Biophysical Chemistry 153 (2010) 104–114

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

"Cold spots" in protein cold adaptation: Insights from normalized atomic displacement parameters (B'-factors)

Alessandro Siglioccolo [a], Raffaele Gerace [a], Stefano Pascarella [a,b,*]

[a] Dipartimento di Scienze Biochimiche "A. Rossi Fanelli", Università di Roma La Sapienza, Italy
[b] Centro di Ricerca per l'Analisi dei Modelli e dell'Informazione nei Sistemi Biomedici (CISB), Università di Roma La Sapienza, Italy

ARTICLE INFO

Article history:
Received 26 August 2010
Received in revised form 13 October 2010
Accepted 13 October 2010
Available online 21 October 2010

Keywords:
B-factor
Local flexibility
Cold adaptation
Psychrophiles
Catalysis

ABSTRACT

Many analyses published in the last decade suggest that enzymes isolated from cold-adapted organisms are characterized by a higher flexibility of their molecular structure. Recently, it has been argued that all cold-adapted enzymes with catalytic efficiency greater than that of their mesophilic counterparts display local flexibility or rigidity that are likely to cooperate, each acting on specific areas of the enzyme structure. Here we report an analysis of the normalized thermal *B*-factor distributions in psychrophilic proteins compared with those of their mesophilic and thermophilic counterparts with the aim to detect statistically significant local variations of relative backbone flexibility possibly linked to cold adaptation. We utilized a strategy based mainly on intra-family comparison of local distribution of normalized *B*-factors. After careful statistical treatment of data, the picture emerging from our results suggests that the distribution of the flexibility in psychrophilic enzymes is locally more heterogeneous than in their respective mesophilic homologues.

1. Introduction

Many terrestrial environments present physical and chemical conditions that are considered extreme from an anthropocentric perspective. The organisms able to thrive in such extreme environments are called extremophiles or polyextremophiles when the environment presents combination of different extreme conditions. Colonization of these environments by life required a large variety of adaptive strategies. The extremophilic organisms adopted several strategies to survive at such conditions involving physiological modifications of the intracellular environment and the synthesis of biomolecules displaying adequate properties. The most widespread extreme condition for life is represented by low-temperature environments. Indeed, the biosphere is dominated by cold environments such as polar regions, mountains, and oceans where temperatures at depths below 1000 m do not exceed 5 °C. Organisms adapted at such extreme environments are named psychrophiles. To survive, they need enzymes able to efficiently catalyze viable reactions at temperatures close to 0 °C, at which most of the other species cannot grow because of their inability to maintain adequate metabolic fluxes [1–3]. Compared with their mesophilic and thermophilic counterparts, most of these enzymes display improved catalytic efficiency at low temperatures that is reflected in their higher turnover number (rise in k_{cat}) at the expense of affinity for the substrate the effect of which is increase of K_M [4]. Several works, which aimed at finding common structural determinants for cold adaptation, have been published [3,5–8]. Because of their higher catalytic efficiency at low temperatures, enzymes extracted from psychrophilic organisms have been attracting scientists' attention for their biotechnological potential, in particular in industrial processes as energy savers and in detergent industry as additives [9–12].

It is now generally accepted that enzymes isolated from cold-adapted organisms are characterized by a higher plasticity or flexibility of their molecular structure to compensate for the lower thermal energy provided by the low-temperature habitat. High structural flexibility of the psychrophilic enzymes could allow better interaction with substrates and could explain their higher catalytic rate (k_{cat}), lower thermostability, and lower activation energy (E_a) requirements, compared with requirements for mesophilic and thermophilic counterparts. More recently, it is emerging that all cold-adapted enzymes with catalytic efficiency greater than that of their mesophilic counterparts display local flexibility/rigidity [13]. Flexibility and rigidity are likely to cooperate, each acting on specific areas of the enzyme structure. Although flexibility and rigidity are properties difficult to quantify for a small and anisotropic material such as a protein molecule, the comparative analysis of the distribution of *B*-factors of Cα in psychrophilic and their homologous structures can give valuable indications on these properties. Cα *B*-factors are indeed considered an important source of information about protein internal dynamics, providing a map of the static flexibility of the ground-state protein conformation [14]. *B*- (or temperature) factors are values that can be applied to the X-ray scattering term for each atom and describe the degree to which the electron density is spread out. This effect can

* Corresponding author. Dipartimento di Scienze Biochimiche, Sapienza Università di Roma, P.le A. Moro 5, 00185 Rome, Italy. Tel.: +39 06 49917694; fax: +39 06 49917566.
E-mail address: Stefano.Pascarella@uniroma1.it (S. Pascarella).
URL: http://w3.uniroma1.it/bio_chem/sito_biochimica/index.html (S. Pascarella).

0301-4622/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2010.10.009

be caused by the vibration of the atoms, or differences between the many different molecules in the crystal lattice. The observed electron density will include an average of all these small motions, yielding a slightly smeared image of the molecule. These motions, and the resultant smearing of the electron density, are incorporated into the atomic model by a *B*-factor. The magnitude of the *B*-value is proportional to the amount of smearing according to the definition $B = 8\pi^2 \times U^2$ where U^2 is the atomic mean displacement. Therefore, these factors can indicate the true static or dynamic mobility of an atom. Atomic *B*-factors are included in the files from Protein Data Bank (PDB) [19] containing the coordinates of a X-ray solved protein structures.

For example, analysis of the distribution of normalized *B*-factors was utilized to characterize the differences in dynamics of thermophilic and mesophilic proteins [15]. Moreover, it was demonstrated that Cα protein backbone flexibility measured by *B*-factor profiles are conserved at family and superfamily levels, even for pairs of proteins with nonsignificant sequence similarity [16]. Correlation of *B*-factor profiles of homologous proteins show that they are correlated with each other with average correlation coefficient equal to 0.80 [17]. Moreover *B*-factors were utilized in the prediction of folding rate of a protein [18]. These findings strongly support the idea that enzyme flexibility and its localization on the structure plays an essential functional role in enzymatic catalysis and protein function in general. Here we report an analysis of the normalized *B*-factor distributions in psychrophilic proteins compared with those of their mesophilic and thermophilic counterparts with the aim to detect significant local variations of relative backbone flexibility possibly linked to cold adaptation.

2. Materials and methods

2.1. Selection of protein structures

The crystallographic structures of the available cold active enzymes were found in the PDB [19]. The search was carried out with the keywords: “psychro”, “cold”, “arctic”, “antarctic” and the like. Only psychrophilic enzyme structures, for which high cold activity and low thermostability were reported either in BRENDA databank [20] or in literature, were considered. Homologous structures from mesophilic and thermophilic organisms were retrieved from PDB by means of the program BLAST [21]. To ensure structural homology, only sequences sharing ≥30% residue identity to the psychrophilic counterpart were considered. Moreover, only structures resolved at least at 2.8 Å resolution were taken into account. Only unique structures were retrieved, and in the presence of alternative structures for the same protein, only those displaying the best resolution and without point mutations, were collected. Growth temperature of the microorganism sources of the selected proteins were taken from the databank DSMZ (http://www.dsmz.de/, Deutsche Sammlung von Mikroorganismen und Zellkulturen).

2.2. B-factor analyses

The *B*-values at the Cα atoms of each selected protein chain were replaced by normalized *B*-factors according to the definition utilized in literature [15]:

$$B' = \frac{B - \langle B \rangle}{\sigma(B)} \tag{1}$$

where $\langle B \rangle$ is the mean *B*-value at Cα atoms and $\sigma(B)$ is the standard deviation. Frequency distributions of *B*′-factors were calculated within different structural environments. Frequencies were counted in 0.5 or 1.0 unit ranges in the interval $(-4.0) - (+7.0)$ $\sigma(B)$. Each residue was assigned to a structural environment according either to its secondary structure (α-helix, β-sheet, β-turn or coil) or its relative solvent accessibility (if ≤5% it was considered buried, if ≥16% exposed) [22]. Secondary structures were assigned with the program DSSP [23] while residue relative accessibility was calculated with the program NACCESS [24]. Frequency of residues within each structural environment and 0.5 σ unit range was defined as:

$$f_{ik} = \frac{n_{ik}}{n_k} \tag{2}$$

where n_{ik} is the number of residue Cα in the *i-th* σ unit range (for example $1.0 \le B' < 1.5$) in the structural environment *k* (for example, α-helix), n_k is the total number of Cα in the structural environment *k* and f_{ik} is the resulting frequency.

