pocket, is different in PE and in SE (see supplementary materials, Table 3S): D186 is missing in SE and V188 and V221 are both replaced by alanine residues in SE. Moreover, V216, which is located on one of the two pocket sides, is present only in the surrounding of PE S189, indicating a wider pocket characterized by few long-chain residues in SE, in agreement with the greater solvent accessibility degree of V216 in SE. The surroundings of V216 and T226 are conserved in both proteins, even if a significant difference in the 217a position, where a positively charged residue in PE (R) is replaced by a short-chain polar amino acid in SE (S), is observed.

4. Conclusions

The possibility of modulating protein–solvent interactions, electrostatic potentials, flexibility and solvent accessibility has been exploited during evolution by enzymes from cold-adapted organisms in order to optimize their catalytic activity at low temperature. Indeed, the picture that is emerging from the investigation of psycrophilic enzymes suggests that not all evolutionary strategies are used simultaneously by cold-adapted enzymes. Instead, evidences indicating that different enzyme families have discovered different adaptive strategies are accumulating [2]. However, more experimental and theoretical data are needed to support and refine current hypothesis, as well as to increase our knowledge of the molecular basis of cold adaptation. Therefore, in light of the above scenario, the systematic investigation of different enzyme families becomes crucial

to unravel the cold-adaptation strategies discovered by specific families.

With the aim of contributing to the elucidation of structure–function relations relevant to cold adaptation in the elastase enzymatic family, we have used MD simulation to investigate and compare the psycrophilic Atlantic salmon elastase and the corresponding mesophilic porcine enzyme. The comparative MD investigation reveals that modulation of the number of protein–solvent interactions is not the evolutionary strategy followed by the psycrophilic elastase to enhance catalytic activity at low temperature. In addition, flexibility and solvent accessibility of the residues forming the catalytic triad and the specificity pocket are comparable in the cold- and warm-adapted enzymes. Instead, the loop regions 91–99a and 143–151, which are characterized by a different amino acid composition in the two enzymes, and are clustered around the active site or the specificity pocket, are characterized by enhanced flexibility in the cold-adapted enzyme, leading to the conclusions that these differences play a crucial role for catalysis at low temperature. These latter observations fully support the hypothesis suggesting that flexibility is the main adaptive character of psychrophilic enzymes, being responsible for the decrease of activation enthalpy $\Delta H^{\#}$ that leads to increased k_{cat} values at low temperatures. Remarkably, the mesophilic porcine elastase is characterized by enhanced flexibility, when compared to the Atlantic salmon enzyme, in scattered regions distant from the functional sites. Indeed, it has been noted that enhanced flexibility in regions that are not directly involved in the catalytic motions has no effect on $\Delta H^{\#}$, but on the contrary may decrease the activation entropy $\Delta S^{\#}$, therefore impairing the gain in k_{cat}. Accordingly, if the difference in activation enthalpy between the psychrophilic and the mesophilic homologues arises from local flexibility in regions close to the catalytic centre, the negative effect of $\Delta S^{\#}$ on k_{cat} can be attenuated in psycrophilic enzymes by local rigidity in regions far from the active site [49]. Therefore, our results are also consistent with a scenario in which both local rigidity in regions far from the functional sites and improved flexibility of the structural components involved in the catalysis [29], can be a positive factor in the adaptation of psychrophilic serine-proteases. Moreover, analogous conclusions have been recently reached for cold-adapted uracil-DNA glycosylases (UDG), where the enhanced local flexibility of a loop, implicated in DNA recognition, correlates well with the k_{cat}/K_m values of different mutants of UDG [9].

Our results are also expected to drive mutagenesis studies aimed at highlighting correlation between flexibility indexes and kinetic parameters. In particular, our conclusions can be tested experimentally by site-directed mutagenesis and chimera construction of residues located in the regions 91–99a and 143–151, which we propose to be correlated to the enhanced catalytic efficiency of SE at low temperatures.

Acknowledgements

We thank Marco Pasi, Irene Fumasoni and Rodolfo Gonella Diaza for technical help and fruitful comments. This research was supported by the "Progetto Nazionale Ricerche in Antartide" and by CINECA (Project 562).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbapap.2006.06.005.