Differences in the frequency of Cα in the different *B*′-factor intervals *i* and structural environments *k* were calculated as:

$$\Delta f_{ik} = f_{ik}^{\psi} - f_{ik}^{h} \tag{3}$$

where f_{ik}^{ψ} is the frequency of residues falling in the *i-th* σ interval of the *B*′-factor distribution (for example 1.0–1.5 σ) in the structural environment *k* (for example, α-helix) in the psychrophilic protein and f_{ik}^{h} the corresponding frequency in each homologous protein. To explore links of Δf_{ik} to sequence or molecular phenotype, Pearson correlation coefficients were calculated between the populations of Δf_{ik} and the corresponding pairwise percentage of identity or difference in optimal temperature.

Average Cα *B*′-factor in every structural environment or in each residue type of a protein was calculated as

$$\overline{B'_r} = \frac{\sum B'_r}{n_r} \tag{4}$$

where B'_r is the *B*′-factor of the residue of type *r* and n_r is the total number of type *r* residues found in the protein considered. The average B'_r-factor was also calculated over all the psychrophilic enzymes and compared with the corresponding values found for the mesophilic and hyper/thermophilic homologs. The average *B*′-factor in a given structural environment was calculated according to the Eq. (4) where the index *r* is replaced by *s*, *i.e.* the structural environment.

Intra-family differences of the average *B*′-factor value between the psychrophilic and each of the corresponding homologous mesophilic and hyper/thermophilic enzymes were calculated according to:

$$\Delta B'_k = \overline{B'}^{\psi}_k - \overline{B'}^{h}_k \tag{5}$$

where *k* means alternatively *r* (residue type) or *s* (structural environment) and $\overline{B'}^{\psi}_k$ and $\overline{B'}^{h}_k$ indicate the average *B*′-factor in the psychrophilic enzymes and in their meso- and hyper/thermophilic homologs, respectively. To explore links of $\Delta B'_k$ to sequence or molecular phenotype, Pearson correlation coefficients were calculated between the population of $\Delta B'_k$ and the corresponding pairwise percentage of identity or difference in optimal temperature.

2.3. Intra-family local B′-factor comparison

The distribution of average Cα *B*′-factors along the sequence of each psychrophilic protein has been compared with that of the corresponding homologous proteins. A structural alignment has been calculated by means of the program SSM [25] within each family and the structural equivalencies were derived. Afterward, a 5-residue sliding window [15] has been applied to each sequence within each family and a local average Cα *B*′-factor calculated as:

$$\overline{B'_i} = \frac{1}{2N+1} \sum_{n=i-N}^{i+N} B'_n \tag{6}$$

Table 1
Protein families collected for the analysis.

Family	Source[a]	Growth T (°C)[b]	Pdb ID[c]	Res. (Å)[d]	Seq. identity(%)[e]	Quaternary structure[f]	Seq. length[g]
Citrate synthase	***Antarctic bacterium***	**5**	**1A59**	**2.09**	**–**	**2**	**377**
	Escherichia coli	37	1K3P	2.20	122/392 (31%)	2	426
	Sulfolobus solfataricus	80	1O7X	2.70	125/382 (32%)	2	379
	Thermotoga maritima	80	2P2W	1.90	151/370 (40%)	2	367
	Thermus thermophilus	85	1IOM	1.50	154/375 (41%)	2	377
	Pyrobaculum aerophilum	98	2IBP	1.60	143/393 (36%)	2	409
	Pyrococcus furiosus	100	1AJ8	1.90	145/370 (39%)	2	371
α-Amylase	***Alteromonas haloplanktis***	**4**	**1AQM**	**1.85**	**–**	**1**	**448**
	Tenebrio molitor	25	1JAE	1.65	199/473 (42%)	1	471
	Sus scrofa	37	1DHK	1.85	230/489 (47%)	4	496
	Homo sapiens	37	1HNY	1.80	219/489 (44%)	1	496
	Homo sapiens	37	1SMD	1.60	221/490 (45%)	1	496
Triosephosphate isomerase	***Moritella marina***	**15**	**1AW2**	**2.65**	**–**	**2**	**255**
	Saccharomyces cerevisiae	27	1YPI	1.90	104/239 (43%)	2	248
	Gallus gallus	37	1TPH	1.80	101/240 (42%)	2	245
	Oryctolagus cuniculus	37	1R2R	1.50	99/239 (41%)	2	248
	Trypanosoma cruzi	37	1TCD	1.83	99/239 (41%)	2	248
	Caenorhabditis elegans	22	1MO0	1.70	113/239 (47%)	2	249
	Escherichia coli	37	1TRE	2.60	166/255 (65%)	2	255
	Entamoeba histolytica	37	1M6J	1.50	107/253 (42%)	2	261
	Plasmodium falciparum	37	1YDV	2.20	96/251 (38%)	2	246
	Leishmania mexicana	37	1AMK	1.83	95/238 (39%)	2	250
	Trypanosoma brucei	41	1TPF	1.80	100/251 (39%)	2	250
	Tenebrio molitor	25	2I9E	2.00	109/244 (44%)	2	259
	Homo sapiens	37	1WYI	2.20	99/239 (41%)	2	250
	Geobacillus stearothermophilus	55	2BTM	2.40	108/251 (43%)	2	250
	Thermotoga maritima	85	1B9B	2.85	103/249 (41%)	2	252
Malate dehydrogenase	***Aquaspirillum arcticum***	**4**	**1B8P**	**1.90**	**–**	**2**	**327**
	Sus scrofa	37	5MDH	2.40	166/330 (50%)	2	333
	Burkholderia pseudomallei	40	3D5T	2.51	253/328 (77%)	2	331
	Thermus flavus	72	1BMD	1.90	202/326 (61%)	2	327
	Thermus thermophilus	85	1IZ9	2.00	202/326 (61%)	2	327
Elastase	***Salmo salar***	**4**	**1ELT**	**1.61**	**–**	**1**	**236**
	Sus scrofa	37	1QNJ	1.10	161/237 (67%)	1	240
Alkaline phosphatase	***Pandalus borealis***	**5**	**1K7H**	**1.92**	**–**	**2**	**476**
	Homo sapiens	37	1EW2	1.82	207/484 (42%)	2	479
	Escherichia coli	37	1ED9	1.75	124/376 (32%)	2	449
Subtilisin	***Bacillus antarctic TA41***	**5**	**2GKO**	**1.40**	**–**	**1**	**309**
	Bacillus clausii	30	1WSD	1.50	120/285 (42%)	1	269
	Bacillus lentus	26	1GCI	0.78	120/285 (42%)	1	269
	Bacillus sp.	30	1DBI	1.80	105/287 (36%)	1	280
	Bacillus subtilis	30	1SEL	2.00	115/283 (40%)	1	274
	Bacillus amyloliquefaciens	30	1GNS	1.80	100/275 (36%)	1	263
	Thermoactinomyces vulgaris	45	1THM	1.37	103/268 (38%)	1	279
	Bacillus sphaericus	50	2IXT	0.80	225/309 (72%)	1	310
	Thermococcus kodakarensis	85	2Z2X	1.70	101/280 (36%)	1	318
Isocitrate dehydrogenase	***Desulfotalea psychrophila***	**6**	**2UXQ**	**1.75**	**–**	**2**	**402**
	Sus scrofa	37	1LWD	1.85	203/405 (50%)	2	413
	Mus musculus	37	2CMJ	1.99	204/398 (51%)	2	410
	Homo sapiens	37	1T09	2.70	203/398 (51%)	2	414
	Thermotoga maritima	80	1ZOR	2.24	232/399 (58%)	2	399
β-Lattamase	***Pseudomonas fluorescens***	**18**	**2QZ6**	**2.26**	**–**	**1**	**358**
	Escherichia coli	37	2BLS	2.00	178/354 (50%)	1	358
	Klebsiella pneumoniae	30	2ZC7	2.40	169/355 (47%)	1	359
	Enterobacter cloacae	30	1GCE	1.80	167/356 (46%)	1	364
Phenylalanine-hydroxylase	***Colwellia psychrerythraea 34H***	**10**	**2V27**	**1.50**	**–**	**2**	**275**
	Chromobacterium violaceum	25	1LTU	1.74	93/216 (43%)	1	297
	Rattus norvegicus	37	1PHZ	2.20	77/227 (33%)	2	429
	Homo sapiens	37	2PAH	3.10	76/227 (33%)	4	335
L-Lactate dehydrogenase	***Champsocephalus gunnari***	**0**	**2V65**	**2.35**	**–**	**4**	**331**
	Cyprinus carpio	25	1V6A	2.30	282/331 (85%)	4	332
	Thermus thermophilus	85	2V6M	2.20	117/301 (38%)	4	310
	Thermotoga maritima	80	1A5Z	2.10	122/309 (39%)	4	319
	Aeropyrum pernix	90	2D4A	2.87	95/303 (31%)	4	308
Aminopeptidase	***Colwellia psychrerythraea***	**10**	**3CIA**	**2.70**	**–**	**1**	**605**
	Homo sapiens	37	1H19	2.10	220/603 (36%)	1	611
D-Gliceraldehyde-3-phosphate dehydrogenase	***Homarus americanus***	**20**	**4GPD**	**2.80**	**–**	**4**	**333**
	Leishmania mexicana	37	1A7K	2.80	194/353 (54%)	4	358
	Oryctolagus cuniculus	37	1J0X	2.40	241/327 (73%)	4	332
	Achromobacter xylosoxidans	30	1OBF	1.70	149/327 (45%)	4	335
	Trypanosoma cruzi	37	1QXS	2.75	194/353 (54%)	4	359
	Homo sapiens	37	1U8F	1.75	236/327 (72%)	4	335
	Kluyveromyces marxianus	25	2I5P	2.30	217/326 (66%)	2	342
	Plasmodium falciparum	37	1YWG	2.60	202/331 (61%)	4	337
	Cryptosporidium parvum	30	3CPS	1.90	204/326 (62%)	4	354