References

[1] R. Margesin, G. Feller, C. Gerday, N. Russel, Cold-adapted microorganisms: adaptation strategies and biotechnological potential, in: G. Bitton (Ed.), The Encyclopedia of Environmental Microbiology, John Wiley and Sons, New York, 2002, pp. 871–885.
[2] D. Georlette, V. Blaise, T. Collins, S. D'Amico, E. Gratia, A. Hoyoux, J.C. Marx, G. Sonan, G. Feller, C. Gerday, Some like it cold: biocatalysis at low temperatures, FEMS Microbiol. Rev. 28 (2004) 25–42.
[3] A.O. Smalas, H.K. Leiros, V. Os, N.P. Willassen, Cold adapted enzymes, Biotechnol. Annu. Rev. 6 (2000) 1–53.
[4] C. Gerday, M. Aittaleb, M. Bentahir, J.P. Chessa, P. Claverie, T. Collins, S. D'Amico, J. Dumont, G. Garsoux, D. Georlette, A. Hoyoux, T. Lonhienne, M.A. Meuwis, G. Feller, Cold-adapted enzymes: from fundamentals to biotechnology, Trends Biotechnol. 18 (2000) 103–107.
[5] R. Cavicchioli, K.S. Siddiqui, D. Andrews, K.R. Sowers, Low-temperature extremophiles and their applications, Curr. Opin. Biotechnol. 13 (2002) 253–261.
[6] P.A. Fields, Protein function at thermal extremes: balancing stability and flexibility, Comp. Biochem. Physiol., A. Mol. Integr. Physiol. 129 (2001) 417–431.
[7] P.A. Fields, G.N. Somero, Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 11476–11481.
[8] K. Miyazaki, P.L. Wintrode, R.A. Grayling, D.N. Rubingh, F.H. Arnold, Directed evolution study of temperature adaptation in a psychrophilic enzyme, J. Mol. Biol. 297 (2000) 1015–1026.
[9] M. Olufsen, A.O. Smalas, E. Moe, B.O. Brandsdal, Increased flexibility as a strategy for cold adaptation. A comparative molecular dynamics study of cold- and warm-active uracil DNA glycosylase, J. Biol. Chem. 280 (2005) 18042–18048.
[10] N. Aghajari, F. Van Petegem, V. Villeret, J.P. Chessa, C. Gerday, R. Hasen, J. Van Beeuman, Crystal structures of a psychrophilic metalloprotease reveal new insights into catalysis by cold-adapted proteases, Proteins 50 (2003) 636–647.
[11] F. Van Petegem, T. Collins, M.A. Meuwis, C. Gerday, G. Feller, J. Van Beeuman, The structure of a cold-adapted family 8 xylanase at 1.3 Å resolution, structural adaptations to cold and investigation of the active site, J. Biol. Chem. 278 (2003) 7531–7539.
[12] M. de Backer, S. McSweeney, H.B. Rasmussen, B.W. Riise, P. Lindley, E. Hough, The 1.9 Å crystal structure of heat-labile shrimp alkaline phosphatase, J. Mol. Biol. 318 (2002) 1265–1274.
[13] A.O. Smalas, E.S. Heimstad, A. Hordvik, N.P. Willassen, R. Male, Cold adaptation of enzymes: structural comparison between salmon and bovine trypsins, Proteins 20 (1994) 149–166.
[14] G.I. Berglund, N.P. Willassen, A. Horvik, A.O. Smalas, Structure of native pancreatic elastase from North Atlantic salmon at 1.61 Å resolution, Acta Crystallogr., D Biol. Crystallogr. D 51 (1995) 925–937.
[15] I. Leiros, O. Lanes, O. Sundheim, R. Helland, A.O. Smalas, N.P. Willassen, Crystallization and preliminary X-ray diffraction analysis of a cold-adapted uracil-DNA glycosylase from Atlantic cod (*Gadus morhua*), Acta Crystallogr., Sect. D: Biol. Crystallogr. 57 (2001) 1706–1708.
[16] B. Asgeirsson, B.N. Nielsen, P. Hojrup, Amino acid sequence of cold-active alkaline phosphatase from Altantic cod (*Gadus morhua*), Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol. 136 (2003) 45–60.
[17] D. Georlette, Z.O. Jonsson, F. Van Petegem, J. Chessa, J. Van Beeumen, U. Hubscher, C. Gerday, A DNA ligase from the

psychrophile *Pseudoalteromonas haloplanktis* gives insights into the adaptation of proteins to low temperatures, Eur. J. Biochem. 267 (2000) 3502–3512.
[18] H.K. Leiros, N.P. Willassen, A.O. Smalas, Structural comparison of psychrophilic and mesophilic trypsins. Elucidating the molecular basis of cold-adaptation, Eur. J. Biochem. 267 (2000) 1039–1049.
[19] S. Kumar, R. Nussinov, Different roles of electrostatics in heat and in cold: adaptation by citrate synthase, Chembiochem 5 (2004) 280–290.
[20] R.J. Russell, U. Gerike, M.J. Danson, D.W. Hough, G.L. Taylor, Structural adaptations of the cold-active citrate synthase from an Antarctic bacterium, Structure 6 (1998) 351–361.
[21] G. Gianese, F. Bossa, S. Pascarella, Comparative structural analysis of psychrophilic and meso- and thermophilic enzymes, Proteins 47 (2002) 236–249.
[22] G. Gianese, P. Argos, S. Pascarella, Structural adaptation of enzymes to low temperatures, Protein Eng. 14 (2001) 141–148.
[23] E. Papaleo, P. Fantucci, L. De Gioia, Effects of calcium binding on structure and autolysis regulation in trypsins. A molecular dynamics investigation, J. Chem. Theory and Comput. 1 (2005) 1286–1297.
[24] C.A. Sotriffer, O. Kramer, G. Klebe, Probing flexibility and "induced-fit" phenomena in aldose reductase by comparative crystal structure analysis and molecular dynamics simulations, Proteins 56 (2004) 52–66.
[25] Y. Pan, B. Ma, O. Keskin, R. Nussinov, Characterization of the conformational state and flexibility of HIV-1 glycoprotein gp120 core domain, J. Biol. Chem. 279 (2004) 30523–30530.
[26] J. Norberg, L. Nilsson, Advances in biomolecular simulations: methodology and recent applications, Q. Rev. Biophys. 36 (2003) 257–306.
[27] W. Wang, O. Donini, C.M. Reyes, P.A. Kollman, Biomolecular simulations: recent developments in force fields, simulations of enzyme-catalysis, protein–ligand, protein–protein and protein–nucleic acid noncovalent interactions, Annu. Rev. Biophys. Biomol. Struct. 30 (2001) 211–243.
[28] T. Hansson, C. Oostenbrink, W. van Gusteren, Molecular dynamics simulations, Curr. Opin. Struct. Biol. 12 (2002) 190–196.
[29] B.O. Brandsdal, E.S. Heimstad, I. Sylte, A.O. Smalas, Comparative molecular dynamics of mesophilic and psycrophilic protein homologues studied by 1.2 ns simulations, J. Biomol. Struct. Dyn. 17 (1999) 493–506.
[30] G.I. Berglund, A.O. Smalas, H. Outzen, N.P. Willassen, Purification and characterization of pancreatic elastase from North Atlantic salmon (*Salmo salar*), Mol. Mar. Biol. Biotechnol. 7 (1998) 105–114.
[31] D. van der Spoel, A.R. van Buuren, E. Apol, P.J. Meulenhoff, D.P. Tieleman, A.L.T.M. Sijbers, B. Hess, K.A. Feenstra, E. Lindhal, R. van Drunen, H.J.C. Berendsen, *Gromacs User Manual version 3.1.1*, Internet: www.gromacs.org, (2002).
[32] M. Schiltz, W. Shepard, R. Fourme, T. Prange, E. de la Fortelle, G. Bricogne, High-pressure krypton gas and statistical heavy-atom refinement: a successful combination of tools for macromolecular structure determination, Acta Crystallogr., Sect. D: Biol. Crystallogr. 53 (1997) 78–92.
[33] G.I. Berglund, N.P. Willassen, A. Hordvik, A.O. Smalas, Structure of native pancreatic elastase from North Atlantic salmon at 1.61 Å resolution, Acta Crystallogr., Sect. D: Biol. Crystallogr. 51 (1995) 925–937.
[34] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gusteren, J. Hermans, Interaction Models for Water in Relation to Protein Hydratation, Reidel, Dordrecht, 1981, pp. 331–342.
[35] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gusteren, A. Di Nola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[36] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraahje, LINCS: a linear constraint solver for molecular interactions, J. Comp. Chem. 18 (1997) 1463–1472.
[37] T.A. Darden, D.M. York, L.G. Pedersen, Particle mesh Ewald. An N.log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[38] R.A. Jarvis, E.A. Patrick, Clustering using a similarity measure based on shared nearest neighbours, IEE Trans. Comp. 11 (1973) 1025–1034.
[39] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[40] S.J. Hubbard, J.M. Thornton, NACCESS, Comp. Program, (1993).
[41] G. Lipari, A. Szabo, Model-free approach to the interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules: 1. Theory and Range of Validity, J. Am. Chem. Soc. 104 (1982) 4546–4559.
[42] D.C. Chatfield, A. Szabo, B.R. Brooks, Molecular dynamics of Staphylococcal nuclease: comparison simulation with 15N and 13C NMR Relaxation Data, J. Am. Chem. Soc. 120 (1998) 5301–5311.
[43] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.
[44] J. Tsai, M. Levitt, D. Baker, Hierarchy of structure loss in MD simulations of src SH3 domain unfolding, J. Mol. Biol. 291 (1999) 215–225.
[45] B. Hess, Convergence of sampling in protein simulations, Phys. Rev. E 652 (2002) 031910/1-031910/10.
[46] G. Caracciolo, G. Amiconi, L. Bencivenni, G. Boumis, R. Caminiti, E. Finocchiaro, B. Maras, C. Paolinelli, A. Congiu Castellano, Conformational study of proteins by SAXS and EDXD: the case of trypsin and trypsinogen, Eur. Biophys. J. 30 (2001) 163–170.
[47] P.H. Hunenberger, A.E. Mark, W.F. van Gusteren, Fluctuations and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations, J. Mol. Biol. 252 (1995) 492–503.
[48] P. Eastman, M. Pellegrini, S. Doniach, Protein flexibility in solution and in crystals, J. Chem. Phys 110 (1999) 10142–10151.
[49] T. Lonhienne, C. Gerday, G. Feller, Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility, Bioch. Biophys. Acta 1543 (2000) 1–10.
[50] V. Brendel, P. Bucher, I. Nourbakhsh, B.E. Blaisdell, S. Karlin, Methods and algorithms for statistical analysis of protein sequences, Proc. Natl. Acad. Sci. U. S. A. 89 (1992) 2002–2006.
[51] P. McCaldon, P. Argos, Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences, Proteins 4 (1988) 99–122.