Table 1 (*continued*)

Family	Source[a]	Growth T (°C)[b]	Pdb ID[c]	Res. (Å)[d]	Seq. identity(%)[e]	Quaternary structure[f]	Seq. length[g]
	Brucella abortus	37	3DOC	2.40	151/325 (46%)	4	335
	Geobacillus stearothermophilus	55	1GD1	1.80	177/325 (54%)	4	334
	Thermus aquaticus	70	2G82	1.65	161/326 (49%)	4	331
	Thermotoga maritima	85	1HDG	2.50	163/334 (48%)	4	332
	Thermus thermophilus	85	1VC2	2.60	161/330 (48%)	4	331
Trypsin	***Salmo salar***	**4**	**1HJ8**	**1.00**	**–**	**1**	**222**
	Oncorhynchus keta	20	1MBQ	1.80	217/220 (98%)	1	220
	Rattus rattus	37	1DPO	1.59	148/222 (66%)	1	223
	Homo sapiens	37	1H4W	1.70	142/222 (63%)	1	224
	Sus scrofa	37	1FNI	1.60	147/222 (66%)	1	223
	Bos taurus	37	1S0Q	1.02	147/222 (66%)	1	223
Proteinase-K like	***Serratia sp***	**4**	**2B6N**	**1.80**	**–**	**2**	**278**
	Vibrio sp. PA-44	30	1SH7	1.84	186/276 (67%)	1	284
	Engyodontium album	25	1IC6	0.98	118/280 (42%)	1	279
	Bacillus lentus	26	1NDQ	1.80	112/257 (43%)	1	269

[a] Boldfaced names and numbers refer to the psychrophilic member of the family.
[b] Growth temperature.
[c] Protein Data Bank code.
[d] Crystallographic resolution.
[e] Number of identical versus aligned residue shared with the psychrophilic homologue. In parentheses: corresponding percentage of identity.
[f] Number of subunits in the biological unit.
[g] Monomer sequence length.

where N is equal to 2, and i is the alignment position of the central residue. For each window, the set of differences $\Delta\overline{B}_i$ between $\overline{B_i'}$ of the psychrophilic sequence and the corresponding one of each homologous proteins, as well as the inverse differences between each mesophilic $\overline{B_i'}$ and the corresponding psychrophilic, were calculated. This population represents the first distribution that, by definition, has mean equal to 0. The reference population with mean equal to 0 consists of the differences $\Delta\overline{B}_i$ between $\overline{B_i'}$ of all the possible mesophilic windows at position i as defined in Eq. (6). If the differences between the psychrophilic and the mesophilic B'-factors are greater than those within the mesophilic population, the variances of the two populations will be different and the variance of the first population would be greater than that of second population. The opposite if the differences were smaller. To evaluate which window $\overline{B_i'}$ is significantly different between the psychrophilic and the mesophilic proteins, the estimated variances of the window populations were tested to be different against the null hypothesis that the population variances were identical. For this purpose a Bartlett's test [26] was applied, with $k=2$ samples to test, with size n_i and sample variance S_i^2:

$$X^2 = \frac{(N-k)\ ln\left(S_p^2\right) - \sum_{i=1}^{k}(n_i-1)\ ln\left(S_i^2\right)}{1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\left(\frac{1}{n_i-1}\right) - \frac{1}{N-k}\right)} \quad (7)$$

where $N = \sum_{i=1}^{k} n_i$ and $S_p^2 = \frac{1}{N-k}\sum_i (n_i-1)S_i^2$ is the pooled estimate for the variance. The test returns a p-value which measures the probability of observing a value greater than X^2 under the null hypothesis, *i.e.*, that variances are equal. Significance threshold was set at p-value = 0.05.

2.4. Software

Calculations were performed with Microsoft Office Excel 2007 or with Perl or Python scripts developed *ad hoc* under Ubuntu Linux. Statistical analyses based on *t*-test or correlation were carried out with the software OriginPro 8.0 (OriginLab corporation, USA, 2009). Molecular graphics relied on the program Pymol [27].

3. Results

3.1. Data

A set of 15 psychrophilic enzymes was collected along with their structural homologues (Table 1) for a total of 90 proteins. Among the proteins homologous to each psychrophilic enzyme are represented mesophilic and, in seven families, hyper/thermophilic species (Table 1). All redundant structures (percentage identity greater than 98%) were removed while retaining the protein at the highest resolution. All the collected structure have a resolution better than 2.8 Å except in the case of 2PHA (human phenylalanine hydroxylase) available only at 3.1 Å resolution. The calculation of accessibility was carried out on the apoenzymes in their quaternary structure.

3.2. B'-factor distribution

The B-factors were normalized to B' as described in Materials and methods and all subsequent analyses were carried out in this form. General distribution of Cα B'-factors in psychrophilic, mesophilic and hyper/thermophilic enzymes in 0.5 σ unit intervals were calculated (Fig. 1). The distribution shape of the three temperature groups display an excess of high B'-factor around the value $+2.0\ \sigma$ and a decreased frequency in the region $-2.0\ \sigma$ in the psychrophilic proteins. To understand the origin of these variations, further distributions were calculated separately for the structural environments: α-helix, β-strand, coil, and residues at relative solvent accessibility (rsa) $\leq$5% or $\geq$16% for buried and exposed positions, respectively. The distributions taken for α-helix, β-strand and coil displayed an enrichment in the region around $+2.0\,\sigma$ and a depletion in the region $-2.0\ \sigma$ of the psychrophilic proteins (Fig. 1) while no apparent difference were detected in the same range of the other structural environments. To assess the statistical significance of the observed variations, differences in the relative frequencies of Cα B'-factors $\geq 1.0\,\sigma$ or, alternatively, $\geq 1.5\,\sigma$ between psychrophiles and each of their homologues were calculated. In this case mesophilic and thermophilic proteins were combined to enhance statistics. The significant values were identified through a one-sample *t-test* using the null hypothesis that the average difference were not significantly greater or less than 0 at 0.05 p-value. If the difference were positive, the frequency of "flexible" B'-factor atoms would be higher in

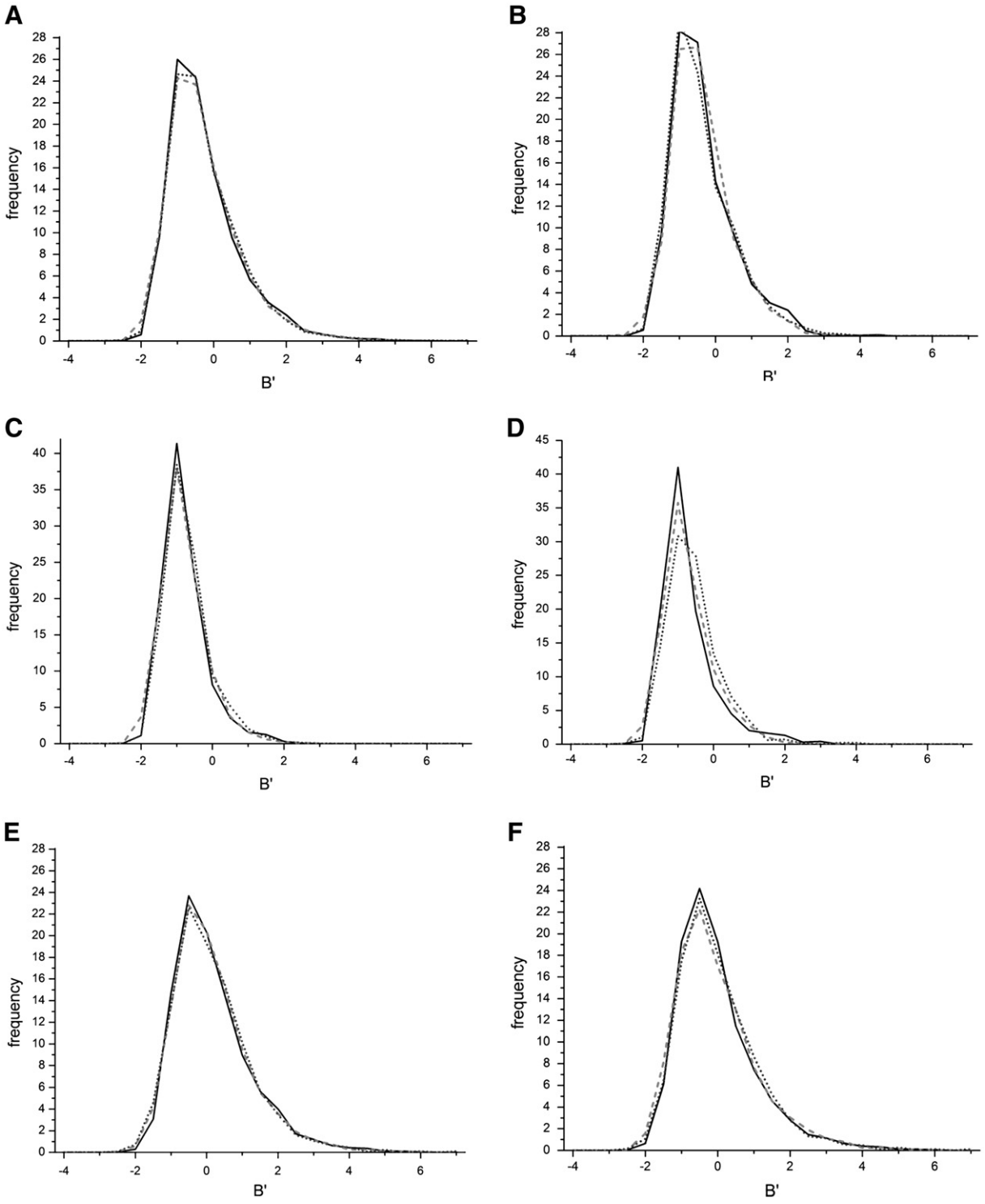

Fig. 1. Comparison of Cα B'-factor distributions. Overall distribution of B'-factors in the three temperature populations taken at 0.5 σ unit intervals in different structural environments. Solid line: psychrophiles; dashed line: mesophiles; dotted line: termophiles. Frequency values are multiplied by 100. The panels refer to A) global distribution; B) α-helix; C) coil; D): β-strand; E) buried 5%; F) exposed 16%.

psychrophiles than in their counterparts. The opposite if the difference were negative. The same analysis has been applied to the relative frequencies of Cα with B'-factors ≤ -1.5 to test whether frequency of "rigid" B'-factors were lower in psychrophiles than in their counterparts. Table 2 reports the p-values of the test under the two opposite hypotheses. Results indicate that psychrophilic β-strands are significantly richer than homologues in "flexible" Cα atoms (B' factor > 1.0) and buried regions are richer in "highly flexible" Cα (B'-factor > 1.5). On the contrary, psychrophilic coil regions appear depleted of flexible atoms. Also, psychrophilic proteins contain less rigid Cα in all the different structural environments except for β-turns. Differences of frequencies of B'-factors in the intervals considered do not show any significant or clear linear correlation to the pairwise percentage of identity or difference in adaptation temperature (Supplementary Table S1). This pattern suggests that the magnitude of the differences does not depend directly on pairwise

Table 2
t-test of the differences between the frequencies f_{ik} of Cα atoms with B'-factor ≥1.0 σ, ≥1.5 σ, and ≤−1.5 σ in psychrophiles and their homologous counterparts (mesophiles plus thermophiles) in different structural environments.

B'-factor range tested	α-helix	β-strand	Coil	Turn	Buried 5%	Exposed 16%
p-value under H_0: $\Delta f_{ik} \leq 0.0$[a]						
≥1.0	0.72	**0.02**	0.99	0.11	0.07	0.93
≥1.5	0.40	**6.83 10^{-4}**	0.99	0.12	**0.01**	0.57
≤−1.5	0.97	0.99	0.99	0.88	0.99	0.99
p-value under H_0: $\Delta f_{ik} \geq 0.0$[a]						
≥1.0	0.27	0.98	**1.8 10^{-4}**	0.88	0.92	0.07
≥1.5	0.60	0.99	**0.01**	0.88	0.98	0.42
≤−1.5	**0.03**	**0.00**	**0.01**	0.11	**0.00**	**0.01**

[a] Boldfaces *p*-values indicate significant values. Underlined numbers indicate a possible trend.

evolutionary or adaptation temperature distances although, globally, it is significant. Therefore, the magnitude itself is possibly related to the catalytic and structural peculiarities of every protein family and not simply to distances.

Overall $\Delta B'_k$ calculated in the each of the six structural environments considered do not show significant differences at 0.05 level except for Cα in turns (data not shown). Likewise, environment $\Delta B'_k$ do not show any significant or clear linear correlation to the percentage of identity or ΔT as well (Supplementary Table S2 and Fig. 2).

3.3. Cα B′-factors of each residue type

Average Cα B'-factor for each of the 20 residues in the psychrophiles has been compared with the corresponding residues in the homologous meso- and thermophiles. Statistical tests show that Ala, Cys, Gly, Ile, Met and Tyr have a Cα B'-factor higher and Leu and Pro lower in psychrophiles than in their homologues (Fig. 3). Amino acid compositions of the psychrophilic, mesophilic and thermophilic populations of the protein sample considered were also calculated (Supplementary Table S3).

3.4. Intra-family B′-factor profile comparison

Comparison of the distribution of Cα B'-factor averaged over a sliding window in the psychrophilic enzymes and their aligned homologues has been calculated only in the 5 families containing at least 4 mesophilic counterparts each: α-amylase, triosephosphate isomerase, subtilisin, D-glyceraldehyde-3-phosphate dehydrogenase and trypsin. This selection was made to avoid possible biases in the generation of reference mesophilic population during the testing procedure if smaller families were included. Moreover, only windows whose central residues are matched to at least four residues from the corresponding mesophilic homologues, were considered for the same reasons. For example, if the central residue of a window of a psychrophilic protein belonging to a family containing four homologues were aligned to three residues and one gap, then that window would be excluded from the calculation. The results reported in Table 3 suggest that there are significant differences in several positions within the alignment of each family. The differences can be greater or less than zero, indicating increasing or decreasing relative local flexibility in the psychrophilic structures with respect to the homologous mesophiles, respectively. In general, relatively "flexible" windows are more represented then "rigid" ones. Moreover, α-helices host more frequently rigid "windows", while β-sheets, turns and coil prefer "flexible" windows. This behavior is paralleled by the higher frequency of "flexible" regions in buried and exposed regions (Table 3). The frequency of amino acids in the pooled flexible and rigid windows has also been calculated. Gly, Lys and Tyr were the most frequent in the flexible windows while Phe, Leu and Thr were more represented in the "rigid" windows (data not shown). To test for the presence of correlation between the magnitude of significant $\Delta\overline{B}_i$ and the psychrophile–mesophile percentage of identity or temperature difference, Pearson correlation coefficients have been calculated. Results confirm that there is no significant linear correlation between the variables (Fig. 4 and Supplementary Fig. S1). Once again, the pattern suggests that the magnitude of the differences does not depend on pairwise evolutionary or adaptation temperature distances, although it is locally significant. Therefore, the magnitude itself is possibly related to the catalytic and structural peculiarities of every protein family and not simply to distances.

Moreover, comparison of the distribution of significant $\Delta\overline{B}_i$ calculated for the psychrophile–mesophile and mesophile–mesophile pairs show that the magnitude of the differences tend to be overall larger in the former than in the latter case.

4. Discussion

Cα B-factors are considered an important source of information about protein internal dynamics, providing a map of the static flexibility of the ground-state protein conformation [14]. For example, analysis of the distribution of normalized B-factors was utilized to characterize the differences in dynamics of thermophilic and mesophilic proteins [15]. Moreover, it was demonstrated that Cα protein backbone flexibility measured by B-factor profiles are conserved at family and superfamily levels, even for pairs of proteins with nonsignificant sequence similarity [16]. Recently it has been pointed out that the observed conservation is due to the most collective motions [28]. Further support to the relevance of thermal factors is that the correlation of B-factor profiles of homologous proteins shows they are correlated with each other with average correlation coefficient equal to 0.80 [17]. Moreover, the relation between thermal motions assessed according to crystallographic B-factors and motion predicted computationally using Normal Mode Analysis has been recently studied. It turned out that the frequency of residues with larger fluctuations in B-factor and Normal Mode Analysis are highly correlated although with a few significant discrepancies [29].

However, for complex molecules like proteins, B-factors can be highly variable within a single structure as a result of several effects, like local packing, structural environment of the atom, temperature of crystallization, *etc.* Because of these considerations, the B-factors in a protein must be normalized before meaningful comparisons among different protein chains can be made [30]. Normalization implies that only relative differences can be seen in the B'-factors distributions and no consideration can be made on the variations of absolute magnitude of the factors. Indeed, the present work was mainly aimed at detecting variations in the distribution of relative thermal factors between psychrophilic proteins and their homologues.

Distributions of Cα B'-factors were calculated in different structural environments for the three temperature population of proteins, namely psychrophilic, mesophilic and thermophilic enzymes. Differences of B'-factor distribution in the structural environments were calculated by accumulation of intra-family differences between the psychrophilic enzymes and their homologous sequences rather than by direct comparison of average values within temperature population. This strategy allows to maintain the characteristics of each family that otherwise would be lost during averaging. For example, an enzyme family may have its Cα B'-factor distribution shape intrinsically skewed toward high values and another family toward low values. The two trends would disappear after the calculation of an overall mean B'-factor in one of the structural environments of a temperature population. On the contrary, intra-family differences between the distribution of B'-factor of each psychrophilic and its homologous proteins, will retain

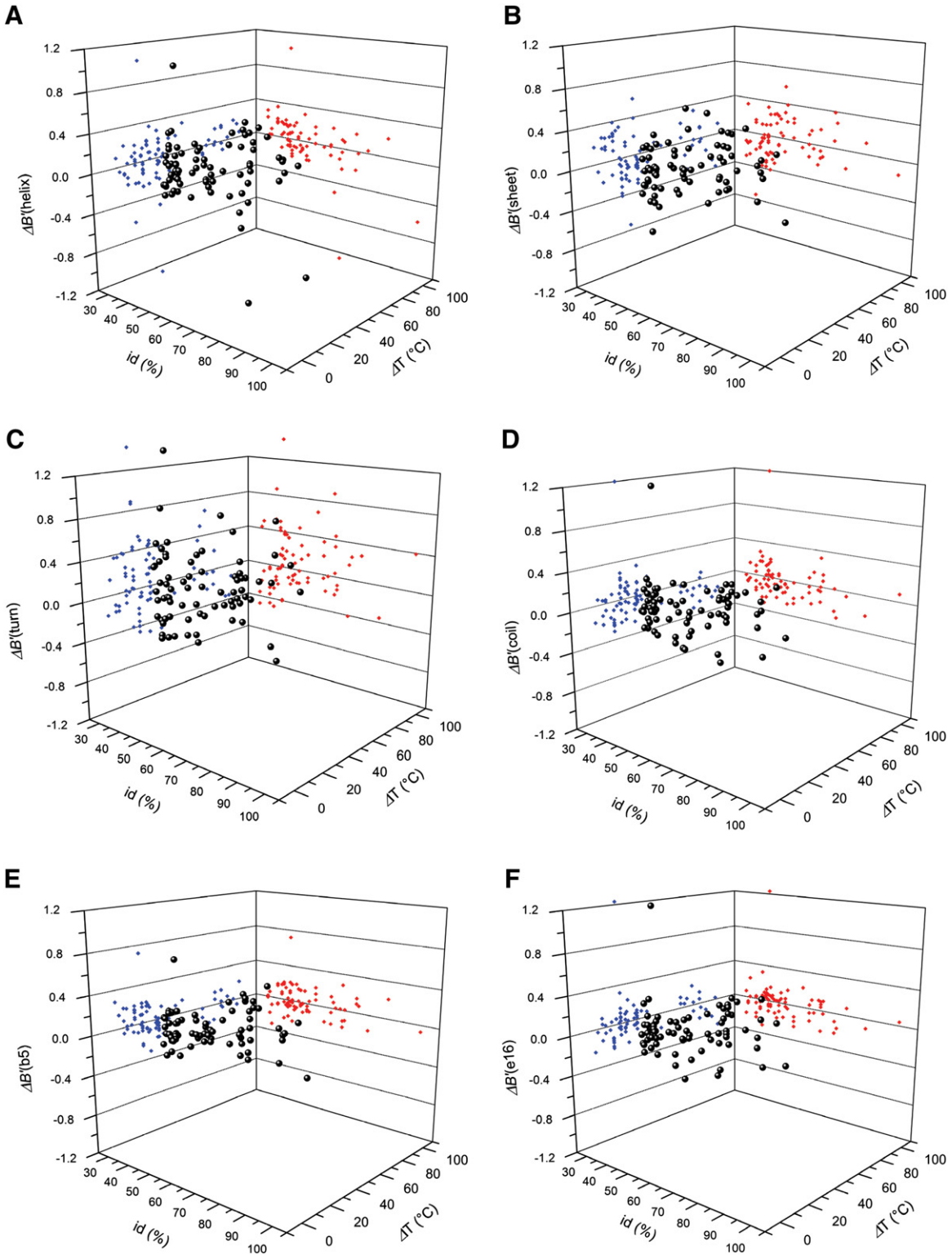

Fig. 2. Correlation between environment $\Delta B'_k$ and pairwise percentage identity and ΔT. Three dimensional graphs of the distributions of overall $\Delta B'_k$ in each structural environment versus pairwise percentage of identity (id%) and ΔT. The panels refer to A) α-helix, B) β-sheet, C) turn, D) coil, E) buried 5% and F) exposed 16%. Data projections on the $\Delta B'_k - id$ and $\Delta B'_k - \Delta T$ planes are reported as red and blue dots respectively.

possible significant trends. It should be also noted that comparison within each family improves filtering of the experimental and artifactual noise of the *B*-factor measurements due to crystal contacts, crystallization conditions, presence of ligands or salts and the like. However, possible links of the changes in frequency or average magnitude of $\Delta B'$ to sequence or molecular phenotype should be checked before generalization of results can be made. Our correlation analyses strongly suggests that the observed differences are not significantly correlated with the pairwise sequence distance (measured as percentage of identity) nor with the difference of optimal temperature in each individual comparison. Therefore, the magnitude of the observed differences is possibly related to the catalytic and

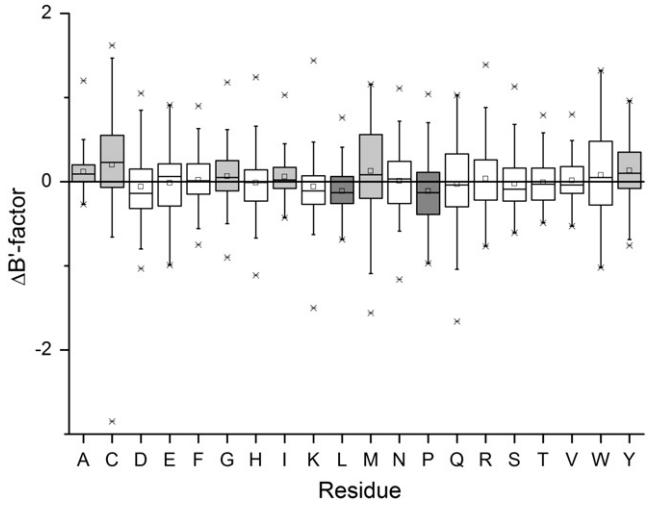

Fig. 3. Cα B'-factor differences in each residue type. Box plot of the distribution of the differences between the average B'-factors of the single residues in psychrophilic and meso- plus hyper/thermophilic proteins. Grey or dark grey boxes indicates residue with average B'-factor significantly greater or lower than expected, respectively, after a *t-test*.

structural peculiarities of every protein family and not to evolutionary or adaptation temperature distances in a simple way. This fact was indeed anticipated by previous studies [16,28] and supports the notion here proposed that the observed differences are tied to cold adaptation.

Overall, the picture emerging from our results suggests that the distribution of the flexibility is locally more heterogeneous in psychrophilic enzymes than in the respective homologues. Indeed, overall distribution shape of Cα B'-factors shows enrichment of flexible B'-factors in the region $+2.0\ \sigma$ and depletion of rigid B'-factors around the region $-2.0\ \sigma$ (Fig. 1). Analysis of the B' distributions in different environments indicate that the frequency of the Cα B'-factors greater than $+1.0$ (flexible) is significantly greater in the β-strands and buried residues of the psychrophilic proteins than in their counterparts (Table 2). Accordingly, Cα B'-factors less than -1.5 (more rigid) are significantly less frequent in all the structural environments of the psychrophilic proteins except for β-turns (Table 2). Likewise, average Cα B'-factor of each residue type display significant differences: in particular, Ala, Cys, Gly, Ile, Met and Tyr appear to be on average more flexible, while Leu and Pro are more rigid than in the mesophilic and thermophilic counterparts (Fig. 2). Among the most flexible residues, interesting is the presence of Gly, Met and Ile. Gly generally attributes local flexibility to the polypeptide chain where it occurs. The role of flexibility enhancer appears to be even greater in the psychrophilic enzymes. On the other hand, Ile and Met are frequently involved in reciprocal stabilizing interactions in the hydrophobic core of proteins mainly at β-strand conformations [31] which contribute to rigidify the entire structure. The increased flexibility of these two residues is reflected by the increased frequency of high B'-factors in β-strands Cα. Pro is frequently represented in the β-turns which tend to be more rigid. Differences in the amino acid composition calculated in our sample reflect some of the patterns described in literature [32]. In particular, Lys declines compared to mesophile and thermophile compositions, while Glu decreases only when compared to thermophiles. A moderate increase of Ala and Met is observed compared to the mesophilic counterparts while Cys, Gly, and again Met, expand when contrasted to thermophiles. As expected, Pro generally decreases with a consequent effect on flexibility. Besides these changes, it is interesting to mention the decrease of Arg content, a charged residue, which suggests a general decrease in the number of salt bridges, another possible strategy to enhance backbone flexibility. As for the hydrophobic residues, a reduction of Ile, Leu and Val content was observed relative to mesophiles and thermophiles, suggesting a decrease in hydrophobic interactions.

The comparison of the of average B'-factor profiles demonstrates that, along the psychrophilic sequence, there are a few flexible or rigid "cold spots" where the flexibility is significantly lower or greater than in the corresponding segments of mesophilic homologous proteins, respectively (Table 3) while the B'-profile is instead expected to be conserved in homologous structures, even at low sequence identity [17]. In general, psychrophiles in our data set tend to contain more frequently flexible than rigid cold spots (Table 3) in agreement with the variations observed in Fig. 1. Helices tend to contain more frequently rigid cold spots while sheets, coils and turns are richer in flexible cold spots. Likewise, buried and exposed positions contain more frequently flexible cold spots. However, cold spots occurring at the N- or C-terminal ends should not be considered because in these regions, residues may have markedly altered flexibility. Besides this, it has been suggested that the B-factors can be affected by the crystal contacts. However, it has been shown [16] that removal of Cα atoms in contacts within the crystal did not change significantly statistics on B'-factor distributions. Moreover, the pairwise intra-family comparison of the profiles contributes removing the potential artifacts, *i.e.* noise generated, for example, by the different crystallization conditions. It should be noted that the cold spots are spread among the different structural environments including buried positions (Table 3) that, by definition, cannot be involved directly in crystal contacts. Several data published in literature supports the idea that local flexibility of psychrophilic proteins is distributed more heterogeneously along the

Table 3
Number of psychrophilic windows that display an average Cα B'-factor significantly greater or lower than the corresponding mesophilic windows in each family structural alignment. Total number of windows is reported along with the distribution in the different structural environments. Overall sum of the "flexible" and "rigid" windows is displayed in the last two lines of the table.

Family	Δ[a]	No. of significant windows[b]	α-Helix[c]	β-Strand[c]	β-Turn[c]	Coil[c]	Buried 5%[c]	Exposed 16%[c]
α-Amylase	>	55	5	21	7	22	29	24
	<	26	15	1	2	8	9	14
Triosephosphate isomerase	>	40	14	1	9	16	11	24
	<	18	10	3	1	4	4	9
Subtilisin	>	15	9	0	1	5	5	8
	<	12	2	3	1	6	6	6
D-Gliceraldehyde-3-phosphate dehydrogenase	>	33	11	8	3	11	16	10
	<	31	15	11	4	1	12	14
Trypsin	>	6	0	0	2	4	2	4
	<	16	7	2	0	7	7	9
Sum	>	149	39	30	22	58	63	70
	<	103	49	20	8	26	38	52

[a] ">" means that average B'-factor of the psychrophilic window is significantly greater that of the corresponding mesophilic windows. "<" indicates the opposite.
[b] Total number of psychrophilic windows whose average B'-factor is significantly different from the mesophilic counterparts.
[c] Number of psychrophilic windows with significantly different B'-factor in each structural environment.

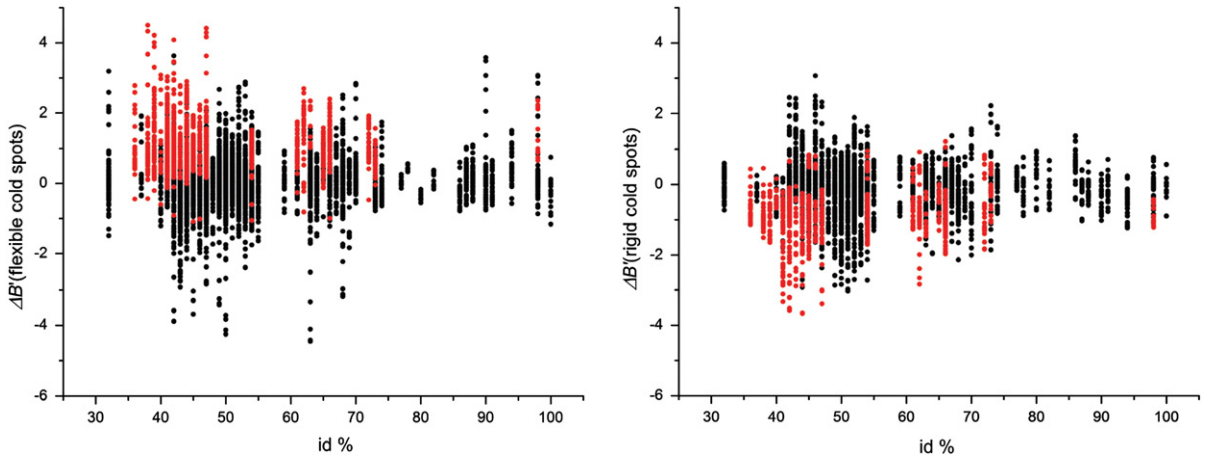

Fig. 4. Distributions of $\Delta\overline{B'}_i$ in flexible and rigid cold spots. Comparison between the distributions of $\Delta\overline{B'}_i$ in every cold spot calculated for the psychrophilic–mesophilic pairs (red dots) and the corresponding generated by the mesophilic–mesophilic pairs (black dots) versus pairwise percentage of identity (id%). Left and right graphs display distributions relative to flexible and rigid cold spots respectively.

polypeptide backbone compared to mesophilic or thermophilic homologues. For example, it should be mentioned that the presence of increased flexibility in small areas of the enzyme lactate dehydrogenase from Antarctic notothenoid fishes was described [33]. The authors proposed that such "hot spots in cold adaptation" allow a higher degree of flexibility in areas that move during catalysis and that an increase in flexibility in these areas should increase k_{cat} by reducing the energetic cost of conformational changes. Likewise, *Aquaspirillum arcticum* malate dehydrogenase has been compared to the structure of the thermophilic homologue from *Thermus flavus* [34]. Use of normalized *B*-factors indicated an increased relative flexibility at and near the active site region of the psychrophilic enzyme. Interestingly, similar conclusions have been reached on another system, the carbonic anhydrase [35] using a combination of experimental and theoretical approaches. In this case, a gain in flexibility in the part of the enzyme that controls the correct folding and an increased rigidity in the portion of the enzyme that anchors the catalytic β-strands into the hydrophobic core compared to the corresponding regions in the mesophilic homologous, were demonstrated [35]. Molecular dynamics simulations suggested that flexible segments are located in different positions in homologous psychrophilic and mesophilic α-amylases [36] with potential influence on catalytic activity. A flexible loop near the active site of a *Pseudoalteromonas haloplanktis* esterase which plays a crucial role in the protein function has also been identified [37]. However, these conclusions have been reached in most cases *via* binary comparison, *i.e.* one psychrophilic enzymes versus one mesophilic or thermophilic homologue. In these conditions, interpretation of results may be compounded by presence of alterations due to specific structural adaptation of the molecule utilized as a reference. Our approach, on the contrary, has the advantage of using at least four references homologues which decreases such potential distortions.

An inverse correlation between flexibility and packing has been recently demonstrated [38] which suggests that psychrophilic enzymes may reach local flexibility in their core regions by perturbation of the residue packing. It should be also mentioned that a relation between residue flexibility and solvent accessibility of adjacent neighbors [39] has been described. As an example of the comparison of *B*′-factor profiles, the difference profile of the α-amylase family is reported in Fig. 5 based on the structural alignment of Fig. 6. The plot displays the cold spots corresponding to the windows bearing a statistically significant deviation of their average *B*′-factor compared to that calculated over the mesophilic counterparts. Our results were compared to those obtained *via* molecular dynamics [36] on the same enzyme: two out of four loops indicated as more flexible in the psychrophilic amylase were identified by our comparison (Figs. 6 and 7). The other two loops were not taken into consideration by our profile comparison because were aligned to insertions and could not pass the filtering procedure we adopted to minimize statistical noise. However, our analysis indicated more cold spots in the α-amylase, represented in Fig. 7. Several of these cold flexible spots are in the surroundings of the active site in buried positions of β-barrel while the rigid cold spots are mainly localized in the small domain exposed to the solvent and therefore may be subject to artifactual interactions.

This work suggests that careful analysis of normalized experimental *B*-factors can provide valuable information about differences of the distribution of static flexibility in homologous proteins adapted at different temperatures. Moreover it suggests that intra-family comparative analysis of local *B*′-factors in psychrophilic enzymes can give valuable indication on their dynamic properties since it was able, at least in the reported case, to reproduce the results obtained by theoretical simulations using only experimental and validated crystallographic data and to indicate further interesting sites. In general, the results suggest that, even if each enzyme family achieve molecular adaptation at cold temperature through different

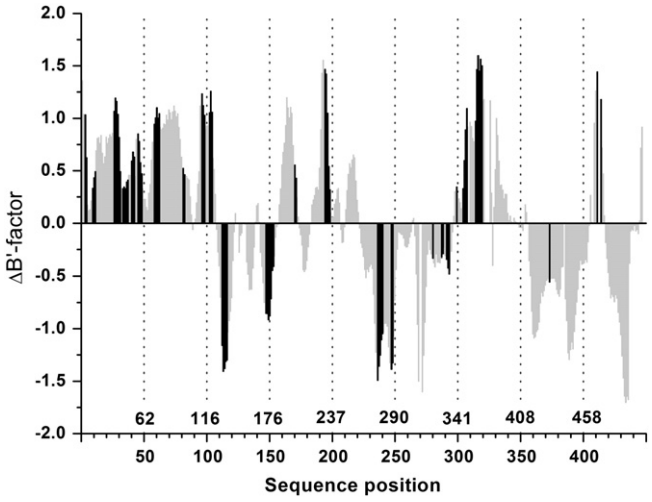

Fig. 5. Cα $\Delta\overline{B}_i'$-factor profile of the enzyme α-amylase (1AQM). Sequence positions at which variance of the distribution of $\Delta\overline{B}_i'$ between the psychrophilic and mesophilic windows differs significantly (according to the Bartlett's test) from the variance of the distribution of $\Delta\overline{B}_i'$ among the mesophilic counterparts, are indicated as black bars. Numbering system above the horizontal bottom axis refers to the numbering above sequences in Fig. 6. Tick labels below axis refer to the sequence itself.

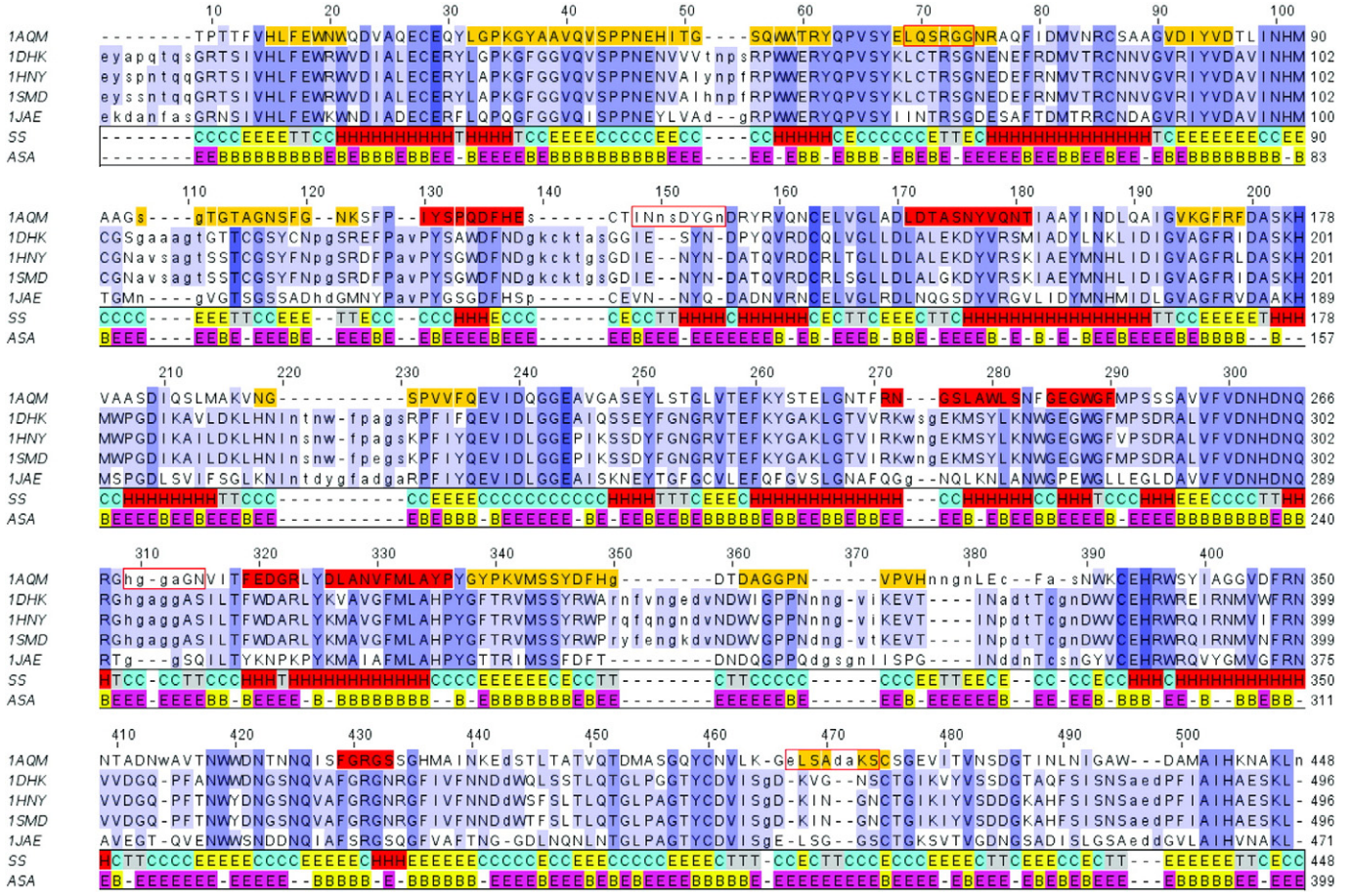

Fig. 6. Multiple structural alignment of the members of the α-amylase family. Secondary structure and accessibility of the psychrophilic sequence are reported in the lines labeled "SS" and "ASA" respectively. Secondary structure is indicated with the following letters: H is α-helix; E β-strand; T turn; C coil. Relative solvent accessibility is reported as B: ≤5%; E ≥16%; blank: between the two ranges; dot: deletion. Each protein is indicated by the corresponding PDB code. Letters with orange or red background on the psychrophilic sequence (1AQM) indicate the window significantly more flexible or more rigid, respectively, than the corresponding mesophilic homologues. Red boxes mark the regions reported to be more flexible [24]. Lowercase amino acid symbols indicate the non-structurally equivalent regions.

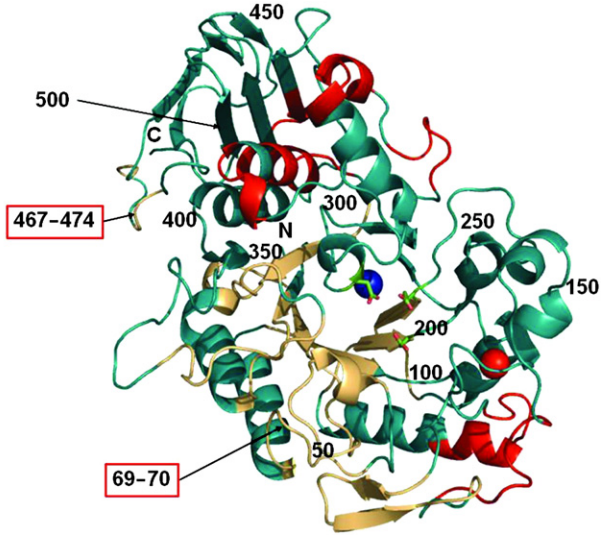

Fig. 7. Cold spots in the psychrophilic α-amylase (1AQM). Ribbon representation of the structure of psychrophilic α-amylase. Orange and red colors indicate respectively the windows significantly more flexible and more rigid than the mesophilic counterparts. Red boxes label the loops identified by the molecular dynamics analysis [24] and by our analysis. Numbering system refers to the numbering above sequences in Fig. 6. The blue and red spheres represent the chloride and the calcium ion, respectively. Active site residues are represented as stick models.

evolutionary strategies [6], a common heterogeneousness in the distribution of *B′*-factors along the structure can be observed in psychrophilic enzymes compared to their homologous mesophilic proteins. Within the distribution of *B′*-factors there are a few regions whose average relative flexibility is significantly different (greater or smaller) than the corresponding ones in the homologous mesophiles. In general, flexible cold spots are more frequent in the psychrophilic enzymes and they tend to occur mainly in strands, turns and coils as well as in buried and exposed positions. Helices seem to have more frequently rigid cold spots. The presence of these regions in the psychrophilic enzymes, although relatively far from the active site, can modify the overall dynamic behavior of the enzyme and make it able to work efficiently at low temperatures. This work was deliberately focused onto enzymes to study the molecular adaptations related to the presence of catalytic function. In this way an homogeneous sample could be collected. Indeed, other non-enzyme proteins are presumably subject to dissimilar structural restrains and therefore may undergo different adaptation mechanisms. In perspective, understanding of the characteristics of the "cold spots" and their effects on the protein structure dynamics will open the possibility to rationally engineering enzymes of biotechnological relevance.

Acknowledgements

This work has been partially supported by the funds from the Italian "Ministero dell'Istruzione, dell'Università e della Ricerca". This work will be submitted by A.S. in partial fulfillment of the

requirements of the degree of "Dottorato di Ricerca in Biochimica" at Sapienza, Università di Roma.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.bpc.2010.10.009.

References

[1] S. D'Amico, C. Gerday, G. Feller, Structural determinants of cold adaptation and stability in a large protein, J. Biol. Chem. 276 (2001) 25791–25796.
[2] S. D'Amico, J.C. Marx, C. Gerday, G. Feller, Activity–stability relationships in extremophilic enzymes, J. Biol. Chem. 278 (2003) 7891–7896.
[3] G. Feller, C. Gerday, Psychrophilic enzymes: hot topics in cold adaptation, Nat. Rev. Microbiol. 1 (2003) 200–208.
[4] K.S. Siddiqui, R. Cavicchioli, Cold-adapted enzymes, Annu. Rev. Biochem. 75 (2006) 403–433.
[5] A.O. Smalås, H.K. Leiros, V. Os, N.P. Willassen, Cold adapted enzymes, Biotechnol. Annu. Rev. 6 (2000) 1–57.
[6] G. Gianese, F. Bossa, S. Pascarella, Comparative structural analysis of psychrophilic and meso- and thermophilic enzymes, Proteins 47 (2002) 236–249.
[7] D. Tronelli, E. Maugini, F. Bossa, S. Pascarella, Structural adaptation to low temperatures—analysis of the subunit interface of oligomeric psychrophilic enzymes, FEBS J. 274 (2007) 4595–4608.
[8] S. Thorvaldsen, E. Hjerde, C. Fenton, N.P. Willassen, Molecular characterization of cold adaptation based on ortholog protein sequences from Vibrionaceae species, Extremophiles 11 (2007) 719–732.
[9] C.J. Marshall, Cold-adapted enzymes, Trends Biotechnol. 15 (1997) 359–364.
[10] N.J. Russell, Molecular adaptations in psychrophilic bacteria: potential for biotechnological applications, Adv. Biochem. Eng. Biotechnol. 61 (1998) 1–21.
[11] R. Cavicchioli, K.S. Siddiqui, D. Andrews, K.R. Sowers, Low-temperature extremophiles and their applications, Curr. Opin. Biotechnol. 13 (2002) 253–261.
[12] C. Vieille, D.S. Burdette, J.G. Zeikus, Thermozymes, Biotechnol. Annu. Rev. 2 (1996) 1–83.
[13] T. Lonhienne, C. Gerday, G. Feller, Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility, Biochim. Biophys. Acta 1543 (2000) 1–10.
[14] D. Ringe, G.A. Petsko, Mapping protein dynamics by X-ray diffraction, Prog. Biophys. Mol. Biol. 45 (1985) 197–235.
[15] S. Parthasarathy, M.R. Murthy, Protein thermal stability: insights from atomic displacement parameters (B values), Protein Eng. 13 (2000) 9–13.
[16] S. Maguid, S. Fernández-Alberti, G. Parisi, J. Echave, Evolutionary conservation of protein backbone flexibility, J. Mol. Evol. 63 (2006) 448–457.
[17] Z. Yuan, T.L. Bailey, R.D. Teasdale, Prediction protein B factor profiles, Proteins 58 (2005) 905–912.
[18] J. Gao, T. Zhang, H. Zhang, S. Shen, J. Ruan, L. Kurgan, Accurate prediction of protein folding rates from sequence and sequence-derived residue flexibility and solvent accessibility, Proteins 78 (2010) 2114–2130.
[19] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.
[20] A. Chang, M. Scheer, A. Grote, I. Schomburg, D. Schomburg, BRENDA, AMENDA and FRENDA the enzyme information system: new content and tools in 2009, Nucleic Acids Res. 37 (2009) D588–D592.
[21] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[22] B. Rost, C. Sander, Conservation and prediction of solvent accessibility in protein families, Proteins 20 (1994) 216–226.
[23] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[24] S.J. Hubbard, J.M. Thornton, NACCESS Version 2.1.1, Dept of Biochemistry and Molecular Biology, University College, London, 1993.
[25] E. Krissinel, K. Henrick, Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2256–2268.
[26] M.S. Bartlett, Properties of sufficiency and statistical tests, Proc. R. Stat. Soc. A 160 (1937) 268–282.
[27] W.L. Delano, The PyMOL molecular graphics system version 1.2, DeLano Scientific, San Carlos, CA, USA, 2002.
[28] S. Maguid, S. Fernandez-Alberti, J. Echave, Evolutionary conservation of protein vibrational dynamics, Gene 422 (2008) 7–13.
[29] S. Hirose, K. Yokota, Y. Kuroda, H. Wako, S. Endo, S. Kanai, T. Noguchi, Prediction of protein motions from amino acid sequence and its application to protein–protein interaction, BMC Struct. Biol. 10 (2010) 20.
[30] D.K. Smith, P. Radivojac, Z. Obradovic, A.K. Dunker, G. Zhu, Improved amino acid flexibility parameters, Protein Sci. 12 (2003) 1060–1072.
[31] Y. Isogai, M. Ota, A. Ishii, M. Ishida, K. Nishikawa, Identification of amino acids involved in protein structural uniqueness: implication for de novo protein design, Protein Eng. 15 (2002) 555–560.
[32] R.A. Goldstein, Amino-acid interactions in psychrophiles, mesophiles, thermophiles, and hyperthermophiles: insights from the quasi-chemical approximation, Protein Sci. 16 (2007) 1887–1895.
[33] 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. USA 95 (1998) 11476–11481.
[34] S.Y. Kim, K.Y. Hwang, S.H. Kim, H.C. Sung, Y.S. Han, Y. Cho, Structural basis for cold adaptation. Sequence, biochemical properties, and crystal structure of malate dehydrogenase from a psychrophile Aquaspirillium arcticum, J. Biol. Chem. 274 (1999) 11761–11767.
[35] R. Chiuri, G. Maiorano, A. Rizzello, L.L. del Mercato, R. Cingolani, R. Rinaldi, M. Maffia, P.P. Pompa, Exploring local flexibility/rigidity in psychrophilic and mesophilic carbonic anhydrases, Biophys. J. 96 (2009) 1586–1596.
[36] M. Pasi, L. Riccardi, P. Fantucci, L. De Gioia, E. Papaleo, Dynamic properties of a psychrophilic alpha-amylase in comparison with a mesophilic homologue, J. Phys. Chem. B 113 (2009) 13585–13595.
[37] V. Aurilia, J.F. Rioux-Dubé, A. Marabotti, M. Pézolet, S. D'Auria, Structure and dynamics of cold-adapted enzymes as investigated by FT-IR spectroscopy and MD. The case of an esterase from *Pseudoalteromonas haloplanktis*, J. Phys. Chem. B 113 (2009) 7753–7761.
[38] B. Halle, Flexibility and packing in proteins, Proc. Natl Acad. Sci. USA 99 (2002) 1274–1279.
[39] H. Zhang, T. Zhang, K. Chen, S. Shen, J. Ruan, L. Kurgan, On the relation between residue flexibility and local solvent accessibility in proteins, Proteins 76 (2009) 617–636.